

microRNAs as metabolic sensors and engineering tools in CHO cells



A dissertation submitted for the degree of Ph.D.

by

Ricardo Valdés-Bango Curell, M.Sc.

The work here described was conducted under the supervision of

Dr. Finbarr O'Sullivan

Prof. Nicole Borth (BOKU)

Prof. Niall Barron (University College Dublin)

National Institute for Cellular Biotechnology

School of Biotechnology

Dublin City University

September 2020

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:



Ricardo Valdés-Bango Curell

ID No.: 15212307

Date: 10th September 2020

TABLE OF CONTENTS

Declaration	i
TABLE OF CONTENTS	iii
ACKNOWLEDGEMENTS	vii
ABBREVIATIONS	xi
ABSTRACT	1
AIMS OF THE THESIS	3
CHAPTER 1 Introduction	5
1. Introduction	7
1.1. Chinese Hamster Ovary Cells – Origins and background	7
1.2. CHO cells as host for biopharmaceutical production	8
1.3. CHO cell engineering: successes and challenges	12
1.4. miRNA-based CHO cell engineering – from targets to tools	15
2. miRNAs and siRNA: biogenesis, regulation and function	17
2.1. miRNA biogenesis	17
2.2. Regulation of miRNA biogenesis	19
2.3. Target Recognition and mechanisms of action.	21
3. Exploiting endogenous miRNA signatures to modulate transgene expression	22
4. Synthetic miRNAs to modulate transgene expression	26
4.1. Inducible expression of synthetic miRNA/siRNAs	26
4.2. Controlling miRNA expression by affecting their processing	27
5. Novel miRNA responsive elements based on new modes of action	28
5.1. miRNA regulating RNA-binding protein motifs	28
5.2. miRNAs to improve translation efficiency	29
6. miRNAs for refined genetic control of transgene expression in CHO cells	30
CHAPTER 2 Materials and Methods	33
1. Molecular Biology Techniques	35
1.1. Nucleic Acid Amplification	35
1.2. Nucleic Acid Isolation and Purification	37
1.3. Nucleic Acid Quantification	38
1.4. Nucleic Acid Quality Control	43
1.5. Nucleic Acid Modification	43
2. Cell Culture Techniques	47
2.1. General Cell Culture Good Laboratory Practices	47
2.2. Cell Line Sub-Culturing	47
2.3. Cell Density and Viability Monitoring	48
2.4. Cryopreserved Cell Lines Revival	49
2.5. Cell Line Cryopreservation	49
2.6. Transient Transfection	50
2.7. Stable Cell Line Generation	51
2.8. GFP detection and quantification	51
2.9. RNA decay assay – Actinomycin D treatment	51
3. In silico Tools	52
3.1. Nucleic Acid Secondary Structure Prediction	52
3.2. Plasmid Cloning/ Maps	52

3.3.	Differential Analysis of miRNA expression	52
3.4.	Unsupervised hierarchical clustering	53
3.5.	miRNA Target Prediction and Pathway Analysis Tools	53

CHAPTER 3 Identification of endogenous miRNA as potential gene switches in CHO cells

1.	Introduction	57
2.	Results	58
2.1.	Glutamine metabolism as guiding a differential factor for miRNA profiling	58
2.2.	RNASeq microRNA profiling reveals expression patterns correlated with cell culture stage	75
3.	Discussion	80
3.1.	Adapting cells to glutamine free conditions – one step is enough	80
3.2.	Glutamine metabolism and culture time as a guiding differential factors for miRNA profiling	82
3.3.	Hsa-miR-1275 and cgr-miR-1195 as ammonia sensitive miRNAs	83
3.4.	miRNAs correlated with ammonia and/or culture time are linked to amino acid and lipid catabolism as well as to membrane-nucleus phosphorylation signalling pathways.	84
3.5.	Towards CHO-specific and robust miRNA profiling	88
4.	Conclusion	89

CHAPTER 4 Temperature-induced control of transgene expression using miRNA sponges

1.	Introduction	93
1.1.	Temperature shift as a way to control transgene expression	93
1.2.	miRNA-sponges as molecular tools for transgene expression modulation	93
2.	Results	94
2.1.	Construction of miRNA-sponges targeting endogenous CHO miRNAs	94
2.2.	Linking transgene expression and miRNA expression using miRNA sponges	96
2.3.	Transient expression of miRNA-sponge sensors - transfection optimisation	99
2.4.	Effect of binding site number and DNA amount on transgene expression and sponge/sensor performance	104
2.5.	Effect of binding site number and complementarity on sponge/sensor performance with changing concentrations of miRNA	108
2.6.	RNA-Seq profiling of CHO cells exposed with mild-hypothermia reveals temperature sensitive microRNAs	111
2.7.	Independent temperature-shift experiments to validate the identified temperature sensitive miRNAs	112
2.8.	Validation of identified temperature sensitive miRNAs by RT-qPCR	114
2.9.	Reducing promoter-related variation upon temperature shift to allow evaluation of miRNAs for transgene expression control	115
2.10.	CMV and EF1a as drivers for miRNA sponge sensors	118
2.11.	EF1a-driven miRNA sponges for temperature sensitive mRNAs	120
2.12.	Optimisation of EF1a-driven miRNA sponge expression under temperature shift to assess miRNA-driven changes in GFP expression	122
2.13.	Evaluation of transiently transfected EF1a-driven miRNA sponges: effects on GFP expression	132
2.14.	Transient evaluation of miRNA-dependent changes in GFP expression in response to temperature shift	134
2.15.	Evaluation of EF1a-driven miRNA sponges in stable mixed pools: effects of miRNA sponge on GFP expression	139
2.16.	Evaluation of temperature shift on miRNA-driven changes in GFP expression using stable mixed pools expressing miRNA-sponges	142
3.	Discussion	149

4. Conclusion	158
CHAPTER 5 miRNA-toehold switches as novel tools to modulate transgene expression	161
1. Introduction	163
2. Results	165
2.1. Design and evaluation of new miRNA-based switches: miRNA-toehold switches	165
2.2. miRNA-Toehold-Switch <i>in silico</i> characterisation	168
2.3. Characterisation of miRNA-toehold switch vectors	171
2.4. Characterisation of miR-toehold switch driven GFP expression in stable pools.	174
2.5. Adaptation to miRNA-toehold stable pools to suspension growth	178
2.6. Evaluation of miR-toehold constructs in stable mixed pools	179
2.7. <i>In silico</i> analysis and transient expression of miR-toehold variants library	189
2.8. Mutations in the miRNA binding site introducing bulges to interfere with miRNA binding/ RISC-mediated degradation	191
2.9. Position of Kozak-start codon region in the hairpin and introduction of a stop codon before the GFP ORF Kozak-start codon	198
2.10. Mutations on the 3'-end of fragment C, corresponding the complementary/non-complementary region	205
2.11. Change of the T rich region from an A rich region in the loop	213
2.12. Are changes in the 5-UTR affecting mRNA, thus decreasing GFP protein expression?	217
3. Discussion	218
4. Conclusion	222
CHAPTER 6 Concluding Remarks and Future Work	225
BIBLIOGRAPHY	233
APPENDICES	I
APPENDIX A – SUPPLEMENTARY MATERIAL FOR CHAPTER 3	I
APPENDIX B – SUPPLEMENTARY MATERIAL FOR CHAPTER 4	IX

ACKNOWLEDGEMENTS

First and foremost I would like to thank **Prof. Niall Barron**. This work would not have been possible without you providing me with the opportunity to be part of your team, and all your support along the way. You have always been there when needed (even after relocating!) and your door has always been open for the often long chats over results and other matters.

Secondly, infinite thanks go to my beloved wife, **Cristiane**, for your invaluable love and support over these years. Your patience, generosity, and permanent smile, especially during the last months of endless hours in the office, have been the fuel that has kept me going all this time. There could not be a better partner for this trip, and the many more to come!

My third pillar over these years have been my NICB *et al.* friends, **Antonio, Berta, Giuseppe, Julie, Krishna, Jesus, Paloma, Peter, Prashant, and Teresa**. Many thanks for our very numerous coffee breaks, laughs, trips, long long chats about data and life in general, and our sometimes intense arguments over lunch. These few lines are only enough to say you are like family now and a bit of this work belongs to you all (you know I'd need much more and could keep going forever!). Thanks as well to the NICB crew, in particular Nga and Srinivas, for their advice and teachings over the years.

Many thanks as well to the **eCHO family: Ankita, Cyrielle, Davide, Linas, Heena, Ly, Nicolas, Nusa, Sara, Théo, Thomas, and Valerie**. Each of you made a very special contribution to my Ph.D. experience and meeting you has been one of the best things of all. Thanks for all the fun-time spent together as well as knowledge shared. Also many thanks to the whole eCHO systems consortium, **PIs and Scientific Advisory Board** alike, for providing an unforgettable and wonderful environment.

I would also like to gratefully acknowledge the funding that made my research possible provided by the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement No 642663 (eCHO systems).

A special mention goes to the members of my Examination Committee, **Dr. Donal O'Gorman and Prof. Mark Smales**, for your time examining my work and your guidance in polishing the final thesis, as well as to **Dr. Rosaleen Devery**, Chair, for undertaking the coordination efforts that have made it possible to achieve graduation in these challenging COVID19 times. Thanks as well to **Dr. Finbarr O'Sullivan** for your support with all the administrative work involved during the last year.

Last but not least, many thanks as well to **my family and friends, especially my parents**, for their continued love and encouragement along the way. Special mention to **my grandparents, Marcel-lí and Pilar**, for their continuous support to study Biotechnology

The Optimist

*"Life handed him a lemon,
As Life sometimes will do.
His friends looked on in pity,
Assuming he was through.
They came upon him later,
Reclining in the shade
In calm contentment, drinking
A glass of lemonade."*

Flynn, Clarence Edwin (November 1940), *The Rotarian*, 57, p. 62

ABBREVIATIONS

- AGO2**, Argonaute II protein;
- ASO**, Antisense Oligonucleotides;
- Cas**, CRISPR-associated;
- CHO**, Chinese Hamster Ovary;
- CRISPR**, Clustered Regularly Interspaced Short Palindromic Repeats
- dsRNA**, double-stranded RNA;
- EPO**, Erythropoietin;
- GFP**, Green Fluorescent Protein;
- HEK**, human embryonic kidney;
- hPSCs**, human Pluripotent Stem Cells;
- HSL**, histone stem-loop;
- iPSCs**, induced pluripotent stem cells;
- IRE**, iron responsive element;
- KO**, knock-out;
- lncRNA**, long non-coding RNA;
- miRISC**, microRNA-induced silencing complex;
- miRNA**, microRNA;
- MMTV-LTR**, mouse mammary tumour virus. long terminal repeat;
- MRE**, microRNA responsive element;
- pORF**, primary Open Reading Frame;
- pre-miRNA**, precursor microRNA;
- pri-miRNA**, primary microRNA;
- RBP**, RNA binding protein;
- RepO**, repressor-dependent operon/promoter;

RNAi, RNA interference;

rtTA, reverse tetracycline-controlled transactivator;

shRNA, short-hairpin RNA;

siRISC, siRNA-induced silencing complex;

siRNA, short interfering RNA;

SLBP, Stem-Loop Binding Protein;

sncRNA, short non-coding RNA;

sxRNA, structural interacting RNA;

TINCR, Tissue Differentiation-Inducing Non-Protein Coding RNA;

TRE, transactivator-responsive element;

uORF, upstream Open Reading Frame;

UTR, untranslated region;

SEAP, Secreted embryonic alkaline phosphatase;

VEGF, Vascular endothelial growth factor;

ABSTRACT

Improvements in the production capabilities of Chinese Hamster Ovary (CHO) cells have relied on traditional genetic engineering strategies, such as gene overexpression and gene knockouts. However, new strategies are likely to require more sophisticated rational approaches and novel molecular tools need to be developed to facilitate more refined control of gene expression. In this thesis, the use of microRNAs to harness control of transgene expression in CHO cells has been investigated. The first part of this thesis aimed to identify miRNA expression profiles that could be used to actuate transgene expression in the context of biopharmaceutical production. miRNA expression data from cell lines with different glutamine requirements was investigated in an attempt to identify glutamine responsive miRNAs. In addition, the analysis of a novel CHO miRNA expression dataset from a fed batch process resulted in the identification of interesting miRNA clusters exhibiting expression profiles that could be matched to particular growth phases. The second part of the work involved the investigation of temperature-induced miRNA expression changes in order to build a temperature dependent transgene expression control system using miRNA sponges. Temperature responsive miRNAs were identified and validated. While providing evidence that miRNA sponges can be used as molecular sensors and modulate gene expression, our results indicate that temperature-driven changes of miRNA expression are unlikely to be used as a robust gene control system. Finally, transgene expression control by combining UTR secondary structure and miRNAs was investigated. We showed that miRNA-toehold switches are able to repress transgene expression repression in a sequence specific manner although a robust ON/OFF function could not be achieved. Using a small library of synthetic 5'-UTR, the effect of several structural and sequence features in the miRNA-toehold was also investigated. Sequence determinants such as upstream ORFs, AU-rich regions and kozak environment showed greater effects on transgene expression than the 5'-UTR local secondary structure features.

In summary, the work described in this thesis represents the first attempt to implement miRNA-based strategies to directly control transgene expression in CHO cells while highlighting the difficulties that need to be overcome for this strategy to be effective.

AIMS OF THE THESIS

The aims of the work described in this thesis were:

- to identify endogenous miRNA whose expression profiles are indicative of the cellular metabolic state, or that can be correlated to specific culture conditions, using pre-existing or de novo generated miRNA expression datasets in Chinese Hamster Ovary cells as well as to investigate the biological functions of these to anticipate potential side effects of using them as switches for controlled transgene expression
- to investigate the use of miRNA sponges, namely synthetic 3'-UTRs containing specific miRNA binding sites as molecular tools to modulate transgene expression using synthetic and endogenous miRNA expression changes in Chinese Hamster Ovary cells
- to investigate the use of novel miRNA responsive elements in expression vectors to achieve miRNA-based modulation of transgene expression using endogenous miRNAs in Chinese Hamster Ovary cells

CHAPTER 1

Introduction

This introduction is an expanded version of the literature review:

Exploring the potential application of miRNA-based genetic circuits in CHO
Valdés-Bango Curell and Barron (2018)

published in Biotechnology Journal (DOI: 10.1002/biot.201700220)

1. Introduction

1.1. Chinese Hamster Ovary Cells – Origins and background

CHO cells grew in popularity between 1957 to the late 1970s for genetic studies, due to their practicality in cell culture and the possibility to generate auxotroph mutants by induced or spontaneous mutations and isolate them using selective media (Puck and Kao, 1982; Wurm, 2013). Chinese Hamster Ovary (CHO) cells were originally derived from ovary tissue from a female Chinese Hamster by Dr. Theodore Puck and collaborators in 1957 (Puck and Fisher, 1956). During the following decade, from an original CHO (CHO-ori), Proline auxotrophic CHO clones were isolated and characterized in vitro (Kao and Puck 1967). During the following years, cells were distributed to various labs around the world, and different cell lines were established, including some that have made it and are widely used today. On the one side, the CHO-K1 lineage, originally described by Kao and Puck and deposited in the ATCC with the CHO-K1 CCL 61 (Kao and Puck, 1967, 1968). From these, the CHO-DXB11 cell line, a thymidine and hypoxanthine auxotroph cell line, was generated by the introduction of a missense mutation by gamma ray induced mutagenesis was derived (Urlaub and Chasin, 1980; Wurm, 2013) in one of the DHFR alleles, and the deletion of the second one. CHO-K1 cells popularity is also based upon the fact that these cells were used in combination with the Glutamine Synthetase selection system (GS-system) (described later on).

On the other side, another subline derived from the original Puck CHO-ori isolate travelled to Los Alamos and was used to generate a CHO cell line adapted to suspension (CHO-S) (Gottesman, 1987; Wurm, 2013). There is a less clear timeline for the development of CHO-S cells, but they can be traced back to Toronto in 1973, when Thompson *et al.* described them already (Thompson and Baker, 1973). These suspension-adapted cells are not to be confused with other CHO cell lines that were later adapted to suspension, including CHO-K1 and CHO-DXB11 (Wurm, 2013).

The Toronto clones generated by Thompson and coworkers were distributed to different labs, and in particular, in the hands of Dr. Siminovitch and Dr. Flintoff, gave rise to a mutant CHO-Mtx^{RIII} that turned to be a useful candidate for further deletion of both DHFR alleles (Flintoff, Davidson and Siminovitch, 1976; Flintoff, Spindler and Siminovitch, 1976). Urlaub and Chasin achieved that in 1983 becoming the fathers of the now popular CHO-DG44 cells (Urlaub and Chasin, 1980; Urlaub *et al.*, 1983).

1.2. CHO cells as host for biopharmaceutical production

The term biopharmaceutical was coined in the 1980s specifically to describe recombinant protein products generated or produced by modern molecular biological methods and to distinguish them from the traditional biological products (Walsh, 2014).

Recombinant proteins can be expressed in a variety of host systems. Bacteria, such as *E. coli*, yeast, such as *Saccharomyces cerevisiae*, and fungi, have been used to express small proteins such as hormones like insulin or human growth hormone, or enzymes such as the β -D-galactosidase, and present great advantages due to fast process timelines and robustness, and low costs. However, they lack the biological ability to express complex glycoproteins with human-like post-translational modifications. Precisely, these complex drugs, such as antibodies or blood factors, have dominated the market over the last decade and are mainly being produced in mammalian cell factories (Walsh, 2014, 2018). While human cell lines, the obvious choice for human-like protein production, have been investigated, Chinese Hamster Ovary cells have been the preferred host for the production of approved therapeutic recombinant proteins over the last 30 years (Walsh, 2014; Lalonde and Durocher, 2017). CHO cells are not necessarily the most adapted cell types to produce large amounts of protein one can find. Indeed, one may wonder why there are no other cell types, such as B cells or other secretory cell types, that have been developed as cell factories. However, CHO cells comprise several phenotypical characteristics that made them thrive in the bioprocess environment. Both, the fact that CHO cells comprise many of the desirable characteristics for a heterologous expression system and their long and proof checked track record explain this success.

First, they are able to cope with genetic manipulation with an outstanding simplicity and flexibility, allowing not only the introduction of a foreign genetic material and the production of a recombinant protein, but also the possibility to select clones that have stably integrated the gene, giving rise to producer cell lines (Kim, Kim and Lee, 2012; Lalonde and Durocher, 2017). The possibility to apply amplification strategies to improve protein expression has also been shown (Lalonde and Durocher, 2017). A second attribute that made CHO cells a suitable host for the development of bioprocesses was their resistance to viral infection, in particular from human viruses. This was important and very much considered in the early years in order to guarantee

the safety of the produced recombinant product. In addition, as already mentioned above, CHO cells are also capable of performing human like post-transcriptional modification on the proceed recombinant proteins, an essential attribute for their proper folding and their in vivo therapeutic activity (Lalonde and Durocher, 2017). Moreover, this capacity has been further increase by cell engineering over the past decades, achieving the expression of fully humanized recombinant proteins (Amann *et al.*, 2019). Last but not least, the possibility to grow CHO cells in suspension and at large scales (>10000 L) makes them a very good host for the development of commercially viable bioprocesses (Chu and Robinson, 2001).

Another aspect to take into account to understand why CHO cells are widely used is their long and successful regulatory track record. Since the approval of the recombinant tissue plasminogen activator (r-tPA) in 1987, many others have chosen to use CHO cells as they are seen as robust and safe hosts, making it easier to commercialise products expressed from them, than to go through a whole new approval process for a new host (Walsh, 2014).

Since then, a wide variety of approved therapeutic recombinant products or biopharmaceuticals have been produced in CHO cells. These include therapeutic enzymes, hormones, cytokines, blood factor, monoclonal antibodies and fusion proteins of some of these (Walsh and Jefferis, 2006; Zhu, Mollet and Hubert, 2007; Walsh, 2010, 2014, 2018) (**Table 1**).

Table 1 – Approved Biopharmaceutical products produced in CHO cells in the US or the EU (compiled from Walsh, 2006-2018 and Zhu, et al 2007)

Product	Company	Year	Product type
rh factor VIII	Baxter Healthcare	1992	Blood factor
rh factor VIII	Aventis Behring	1993	Blood factor
Human factor IX	Wyeth	1997	Blood factor
Antihemophilic factor VIII	Wyeth	2000	Blood factor
thrombin-rh factor IIa	ZymoGenetics	2008	Blood factor
rh coagulation factor IX	Aptevo BioTherapeutics	2015	Blood factor
von Willebrand factor (recombinant)	Baxalta	2015	Blood factor
rh coagulation factor IX	Novo Nordisk	2016	Blood factor

recombinant coagulation factor IX)	Novo Nordisk	2017	Blood factor
rh coagulation factor IX	Novo Nordisk	2017	Blood factor
factor VIII humano de coagulación	Baxalta	2018	Blood factor
coagulation factor Xa recombinant inactivated-zhzo	Portola	2018	Blood factor
coagulation factor Xa recombinant inactivated-zhzo	Portola	2018	Blood factor
rh factor IX–albumin fusion protein	CSL Behring	2016	Blood factor-Fusion protein
Interferon beta 1a	BiogenIdec	1996	Cytokine
Interferon beta-I a	Serono	2003	Cytokine
rh PEGylated IFN- β -1a	BiogenIdec	2014	Cytokine
Etanercept	Amgen	1998	Fusion protein
Alefacept	BiogenIdec	2003	Fusion protein
Etanercept	Samsung Bioepis	2016	Fusion protein
Etanercept	Pfizer	2017	Fusion protein
Epoetin alfa	Amgen	1989	Hormone
Epoetin alfa	Ortho Biotech	1990	Hormone
Somatropin	Novo Nordisk	1995	Hormone
Follicle stimulating hormone (FSH)	Serono	1997	Hormone
FSH	Organon	1997	Hormone
Thyrotropin alfa	Genzyme	1998	Hormone
Darbepoetin alfa	Amgen	2001	Hormone
Choriogonadotropin alfa	Serono	2003	Hormone
rh erythropoietin alfa	Medice Arzneimittel Putter	2007	Hormone
rh erythropoietin α	Sandoz	2007	Hormone
rh erythropoietin- α	Hexal	2007	Hormone
methoxy polyethylene glycol-epoetin β	Roche	2007	Hormone
epoetin theta rh EPO	Ratiopharm (now Teva)	2009	Hormone
modified rhFSH in which the C-terminal peptide of the β -subunit of hCG	Merck Sharp & Dohme	2010	Hormone-Fusion Protein
Rituximab	Genentech/BiogenIdec	1997	mAb

Trastuzumab	Genentech	1998	mAb
Alemtuzumab	MilleniumIdex	2001	mAb
Adalimumab	Abbot Laboratories	2002	mAb
Efalizumab	Genentech	2003	mAb
Bevacizumab	Genentech	2004	mAb
obinutuzumab	Roche (Genentech)	2013	mAb
siltuximab	Jansen	2014	mAb
vedolizumab	Takeda/Millenium	2014	mAb
ixekizumab	Eli Lilly	2016	mAb
rituximab	Celltrion	2017	mAb
guselkumab	Jansenn	2017	mAb
pertuzumab	Roche/Genentech	2017	mAb
erenumab-aooe	Amgen	2018	mAb
benralizumab	AstraZeneca	2018	mAb
infliximab	Sandoz	2018	mAb
trastuzumab	Celltrion	2018	mAb
trastuzumab emtansine	Roche/genentech	2013	mAb-confugate
Ibritumomab tiuxetan	BiogenIdex	2002	mAb-conjugate
Zoster vaccine	GSK	2018	Recombinant protein / Vaccine
Hepatitis B vaccine	GSK	1989	Recombinant Protein/ Vaccine
Alteplase recombinant (tPA)	Genentech	1987	Therapeutic enzyme
Dornase alfa	Genentech	1993	Therapeutic Enzyme
Imiglucerase	Genzyme	1994	Therapeutic Enzyme
Retepase (tPA)	Centocor	1996	Therapeutic enzyme
Tenecteplase	Genentech	2000	Therapeutic enzyme
Laronidase	Biomarin	2003	Therapeutic Enzyme
Agalsidase beta	Genzyme	2003	Therapeutic Enzyme
Velmanase alfa	Chiesi	2018	Therapeutic enzyme

Most of these products have benefited from the increases in titres and yields after downstream obtained to date. However, for some, such as the more labile blood factors, or other products that would prove to be cytotoxic for cells, pushing cells to produce more product might not only not improve the overall process yield but also make it worse. For these, time regulated expression or biphasic cultures where product expression is suppressed during growth and switched on during the production phase would be more suited.

1.3. CHO cell engineering: successes and challenges

Chinese Hamster Ovary (CHO) cells are the preferred system for the production of biopharmaceuticals (Walsh, 2014). Improving their production capabilities has always been desirable in order to reduce the cost of producing therapeutic proteins (Butler and Meneses-Acosta, 2012). Significant improvements in cell growth and productivity have been achieved in the last few decades, mainly via bioprocess and media optimization, increasing protein titres from the initial few milligrams per litre to more than 10 g/L. Although there have been large efforts to push the boundaries further by cell engineering, such strategies to date have been limited in terms of industrial application.

Cell engineering efforts have been focused in three main areas of improvement of the CHO expression platform: selection systems, productivity improvements and product quality improvements. The different cell engineering strategies implemented to achieve those gains can be classified in mainly 3 groups: gene overexpression, gene knock outs, and cell engineering using short non-coding RNAs such as siRNA or miRNA.

Briefly, gene overexpression relies on the introduction of an expression vector carrying an expression cassette for a transgene that will be expressed transiently or be integrated stably into the genome of CHO cells, granting them superior characteristics. Improvements in the productivity of CHO cells have been made by overexpressing genes that extend the lifespan of the culture, such as anti-apoptotic genes or pro-proliferation genes (Fischer, Handrick and Otte, 2015). A classical example of these would be the overexpression of human BCL-like antiapoptotic genes that have been shown to extend culture and increase product titers by several groups (Chiang and Sisk, 2005; Majors *et al.*, 2012). More recently, overexpressing CHO proteins such as Heat shock protein 27 (HSP27) have also shown to improve culture viability and peak cell densities, thus improve mAb titers (Tan *et al.*, 2015). Gene overexpression has also been useful for

improving the quality of recombinant proteins expressed in CHO cells. An early study by Lee et al. where they overexpressed the gene encoding for a beta-galactoside alpha 2,6-sialyltransferase showed that post-translational modification of recombinant proteins, in particular glycosylation, could be altered (Lee, Roth and Paulson, 1989). A more recent study by Amman et al. showed how it is possible to obtain fully humanized N-glycosylation profiles on recombinant proteins expressed in CHO, not only by overexpressing specific glycosylase enzymes but also by knocking-out others using CRISPR/Cas9, therefore permanently removing them (Amann *et al.*, 2019).

This last example brings us to the use of gene knock-out strategies to improve the CHO protein expression platform. Staying with the last mentioned example, Amman et al. demonstrated a 10-gene knock-out to eliminate the CHO specific glycosylation enzymes and thus achieve two fully humanized recombinant products, alpha-1-antitrypsin and C1 esterase (Amann *et al.*, 2019). These permanent gene deletions or knock-out have been achieved using nowadays state-of-the-art technologies such as zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and more recently CRISPR/Cas9 technology, the latest having been a major break-through allowing us to very precisely edit virtually any single location in the genome. Another example of a gene knock-out used to improve product quality would be the inactivation of the GDP-fucose transporter gene using various gene editing technologies (TALENs, ZFNs, and CRISPR/Cas9) by Chan *et al.*, thus generating CHO cells to generate fucose-free antibodies, which exhibit better cytotoxicity profiles (Chan *et al.*, 2016). Knock-out strategies have also been used to generate improved CHO cell phenotypes aiming to improve growth and productivity. Once more with an example targeting pro-apoptotic pathways, Cost et al. showed how culture lifespan and titer could be increased by deleting the BAK and BAX pro-apoptotic genes using ZFN (Cost *et al.*, 2010). More recent studies have used CRISPR/Cas9 to improve the catabolic metabolism of CHO cells and reduce by-product formation (Ley *et al.*, 2019). Gene knock-out strategies have been essential for the generation of selection systems, crucial for the development of recombinant stable cell lines expressing recombinant proteins. While they were both generated in the early days by non-specific mutagenesis, DHFR-deficient and Glutamine Synthetase-deficient cell lines are probably the most prominent successes in terms of gene-knock engineering, both in terms of impact, allowing all other cell engineering

strategies to follow, and also actual industrial applicability (Fischer, Handrick and Otte, 2015).

Last but not least, siRNA/miRNA have also been used to manipulate CHO cell phenotypes. Since the identification of the first CHO-specific process relevant miRNA, there has been a significant effort to identify miRNAs involved in the regulation of bioprocess-relevant pathways as potential engineering targets to improve CHO cellular phenotypes (Gammell *et al.*, 2007; Niall Barron *et al.*, 2011). Several miRNAs have been identified and manipulated to improve process relevant cell characteristics such as cell growth, apoptosis, and recombinant protein production (N. Barron *et al.*, 2011; Druz *et al.*, 2013; Strotbek *et al.*, 2013; Jadhav *et al.*, 2014; Fischer, Mathias, *et al.*, 2015; Kelly *et al.*, 2015). In addition, several studies have already highlighted miRNAs that exhibit expression profiles which vary depending on specific conditions, such as different culture stages (Gammell *et al.*, 2007; Bort *et al.*, 2012), different culture conditions like a temperature shift (N. Barron *et al.*, 2011; Emmerling *et al.*, 2016) or different cellular phenotypes such as high/low Qp, producers/non-producers and fast/slow growth rate (Clarke *et al.*, 2012; Maccani *et al.*, 2014).

Little attention has been dedicated to utilizing miRNA molecules as control elements, rather than direct targets, in order to implement more refined cell engineering strategies. In order to do so, re-evaluating these datasets from this new perspective could help to identify specific miRNA signatures that would allow implementation of some of the reviewed miRNA-based circuits in CHO cells for particular bioprocess relevant applications, such as conditional expression after a temperature shift or selective killing of low producers to improve selection methods. There are some challenges that can be anticipated in terms of introducing miRNA-based gene control in CHO cells. For instance, the use of miRNA-based genetic switches could result on the knock down of the 'controller' miRNAs, i.e. the sponge effect, potentially leading to unwanted cellular effects. In addition, full ON/OFF control is likely to be difficult to achieve using the currently available miRNA-based gene control technologies, so combinations with other gene control elements such as promoters or other genetic elements should be further investigated.

1.4. miRNA-based CHO cell engineering – from targets to tools

Since the first CHO genome was published in 2011, advances in next-generation sequencing and other ‘omics technologies such as proteomics or metabolomics, have improved our ability to study CHO cells on a systematic and genome-wide scale (Xu *et al.*, 2011; Clarke *et al.*, 2012; Datta, Linhardt and Sharfstein, 2013; Fischer, Handrick and Otte, 2015). In addition, the CHO field has recently entered the systems biology era with the development of the first consensus genome-scale model, integrating the available genomic, transcriptomic, proteomic and metabolomic data. This model should facilitate a better understanding of CHO cell biology, leading to novel rational genetic engineering approaches to optimize CHO cell lines for bioprocessing (Kaas *et al.*, 2014; Harcum and Lee, 2016).

For instance, several studies have acknowledged that improvements in cellular productivity will not likely be achieved by engineering a master gene or small subset of genes, but rather by coordinating and balancing a variety of complex cellular pathways in an optimal way (Clarke *et al.*, 2012; Harreither *et al.*, 2015; Ley *et al.*, 2015; Wilkens and Gerdtzen, 2015). In this context, it is likely that traditional cell engineering strategies, such as crude gene overexpression, gene knock-down or gene knock-out might not be nuanced enough to address some of these challenges (Fischer, Handrick and Otte, 2015).

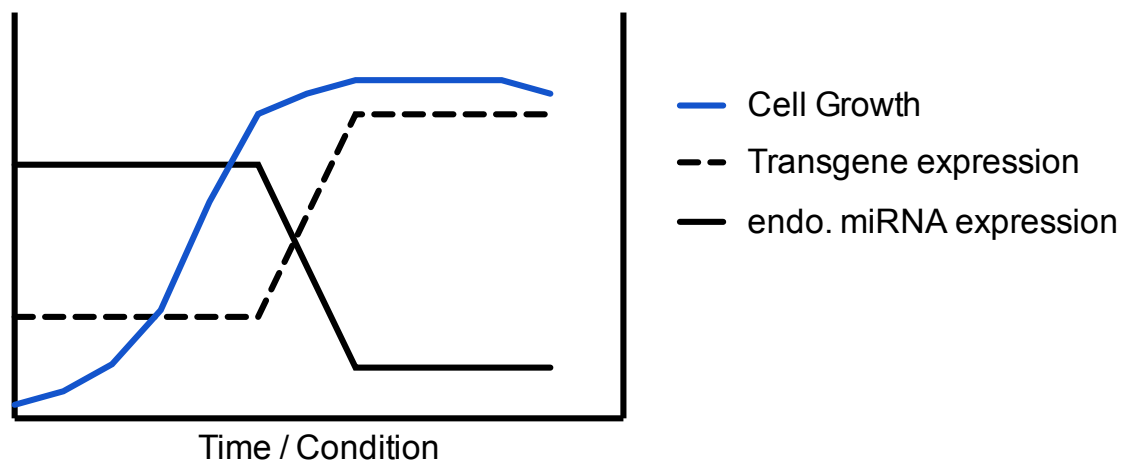


Figure 1 – A theoretical example of transgene expression control based on an endogenously expressed microRNA. Transgene expression is repressed during growth phase, when microRNA levels are high, and it is turned on to the right dose when the culture enters in stationary phase and the “controller” microRNA levels drop.

More sophisticated approaches, such as metabolic pathway rewiring, or gene dosage engineering will be required. For these to be implemented, gene control systems beyond the ON/OFF paradigm and allowing fine-tuning of gene expression need to be established. These systems would ideally utilize expression vectors capable of linking transgene expression to endogenous molecules, for example microRNAs (miRNAs), which show specific expression profiles, in order to provide the right transgene dose at the right time, in response to a specific metabolic or environmental signal (Figure 1).

The field of synthetic biology offers the most promising approach to gain better control over gene expression, and ultimately be able to rationally engineer organisms and control their behaviour (Purnick and Weiss, 2009). A variety of DNA- or RNA-based expression control systems allowing fine control of gene expression has been developed. DNA-based systems are mainly transcriptional control elements or inducible promoters. Although the number of inducible promoters available in mammalian cells is small when compared with *E. coli* or yeast, there have been several examples of inducible promoters applied to the production of biopharmaceuticals in mammalian cells (Wang *et al.*, 2012; Vogl and Glieder, 2013; Weinhandl *et al.*, 2014). Some examples include the doxycycline-inducible Tet-ON / Tet-OFF system, the metallothionein expression system, responsive to heavy metals, the glucocorticoid-inducible mouse mammary tumour virus long terminal repeat (MMTV-LTR) or a biotin inducible promoter (Huang, Marquis and Gray, 2004; Lipscomb, Mowry and Kompala, 2004; Figueroa *et al.*, 2007; Weber *et al.*, 2009; Santos, Parreira and Resende, 2016). RNA-based systems are mainly riboswitches. Originally discovered in bacteria, these RNA-domains can undergo changes in their secondary structure after binding of a small-molecule ligand to modulate transgene expression at both transcriptional and post-transcriptional levels (Serganov and Nudler, 2013). In mammalian cells, riboswitches responsive to small molecules, but also to proteins, have been developed (Ausländer *et al.*, 2014; Kennedy *et al.*, 2014; Beilstein *et al.*, 2015). One of the main drawbacks of these types of inducible elements is the need to add a controller molecule to the culture in order to modulate transgene expression. A more interesting group of RNA-based inducible systems is the one in which the regulating molecules are endogenously expressed miRNAs or synthetic siRNAs. This review focuses on the state-of-the-art miRNA/siRNA-based genetic switches and on their potential use to improve control over transgene expression in CHO cells. In addition, we discuss the challenges in adapting the currently developed miRNA-

based gene control strategies, which mainly focus on cell therapy or diagnostic applications, to the context of biopharmaceutical bioprocesses and CHO cells.

2. miRNAs and siRNA: biogenesis, regulation and function

2.1. miRNA biogenesis

miRNAs are a type of short non-coding RNA (sncRNA) (21-25 nt) that play an important role in the regulation of gene expression at the post-transcriptional level (Bartel, 2009). The canonical biogenesis pathway of miRNAs involves several steps. First, a miRNA gene is transcribed in the nucleus by RNA polymerase II to produce a primary miRNA transcript (pri-miRNA), which comprises a hairpin domain flanked by two single stranded RNA domains (Lee *et al.*, 2004). In terms of their genomic location, and taking human and mouse as reference organisms, miRNAs are mainly found in intronic regions of protein-coding transcripts and intergenic non-coding regions of the genome (Georgakilas *et al.*, 2014; de Rie *et al.*, 2017).

After transcription, pri-miRNAs undergo a processing step in the nucleus. First, the RNase-III enzyme DROSHA, with its essential co-factor protein DGCR8, together known as the Microprocessor protein complex, crops the pri-miRNA molecule to form a 60-80 nt long precursor miRNA (pre-miRNA), that conserves the hairpin-like secondary structure and can be exported to the cytoplasm by a transport complex formed by Exportin5 and RAN-GTP (Yi *et al.*, 2003; Zeng and Cullen, 2005; Ha and Kim, 2014). As explained below, nuclear processing is one of the main points where miRNA biogenesis regulation occurs. In addition, as the pri-miRNA cleavage by Drosha will determine the miRNA terminus, it is also a critical step in determining its specificity (Ha and Kim, 2014). Once in the cytoplasm, the pre-miRNA is cleaved near its terminal loop by a second RNase-III, Dicer, to form a dsRNA intermediate, from which the mature miRNA strand will be loaded into Argonaute proteins (AGO), as part of the miRNA-induced silencing complex (miRISC). Both, this second processing step and the loading of the dsRNA intermediate onto Argonaute proteins are two a key points where miRNA biogenesis is regulated (Kim, Han and Siomi, 2009; Ha and Kim, 2014).

Short interfering RNAs (siRNAs) can be derived from different biogenesis pathways, such as cleavage of long double stranded RNA (dsRNA) by Dicer, or be introduced artificially in the cell (Tam *et al.*, 2008; Kim, Han and Siomi, 2009). siRNAs are also loaded into AGO2

proteins and guide the binding of the siRNA silencing complex (siRISC) to their target mRNA molecules. Whereas typically miRNAs can have multiple targets due to their imperfect complementarity, siRNAs are specific to their target as they are fully complementary to their binding site. Perfect binding between siRNA and their target leads to its endonucleolytic cleavage by the RNase Dicer, a mechanism known as RNA interference (RNAi) (Kim, Han and Siomi, 2009). To date, the role of endogenous siRNA in mammalian cells is still poorly understood. However, exogenously added siRNA have been broadly used as tools to investigate individual gene function (Cullen and Arndt, 2005; Papageorgiou *et al.*, 2015).

For clarity, in this introduction, the term 'miRNA' will refer to any molecule originated from a miRNA precursor (pre-miR) either transcribed from cellular genes and being endogenously expressed in the cell, or to a miRNA that comes from an exogenously introduced plasmid containing the pre-miR hairpin sequence and is therefore expressed and processed through the miRNA expression pathway. siRNA will only refer to exogenously short double stranded RNA introduced by transfection or electroporation.

Formation of the pre-RISC and later mature RISC are the last steps before miRNA can go bind their target transcripts and are also tightly regulated. Here, two distinct steps have been characterized: the loading of the RNA duplex and its subsequent unwinding leaving only the mature strand loaded into RISC.

The loading step involves the binding of the RNA duplex intermediate to the AGO protein. This is an active ATP-dependent process mediated by chaperone proteins and. Several studies have shown that in certain species (*Drosophila*) specific AGO proteins (AGO1) show preferential affinity for miRNA duplexes while in humans miRNA and siRNA can be loaded indistinctively into several types of AGO proteins (AGO1-4) (Kawamata, Seitz and Tomari, 2009; Ha and Kim, 2014). Studies in *Drosophila*, *C. elegans* and human cells have shown that the loading step is at least partially dependent on structural features of the miRNA duplexes, with mismatches in the central positions of the duplex promoting loading into AGO proteins (Steiner *et al.*, 2007; Kawamata, Seitz and Tomari, 2009; Yoda *et al.*, 2010).

The last step for the formation of the mature RISC is the miRNA duplex unwinding step, which leads to discarding the "passenger" strand of the miRNA duplex and leaves the mature strand bound to one of the AGO proteins. This unwinding step has been shown

to be ATP-independent and determined by structural features of the seed and 3'-mid regions of the miRNA duplex, with mismatches in these regions favouring strand separation (Kawamata, Seitz and Tomari, 2009; Yoda *et al.*, 2010).

2.2. Regulation of miRNA biogenesis

As outlined above, during miRNA biogenesis, several steps are susceptible of being regulated and together control the formation of new miRNAs in the cell. These can broadly be divided into transcriptional and post-transcriptional regulatory mechanisms and will be briefly described here.

Our understanding of transcriptional regulation of miRNA expression relies on our knowledge around miRNA promoter regions and how their structural features determine binding of transcription or epigenetics factors that regulate transcription of the pri-miRNA. Chromatin immunoprecipitation experiments have shown that many of the miRNA promoter features such as the relative frequencies of CpG islands, TATA box, TFIIB recognition, initiator elements and histone modifications, are similar to the promoters of protein coding genes (Ozsolak *et al.*, 2008; Corcoran *et al.*, 2009; Davis and Hata, 2010). A good example of transcriptional regulation of miRNA can be seen in studies relating c-Myc to the expression of several tumorigenic miRNAs (Chang *et al.*, 2008). Similarly, epigenetic mechanisms that control the expression of protein-coding genes such as DNA methylation or histone modification seem to apply for the transcriptional regulation of miRNA expression (Morales, Monzo and Navarro, 2017). Methylation of CpG islands, either in the miRNA promoter region, or in distal genomic regions upstream of the miRNA locus or even for miRNAs which are embedded in CpG rich regions, has been shown to affect their baseline levels of transcription in several cancer studies (Morales, Monzo and Navarro, 2017). In addition, the role of epigenetic mechanisms such as histone modification have also been described in relation with miRNA regulation, also mainly in cancer. As an example, a study in pancreatic cancer cells showed how the interaction of the protein Enhancer of zeste homolog 2 (EZH2) with the lncRNA HOTAIR drives the formation of heterochromatin around the miRNA miR-34a, thus silencing its expression (Li *et al.*, 2017).

As previously mentioned, miRNA biogenesis is also regulated at a post-transcriptional level. These mechanisms involve both, *cis*-acting sequence elements and *trans*-acting factors that influence each miRNA processing steps, namely pri-miRNA processing by

the microprocessor (DROSHA-DGCR8), and processing by Dicer of the pre-miRNA in the cytoplasm, have been recently reviewed and will be briefly described below (Michlewski and Cáceres, 2019).

In relation to the *cis*-acting sequence elements related to the regulation of the Microprocessor processing step, just remind us that the microprocessor complex, through the dsRNA binding domains and the RNA-binding heme domain of DGCR8 recognises the loop and stem regions of the pri-miRNA, respectively, efficiently binding to it. DROSHA acts a measuring tape to cleave the stem-loop of the pri-miRNA at an 11-nucleotides distance from the basal single stranded RNA. For the above to occur, the presence of a basal UG motif and an apical UGUG motif on the pri-miRNA are critical for the correct orientation of the Microprocessor complex (Nguyen *et al.*, 2015). Other important identifiers of pri-miRNA hairpins are a CNNC motif that binds the SR protein, SRSF3 protein (Kim *et al.*, 2018) or the presence of a N6-methyladenosine mark close to the stem-loop that is recognised by the reader protein hnRNP protein A2/B1 which stimulates miRNA processing by binding to DGCR8 (Knuckles *et al.*, 2017).

Regarding the regulation of DICER processing of the pre-miRNA in the cytoplasm, both structural and sequence *cis*-acting determinants have been shown to play an important role in defining the cleavage point. The presence of the apical stem-loop or other asymmetric secondary structures in that region have been shown to allow Dicer to measure the distance for its cleavage point (MacRae, 2006; Starega-Roslan *et al.*, 2011). However, sequence determinants have also been shown to influence Dicer activity, affecting the precise position in which Dicer will cut, resulting in a variety of RNA duplexes arising from the processing of the same pre-miRNA (Starega-Roslan, Galka-Marciniak and Krzyzosiak, 2015).

As mentioned previously, trans-acting factors have also been shown to regulate miRNA biogenesis. These are directly related with the *cis*-acting sequence elements mentioned above. These trans-acting regulators have been recently reviewed and mainly include RNA-binding proteins (RBP) and long-non coding RNAs (Ratnadiwakara, Mohenska and Änkö, 2018; Michlewski and Cáceres, 2019). In short, these trans-acting levels act by promoting or repressing processing of the pri-miRNA in the nucleus or the pre-miRNA in the cytoplasm, by binding to the non-coding RNA molecule. Other trans-regulatory

mechanisms involve post-transcriptional modifications of the processing complexes DROSHA and DICER by phosphorylation as well as ubiquitination (Ha and Kim, 2014).

2.3. Target Recognition and mechanisms of action.

Based on an imperfect sequence complementary between the loaded mature miRNA and its mRNA target, the miRISC complex can bind to the untranslated regions (UTR) of the target mRNA transcript and inhibit its translation, or in some cases of high level of complementarity between the miRNA and its target lead to endonucleolytic cleavage of the transcript (Lytle, Yario and Steitz, 2007). The minimal and most critical element for miRNA to recognise and bind to their target has been known as the seed region. This seed region comprises the 2-7 nucleotides at the 5'-end of the mature miRNA loaded into the miRISC complex (Bartel, 2018; Chipman and Pasquinelli, 2019). Indeed, recent studies have shown that loaded miRNA into AGO2 as part of the miRISC only have partly exposed nucleotides of this seed region and that pairing of these first 2-5 nucleotides with a target mRNA generates conformational changes that expose the rest of the miRNA sequence in a step-wise fashion, allowing further interactions to occur (Schirle, Sheu-Gruttadauria and MacRae, 2014).

Two main miRNA mechanisms of action have been shown upon binding of miRNA/miRISC to a target mRNA: miRNA mediated target degradation and miRISC mediated translation inhibition and mRNA decay. The degree of complementarity between the miRNA and its target mRNA determines which one from the two occurs. Extensive miRNA:mRNA complementarity is able to trigger the degradation of the target mRNA by the slicing activity of the AGO proteins, which cleave the miRNA:mRNA duplex (Jo *et al.*, 2015). However, in animals, non-fully complementary miRNA:mRNA interactions are the norm, and these do not allow cleavage by AGO protein. These imperfect interactions allow miRNA-mediated gene silencing by promoting a combination of translational repression and mRNA destabilization (Jonas and Izaurralde, 2015). Translation inhibition is believed to be an early silencing mechanism that lasts until a strong full repression is established. mRNA destabilization and subsequent decay through the 5'-to-3' mRNA decay pathway after deadenylation and decapping of the mRNA target have been shown to account for most of miRNA mediated repression with the miRISC, in particular the AGO proteins interacting with the GW182 protein, which in

turns interacts with cytoplasmic deadenylase complexes that start the mRNA decay process (Eichhorn *et al.*, 2014).

3. Exploiting endogenous miRNA signatures to modulate transgene expression

Several studies have reported on using miRNAs to modulate the expression of transgenes in specific cell types or tissues. After identifying specific miRNA signatures in the desired cell type, miRNAs have been used as inputs for genetic switches containing a transgene, typically a reporter gene, under the control of microRNA-responsive elements (MREs). These MREs are simply miRNA-binding sites located in the UTR of the transgene.

One of the first examples of miRNA-controlled cell-specific transgene expression relied on a miRNA-sensitive OFF-type sensor and was used to study miRNA expression patterns during the development of mouse embryo (Mansfield *et al.*, 2004). An OFF-type switch comprises a reporter gene, a Green Fluorescent Protein (GFP) or β -Galactosidase, which contains specific fully complementary miRNA-binding sites in its 3'-UTR. In the presence of the corresponding miRNA, the reporter gene mRNA transcript is sliced by RNAi, hence marking the areas where the miRNA is being expressed by the absence of fluorescence (GFP reporter) or lack of coloration in the presence of X-Gal (β -Galactosidase reporter) (Figure 2 – A). Two later consecutive studies by Brown *et al.* used a similar approach in order to obtain selective transgene expression in different human cell types based on their specific miRNA expression signatures. Again, the vector comprised a GFP reporter transgene, and specific 3'-UTR miRNA-binding sites, which lead to a downregulation of the transgene in the presence of the complementary miRNA. By using a double reporter vector, with one reporter gene expressed constitutively and the second one under the control of miRNA-binding sites, specific populations of closely related cell types (including dendritic cells, hematopoietic and embryonic stem cells, and their progeny, which exhibit different miRNA signatures) or specific tissues, like retina, could be segregated based on the differential miRNA-controlled expression of the reporter transgene (Brown *et al.*, 2006a, 2007; Karali *et al.*, 2011). More recent studies used this type of miRNA-based transgene expression vector to classify neuron-type and brain-region-specific mice cells based on their cell-type specific miRNA signature (Sayeg *et al.*, 2015).

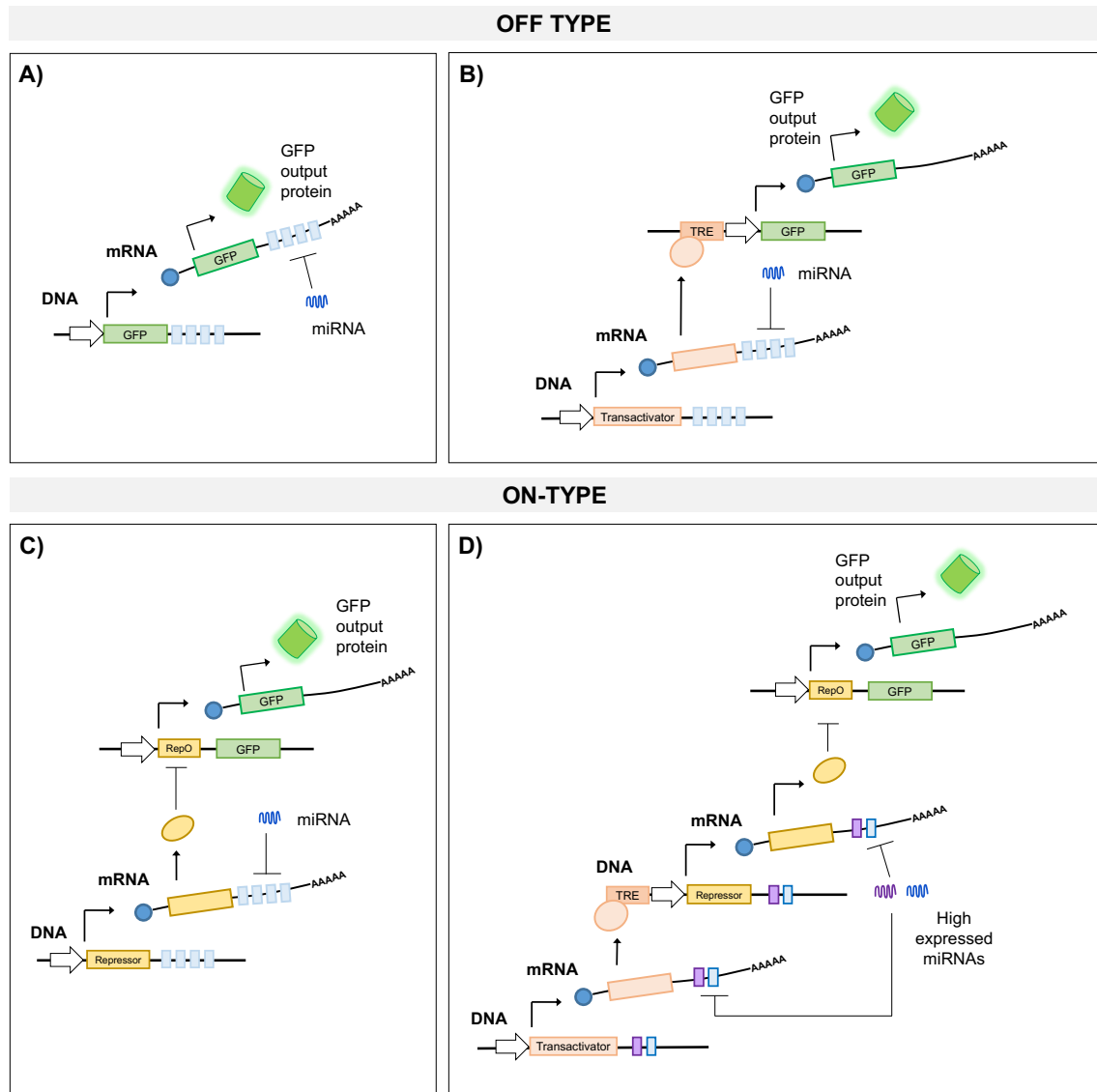


Figure 2 – Schematic representation of different miRNA-based genetic switches or circuits able to react to specific miRNAs signatures. OFF-type switches or circuits are able to repress the expression of the output gene in the presence of specific miRNAs signatures. (A) – The output gene is controlled by miRNA-binding sites at its UTR, which lead to translation inhibition in the presence of the specific miRNA(s) signature. (B) – The output gene is controlled by a transactivator-responsive element (TRE). The transactivator protein is controlled by miRNA-binding sites at its UTR, which inhibit its translation in the presence of the specific miRNA signature. ON-type switches or circuits are able to allow the expression of the output gene in the presence of specific miRNA signature. (C) - The output gene is controlled by a repressor-dependent promoter (RepO). The repressor protein is controlled by miRNA-binding sites at its UTR, which inhibit its translation in the presence of the specific miRNA(s) signature. (D) – Double inversion module. The output gene is controlled by a repressor-dependent promoter. The repressor protein is controlled by a transactivator protein. In addition, both the repressor protein and the transactivator protein have miRNA-binding sites in their 3'-UTR, specific to high abundance miRNAs. Only in the presence of the specific high miRNA the output protein can be expressed.

Another example of OFF-type miRNA-based sensors is a study that identified and sorted different cell populations from complex mixtures of differentiated human pluripotent stem cells (hPSCs), using two types of OFF-switches: one based on miRNA-binding sites in the 3'-UTR of the reporter gene and one with a single miRNA-binding site in the 5'-UTR. The authors demonstrated the possibility of specifically selecting a given cell type, namely cardiomyocytes, using a miR-based switch driving the expression of the pro-apoptotic gene Bim in cells that do not show the proper miRNA signature (high miR-1 or high miR-208a). In addition, the authors proved that the obtained cells were still fully functional and could be used for medical applications such as in vivo transplantation therapy. The same technology was used to purify other cell types. In addition, the authors demonstrate the possibility of combining miRNA switches in order to simultaneously purify two different types of cells (Miki *et al.*, 2015). In a more complex version of a miRNA-based OFF-type switch, developed by Xie *et al.*, the expression of the output gene is controlled by a transactivator placed under the control of a particular miRNA by the presence of miRNA-binding sites at the 3'-UTR. The binding of the specific miRNA inhibits the expression of the transactivator, hence blocking the expression of the output gene (Figure 2 – B) (Xie *et al.*, 2011). A last example of a practical application of these miRNA-based genetic switches is a cell-type-specific genome editing platform developed based on a miRNA-responsive element developed using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated protein (Cas9) system, which was able to drive targeted genome editing of human cell lines (HeLa and induced pluripotent stem cells) in a modulated fashion in response to endogenous miRNA signatures (Hirosawa *et al.*, 2017). The system comprises a Cas9 gene, modified by adding a miRNA-binding site to the coding sequence, complementary to the miRNA of interest. These miR-Cas9 switches are able to control genome editing based on the presence/absence of specific miRNAs. In the presence of the miRNA, the Cas9 transcript is degraded and no genome editing activity can occur. In the absence of the particular miRNA, the transcript is not degraded and can be translated, thus the Cas9 protein can drive the desired targeted genome edition based on the provided single guide RNA. The authors show how a miR-21-Cas9 switch responded to the cell-specific miR-21, which is highly expressed in HeLa cells, switching OFF the CRISPR-Cas9 KO of a stably transfected GFP reporter gene. On the other hand, a miR-302-Cas9 switch was not functional, as miR-302 is not expressed in these cells, therefore resulting in the GFP

gene being knocked out. The authors also provide evidence that the presence of specific miR-21 inhibitors switched ON Cas9 protein expression, re-instating genome editing activity. The authors used the same miR-Cas9 switch to target other endogenous regions in the genome, including the Alu1 element, leading to selective cell death based on the chosen miRNA-binding site present in the switch.

All the above examples rely on OFF-type miRNA-based switches, in which specific miRNA signatures act as the input, switching OFF the expression of the transgene. There are other examples of miRNA-based circuits whereby the presence of specific miRNA input turns ON the expression of the transgene. An example of this type of ON-circuit has been developed and tested in human embryonic kidney cells (HEK) (Rinaudo *et al.*, 2007). It comprises an endogenous miRNA which controls the expression of a synthetic miRNA using an OFF-switch configuration as explained above. The output gene also contains binding sites for the synthetic miRNA/siRNA in its UTR. Hence, in the presence of the endogenous miRNA, the expression of the synthetic miRNA is repressed, and the output protein can be expressed. Another example of miR-based ON-switches comprises a repressible promoter driving the expression of an output protein and a repressor protein containing miRNA-binding sites in its UTR. In the presence of the input miRNA, the expression of the repressor protein is blocked, allowing the output protein to be expressed (Figure 2 – C).

Combining these different types of miRNA-based switches, it is possible to build more complex multi-input circuits, which can react to specific miRNA signatures and allow the final expression of an output protein. Xie *et al.* describe a cell classifier able to detect specific cancer cells in a mixed pool of different human cancer cell lines, based on more complex miRNA signatures, for example the one found in HeLa cells, which show high expression of miR-17, miR-21 and miR-30a and low expression of miR-141, miR-142(3p) and miR-146a. By using a double inversion module circuit design, a reverse tetracycline-controlled transactivator (rtTA) controls the expression of an inducible LacI repressor, which represses the transcription of the output gene. In addition, both trans-activator and repressor have miRNA binding sites specific to the highly expressed miRNAs (17, 21, and 30a). In the presence of these miRNAs, both transactivator and repressor proteins are repressed due to specific binding sites in their UTRs and the output protein can be expressed (Figure 2 – D). To incorporate a sensing capacity for low expressed miRNAs,

the output gene also contains binding sites for the low expressed miRNAs (hsa-miR-141, hsa-miR-142(3p) and hsa-miR-146a). When low expressed, these miRNAs are not sufficient to efficiently inhibit the output gene; therefore, it is expressed. Any change in expression of any of the controller miRNAs (high expression changing to low expression or vice versa) will result in repression of the output gene (Xie *et al.*, 2011). In a similar way, other cell classifiers able to detect complex miRNA signatures in different human cell lines (HeLa cells and HEK293 cells) have been developed using different circuit designs. For example, Wroblewska *et al.* developed a cascade circuit, in which a synthetic siRNA controlled the expression of a RNA-binding protein (RBP), which in turn controlled the expression of the output gene, and a more complex two-way signal transmission and feedback circuit, in which two different siRNA inputs are able to modulate the expression of two different reporters, by controlling different RBPs (Wroblewska *et al.*, 2015). Other examples can be found as well, in which miRNA-based circuits are combined with other genetic control elements such as transcription repressors or transactivators and which are able to react to specific miRNA signatures (Prochazka *et al.*, 2014; Li *et al.*, 2015).

4. Synthetic miRNAs to modulate transgene expression

4.1. Inducible expression of synthetic miRNA/siRNAs

Synthetic miRNAs have also been widely used in genetic circuits to modulate transgene expression. Some studies have combined them with inducible promoters to be able to control expression in a dose and time-dependent manner. Leisner *et al.* have built and characterized different synthetic genetic circuits which comprise miRNAs controlled by transcription factor-dependent promoters, which down-regulate the expression of an output reporter gene by RNAi. They show how multiple combinations of different modular circuits can lead to different programmable protein expression patterns, based on the different inputs present in the circuit (Leisner *et al.*, 2010).

In another study, Malphetes *et al.* described the first adjustable RNA polymerase II promoter-based siRNA expression system, which allows the expression of siRNA hairpins (or miRNAs) based on the absence/presence of an antibiotic (erythromycin or tetracycline), in a dose-dependent manner, for the precise control of transgene expression in CHO cells. In a later study, the previously described siRNA expression

system is used to fine tune the expression of different transgenes in CHO cells, including secreted embryonic alkaline phosphatase (SEAP), vascular endothelial growth factor (VEGF), or erythropoietin (EPO) (Malphettes and Fussenegger, 2006). This study is the only one in CHO cells where a systematic characterization of different circuit relevant parameters, such as the number of binding sites for the specific siRNA is performed. Combining the above controllable siRNA expression module with other transcriptional regulatory elements such as a tetracycline-dependent inducible promoter allowed direct control of transgene expression.

In a more recent study, an OFF genetic control device based on RNAi that combines ligand-responsive ribozyme switches with synthetic miRNA regulators has been built and characterized (Bloom, Winkler and Smolke, 2015). The genetic device comprises a ribozyme, where cleavage is dependent on the presence/absence of a ligand and which controls the expression levels of a synthetic miRNA. Finally, synthetic miRNAs have been used as add-ons to improve other transgene expression control systems. miRNAs expressed from an artificial intron in the trans-regulator gene and targeting the output reporter gene have been used to reduce leaky expression, leading to tighter control of transgene expression in the OFF state (Greber, El-Baba and Fussenegger, 2008; Xie *et al.*, 2011; Lapique and Benenson, 2014).

4.2. Controlling miRNA expression by affecting their processing

Another interesting strategy to control the expression of synthetic or ectopic miRNAs or siRNAs is to fuse them with aptamers. An aptamer is an RNA domain which can bind to a specific ligand and change its secondary structure. One of the best-characterized aptamers is based on a previously described riboswitch, which is able to drive the termination of the transcription of a gene of interest in the absence of its binding ligand theophylline (Wachsmuth *et al.*, 2013). By introducing an aptamer sequence in between functional regions of the miRNA gene, such as the hairpin loop region or the basal stem region, it is possible to alter its secondary structure, hence affecting miRNA processing and consequently, its expression (An, Trinh and Yokobayashi, 2006; Beisel *et al.*, 2008, 2011).

A recent study used this strategy to modulate the expression of the Tissue Differentiation-Inducing Non-Protein Coding RNA (TINCR), a well-known long non-coding RNA (lncRNA) that acts as a master regulator in somatic differentiation and which

is up-regulated in bladder cancer tissues and cells, contributing to oncogenesis and cancer progression (Kretz *et al.*, 2013; Xu *et al.*, 2015). Here, the expression of the transgene could be controlled by the theophylline aptamer using two newly developed genetic switches, both based on the conditional expression of specific siRNAs (Chen *et al.*, 2016). The OFF-switch relies on the expression of a mature miRNA domain, in the presence of theophylline, which silences the TINCR lncRNA, switching it OFF and causing a reduction in cell growth. The ON-switch is based on the conditional expression of a mature miRNA, which is able to inhibit the processing of a second short-hairpin RNA (shRNA), which acts as a repressor of TINCR. By conditionally silencing TINCR, cell growth can be affected. This is another good example of how miRNA-based genetic switches can be used to modulate the expression of relevant target genes and lead to programmed cellular behaviour.

5. Novel miRNA responsive elements based on new modes of action

So far, most examples of miRNAs being incorporated in genetic circuits relied on the best-known mechanism of action of miRNAs or siRNA in the cell: translation inhibition or RNAi-mediated transcript degradation. However, some studies have exploited different mechanisms of action of miRNAs, providing new miRNA responsive elements for use in synthetic genetic circuits.

5.1. miRNA regulating RNA-binding protein motifs

Doyle *et al.* proposed alternative mechanisms through which miRNA can bind to mRNA molecules and affect how certain RBPs interact with them (Doyle and Tenenbaum, 2014). The model proposes that miRNA can act as structural interacting RNAs (sxRNA) and participate in forming 3-way RNA junctions to regulate gene expression. Post-transcriptional regulation of histone genes, by means of a histone stem-loop (HSL) structure located at the 3' end of the mRNA transcripts is used as an example. The Stem-Loop Binding Protein (SLBP) is able to bind to this structure and greatly increase the translation of the associated mRNA. By using secondary structure and nucleic acid interaction prediction software, the authors explored how miRNA could be responsible for generating or stabilizing the HSL structure and identify several target miRNAs which could potentially do it.

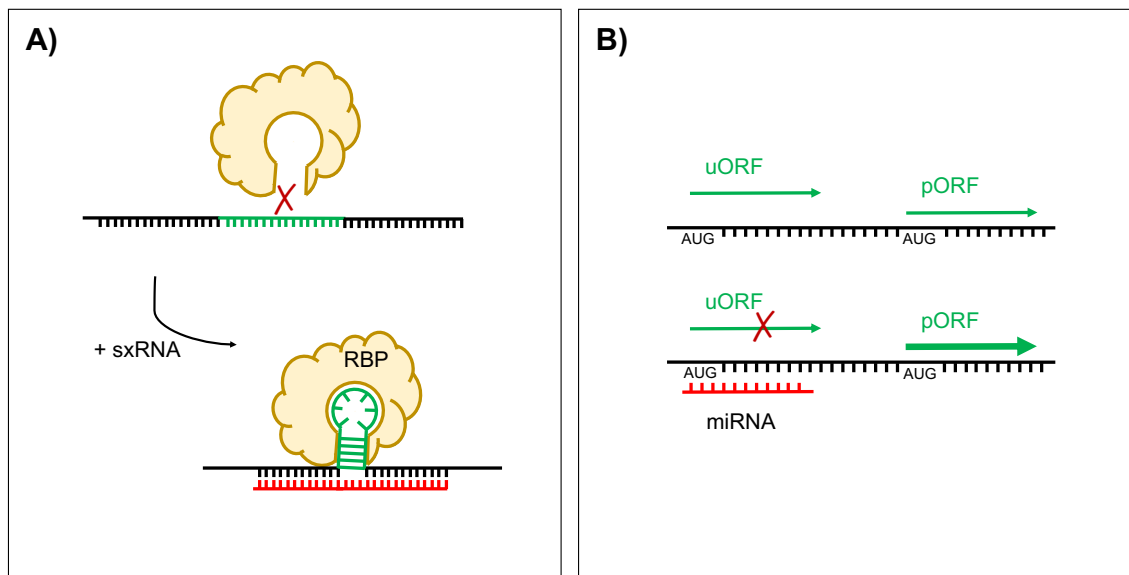


Figure 3- Schematic representation of potential novel miRNA-based responsive elements. (A) – The binding of a Structural Interacting RNA (sxRNA) leads to the formation of a particular stabilized secondary structure, which can be recognized by an RNA-binding protein (RBP). (B) – The presence of upstream open reading frames (uORF) reduces the translation efficiency of the primary ORF (pORF) leading to lower levels of full-length protein. The binding of an antisense oligonucleotides (ASO) complementary to the start codon region of the uORF blocks translation initiation events on this ORF, hence increasing the translation of the pORF.

In a similar way, the authors identified miRNAs that could potentially help form sxRNA structures in the iron- responsive element (IRE), which is targeted by IRE-binding proteins to regulate post-transcriptional expression of iron metabolism-related genes. The authors predict how specific miRNAs can influence alternative splicing of the ALA synthase gene by forming two different sxRNAs structures.

In a follow-up study from the same group, a sxRNA-based reporter consisting of a luciferase gene upstream of the sxRNA switch, was designed, built and tested in vitro, using *in silico* folding studies and gel-based and cell-based binding studies (Doyle *et al.*, 2017). The sxRNA switch comprises the sequence required to form the functional stem-loop structure, in order to bind the RBP, flanked by two paired regions complementary to the desired trigger miRNA (Figure 3 – A).

5.2. miRNAs to improve translation efficiency

A recent study has shown the possibility of using competing upstream open reading frames (uORF) and antisense oligonucleotides (ASO) to modulate expression of a transgene at the post-transcriptional level (Figure 3 – B) (Liang *et al.*, 2016). uORFs are

caused by the presence of initiation sequences and/or start codons upstream of the correct initiation sequence and start of the gene of interest, which initiates the primary ORF (pORF). The presence of these alternative uORFs down-regulates protein synthesis from the pORF due to competing translation events. To counteract this, ASOs can be designed to specifically target uORFs, hence blocking translation events starting in those without leading to any transcript degradation. This leads to the up regulation of the translation occurring from the main pORF. If we consider that the proposed ASOs are synthetic 16-18bp oligonucleotides (DNA/RNA), it is possible to imagine scenarios in which uORFs could be designed to be targeted by endogenous miRNAs and incorporated upstream of a transgene to modulate its expression.

6. miRNAs for refined genetic control of transgene expression in CHO cells

In the context of the more complex challenges that will arise from the new systems biology approaches in studying CHO cells, it is necessary to start thinking about developing more refined solutions to improve the CHO cell platform. Further developments and refinement of the current CHO genome-scale models will surely be crucial not only to identify which targets/pathways require these interventions but also to assess the potential benefits that would arise from them. Hypothetical examples of these more sophisticated approaches include conditional expression of specific product genes or gene-dosage balancing to fine-tune the expression of specific genes or pathways and achieve a more energetically efficient CHO cells free of unnecessary metabolic burden. In order to implement these new engineering approaches, the proper molecular tools have to be available.

Synthetic biology advances have provided us with a large variety of miRNA-based genetic switches or circuits, which allow us to gain tighter control over transgene expression. Most of the available miRNA-based gene expression control elements are based on tissue-specific miRNA signatures or diseased versus control cell profiles. Although miRNAs have been typically studied in CHO cells as targets for cellular engineering, we anticipate that meaningful miRNA signatures or expression patterns in the context of CHO cells and bioprocessing can be identified from the CHO miRNA profiling datasets that have been published to date. This should present new opportunities in terms of using miRNA-based genetic-switches in the next generation of

rational genetic engineering strategies and to facilitate more refined control over the CHO cell as a therapeutic protein production platform.

CHAPTER 2

Materials and Methods

1. Molecular Biology Techniques

1.1. Nucleic Acid Amplification

1.1.1. Taq Polymerase Chain Reaction

For testing primers and routine analysis of nucleic acids, the MyTaq 2X Kit (Bioline) was used. All reactions were set up in ice. Primer concentrations ranged from 0.2-1 μM . Template concentrations varied depending on the concrete experiment but ranged between 20-200 ng of DNA. A typical protocol for this type of reactions was:

10.0	μL MyTaq Master Mix (2X)
1.0	μL Forward Primer (10 μM)
1.0	μL Reverse Primer (10 μM)
X	μL Nuclease Free H_2O
X	μL Template DNA
<hr/>	
20.0	μL Total volume

PCR cycling conditions varied according to the template/fragment to be amplified and were adjusted starting from the following and taking into account the manufacturer recommendations.

Step	Temperature	Time	Cycles
Initial Denaturation	95°C	1 min	1
Denaturation	95°C	15 s	25-40
Annealing	User determined	15 s	
Extension	72°C	10 s	

1.1.2. High-Fidelity Polymerase Chain Reaction (PCR)

For high-fidelity amplification of specific DNA fragments from plasmid DNA template for further cloning, the Phusion® High-Fidelity PCR Master Mix (Thermo Fisher) was used.

All reactions were set up in ice. Primer concentrations ranged from 0.2-1 μM . Template concentrations varied depending on the concrete experiment but ranged between 20-200 ng of DNA. A typical protocol for this type of reactions was:

10.0	μL Phusion Master Mix (2X)
1.0	μL Forward Primer (10 μM)
1.0	μL Reverse Primer (10 μM)
X	μL Nuclease Free H_2O
X	μL Template DNA
<hr/>	
20.0	μL Total volume

PCR cycling conditions varied according to the template/fragment to be amplified and were adjusted starting from the following and taking into account the manufacturer recommendations.

Step	Temperature	Time	Cycles
Initial Denaturation	98°C	30 min	1
Denaturation	98°C	5-10 s	25-40
Annealing	User determined	10-30 s	
Extension	72°C	15-30 s/kb	
Final extension	72°C	5-10 min	1
	4°C	Hold	

1.1.3. Bacterial Transformation

For plasmid amplification, bacterial transformations were performed in a 1.7 mL micro centrifuge tube using 30-100 µL of Competent DH5alpha Sub Cloning Efficiency E. coli (Invitrogen), +10% of the volume of the ligation reaction or 1% of the volume for pure plasmid amplification. The mix was incubated in ice for 30 minutes. The mix was then subjected to a heat shock at 42°C for 40 seconds in a water bath, then incubated again in ice for 2 minutes. 300-500 µL of S.O.C. medium (Invitrogen) were added to the mix. Transformed cells were incubated for 1 hour at 37°C in a shaken platform at 220 rpm. Samples were then centrifuged at 4.000 rpm for 5 minutes. All the Supernatant, but 50-100 µL was discarded, and cell pellets re-suspended in the remaining medium. The obtained bacterial suspensions were plated in pre-warmed LB-Agar containing the corresponding selection antibiotic, by spreading 5-50 µL of bacterial suspension. Plates were incubated overnight at 37°C, lid down to avoid condensation drops to fall on the agar surface. Colonies could be observed after the incubation, and plates could be stored at 4°C for up to 2 weeks.

1.2. Nucleic Acid Isolation and Purification

1.2.1. Plasmid DNA ExtractionMini, Midi, Maxi Preps

1.2.1.1. Mini, Midi, Maxi Preps

Plasmid DNA was isolated and/or amplified using Mini, Midi or Maxi Prep Kit (Qiagen or Invitrogen) following the manufacturers protocols.

1.2.1.2. Gel extraction and PCR Purification

To isolate and purify amplified or digested DNA fragments of a specific size, the PureLink™ Quick Gel Extraction and PCR Purification Combo Kit was used following the manufacturers protocols.

1.2.2. Total RNA Extraction

Total RNA from cells was isolated using the Trizol® Reagent (Ambion-Life Technologies). Cell suspension containing 2.0×10^5 – 5.0×10^6 cells was centrifuged at 1000 x g for 5 minutes and supernatant was discarded. Cell pellets were lysed by adding 300-1000 µL of room temperature Trizol® reagent, mixing the solution by pipetting up and down

several until obtaining a homogeneous solution. Samples were incubated for 5 minutes at room temperature for immediate processing or stored at -80°C for later processing (after thawing them in ice!). 200 µL of Chloroform per mL of Trizol® was added to each sample. Samples were shaken (or vortex) vigorously for 15 seconds, then incubated for 2-3 minutes at room temperature. For phase separation, samples were centrifuged at 12,000 x g for 15-20 minutes at 4°C. Aqueous phase (top) was carefully transferred to a clean micro centrifuge tube, avoiding drawing any interphase or organic layer into the pipette. For RNA precipitation, 500 µL per 1 mL of used Trizol® were added to the samples and incubated for 10 min at room temperature. Samples were centrifuged at 12,000 x g for 10 minutes at 4°C. After decanting the supernatant, RNA pellets were washed using 1 mL of 75% Ethanol (prepared with DEPC-treated Nuclease Free water) and centrifuged at 7,000 x g for 5 minutes at 4°C. This was done twice. After the second wash, all supernatant was carefully removed using a pipette, and pellets were air dried for 5-10 minutes, before being re-suspended in 20-50 µL of pre-warmed (55°C) Nuclease-free water (Ambion®). Samples were further incubated at 55°C for 5-10 minutes, then quantified using Nanodrop® (1.3.1) and stored at -20°C (short term) or -80°C (long term).

1.3. Nucleic Acid Quantification

1.3.1. Nanodrop® Spectrophotometer

DNA and RNA samples were quantified using the Nanodrop® spectrophotometer. The instrument was blanked using the appropriate re-suspension solution (TE buffer or Nuclease free water) prior to measuring the samples. For accurate measurement, each sample was measured 2-3 times and the average was calculated. When possible, samples were always measured in the appropriate range 2 ng/µL – 3000 ng/µL for accuracy.

1.3.2. Gene expression (SYBR)

1.3.2.1. Primer design

In order to quantify gene expression, specific primers for the gene(s) of interest were designed using the Primer3 in silico software (<http://primer3.ut.ee>) on the transcript sequence (Koressaar and Remm, 2007; Untergasser *et al.*, 2012). Usually 2-3 pairs of

primers were designed and ordered to Integrated DNA Technologies (IDT). On arrivals, lyophilized primers were re-suspended in nuclease-free water for a stock final concentration of 100 μ M following manufacturer's instructions, mixed and spun down. A working stock for each pair of primers was prepared by dilution the original stock 1:10 for a concentration of 10 μ M. Stocks of individual primers or pairs of primers were prepared. All primer stocks and working stocks were stored at -20°C. Before use, primers were thawed in ice, vortex for 10-15 seconds and spun down for proper re-suspension.

1.3.2.2. Primer testing.

Testing the specificity of a pair of primers for gene expression quantification (or other applications) was done by PCR reaction (1.1.1) on cDNA samples using the same RT-qPCR conditions. Samples were analysed by agarose electrophoresis (1.4.2) and pairs primers yielding a single band were selected. Optimization when required was performed in the same way.

1.3.2.3. cDNA Synthesis

For gene expression quantification, total RNA had to be reverse transcribed to cDNA. For that, the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) was used, which relies on random primers for the generation of single stranded cDNA molecules. The reverse transcription reactions were set up by preparing a master mix for the total number of reactions (plus 10-20% for pipetting errors), using the following reagents per reaction:

2.0	μ L RT Buffer (10X)
0.8	μ L dNTP (25X)
2.0	μ L RT Random Primers
1.0	μ L Multiscribe [®] Reverse Transcriptase
1.0	μ L RNase Inhibitor
3.2	μ L Nuclease-free H ₂ O
10.0	μ L Template total RNA (2 μ g)
<hr/>	
20.0	μ L Total volume (1 reaction)

Samples were run in a bench top thermocycler using the following cycling program:

	Temperature	Time
Step 1	25°C	10 min
Step 2	37°C	120 min
Step 3	85°C	5 min
Step 4	4°C	Hold

1.3.2.4. Real Time-quantitative PCR

Real Time-quantitative PCR was done using the Fast SYBR® Green Master Mix (Applied Biosystems). For each run, the appropriate control primers (typically for housekeeping genes such as GAPDH or Beta-Actin) were included, together with the target genes. This allowed to analyse the data and obtain a relative quantification of the genes of interest for the different targets and to assess that the same amount of template was loaded in all the wells. For one RT-PCR reaction, the following recipe was used:

10.0	µL Fast SYBR® Green master Mix (2X)
1.0	µL Forward and Reverse Primer Mix (10 µM)
1.0	µL cDNA template (directly from cDNA reaction, recommended 20 ng)
9.0	µL Nuclease-free H ₂ O
10.0	µL Template total RNA (2 µg)
3.2	µL Nuclease-free H ₂ O
<hr/>	
20.0	µL Total volume (1 reaction)

In order to reduce variability among samples due to reaction preparation, a master mix containing both the Fast SYBR® Green master Mix (2X) and the cDNA template, with a

20% excess to account for pipetting errors, was prepared for each condition to be analysed, taking into account the number of primers sets to be used. The mix was then split according to the number of primers, and each primer mix was added. Each reaction was prepared to be ran in triplicate and loaded in a MicroAmp® Fast Optical 96-well Reaction Plate (Applied Biosystems), which was sealed using an optical adhesive film, and centrifuge to ensure all sample was at the bottom of the plate before the measurements were performed. Samples were run in an Applied Biosystems 7500 Instrument, using the following fast cycling program:

Step	Temperature	Time	Cycles
AmpliTaq® Fast DNA Polymerase, UP Activation	95°C	20 sec	Hold
Denature	95°C	3 sec	40
Anneal/Extend	60°C	30 sec	

1.3.3. MicroRNA expression (TaqMan® Assays)

MicroRNA expression levels were analysed using the TaqMan® Small RNA Assay Kit (Applied Biosystems). This kit relies on stem-loop primers specific for each miRNA target for both steps, the reverse transcription and the RT-qPCR. Reverse transcription reactions are set up as follows. First prepare a RT master mix, using the following amount of reagent per reaction and adding 10-20% excess to compensate pipetting losses:

0.15	μL dNTPs (with sTTP) (100 mM)
1.0	μL MultiScribe® Reverse Transcriptase, 50 U/μL
1.50	Reverse Transcription Buffer (10X)
0.19	μL RNase inhibitor, 20 U/μL
4.16	μL Nuclease-free H ₂ O
<hr/>	
7.0	μL Total volume (1 rxn)

Keep RT master mix in ice.

To set up the RT reactions, split the RT master mix according to the number of RNA samples/miRNA targets to analysed, and mix with 5 µL of single stranded RNA sample previously adjusted to 20 ng/µL. Add 3 µL of 5X RT primers/ reaction. Incubate in ice for 5 minutes before loading the thermal cycler.

Run the samples using the following thermal cycler program:

Step	Temperature	Time
1	16°C	30 min
2	42°C	30 min
3	85°C	5 min
4	4°C	Hold

Amplification curves and PCR Efficiency were used as a measure to evaluate specificity and appropriate detection of miRNA with each particular assay. PCR efficiencies were calculated as follows: $PCR\ efficiency = 10^{(1/slope)} - 1$, when the logarithm of the initial template concentration (the independent variable) is plotted on the x axis and Ct (the dependent variable) is plotted on the y axis (Bustin *et al.*, 2009).

1.3.4. Methods for mRNA relative quantification after RT-qPCR

In order to determine the relative amounts of transcripts between samples/conditions, Ct values were used to calculate the Relative Quantification (RQ) using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). This method allows the calculation of the relative amounts for each transcript in relation to a reference control sample/condition. The method also uses Ct values of an endogenous control known not to change between samples or conditions. GAPDH, B-Actin and GUSB were routinely used in this work as endogenous controls. RQ values are calculated using the following formula:

$$RQ = 2^{-\Delta\Delta Ct}$$

where $-\Delta\Delta Ct = (Ct_{target} - Ct_{endo})_{Condition\ 1} - (Ct_{target} - Ct_{endo})_{Reference\ condition/sample}$

For samples obtained during the actinomycin D experiments, it was not possible to have an endogenous control gene that would not change upon treating cells with actinomycin D. For the relative quantification of mRNA transcripts in these experiments a variation of the above-mentioned method named $2^{-\Delta Ct}$ was used where:

where $-\Delta Ct = (Ct_{\text{target}})_{\text{Time 1}} - (Ct_{\text{target}})_{\text{Time 0}}$

1.4. Nucleic Acid Quality Control

1.4.1. Nanodrop® Spectrophotometer

Quality of the purified nucleic acids (DNA and/or RNA) was assessed at the same time as the quantification by Nanodrop®. The instrument calculates two quality ratios: A260/A280 for protein contamination and A260/A230 for organic compounds contamination. Pure and good quality nucleic acids were expected to have an A260/A280 between 1.8-2.0 and a A260/A230 between 1.8-2.3.

1.4.2. Agarose Electrophoresis

For a qualitative evaluation of nucleic acid quality and size (for isolated and purified fragments of specific sizes, i.e. after restriction digestion), an agarose electrophoresis separation was performed. 0.8-1.2% agarose (Sigma Aldrich) was dissolved in the appropriate volume of TAE buffer (1X) (Thermo Scientific), heated up until boiling in a microwave and let cool down to handling temperature. Before pouring the gel on the frame, 1 µL of SafeView nucleic acid staining (NBS Biologics) per 100 mg of agarose used was added to gel solution and mixed by gently swirling.

1.5. Nucleic Acid Modification

1.5.1. Restriction Enzyme Digestion

DNA can be cut in a sequence specific by restriction enzymes, which are able to generate either complementary “sticky” end, which can be ligated together, or blunt ends, which can be ligate independently of the sequence but with a lower efficiency. All enzymes used here were purchased from New England Biolabs or Fisher Scientific and used with their specific buffers and incubation temperatures and times. Typically, analytical restriction digestions were set up in 20 µL reactions with up to 1 µg of DNA and less than 10% of the volume between the different restriction enzymes (to avoid Glycerol

inhibition) and directly analysed by agarose electrophoresis (1.4.2). Preparative restriction digestions for further purification of the desired DNA fragments were set up in 50-100 μ L. Of the obtained digested sample, 10 μ L were run in an agarose gel (1.4.2) and the remaining processed for gel purification.

1.5.2. Ligation

All ligation reactions were performed using the T4 Ligation Kit (Roche). The online tool NEBioCalculator (<https://nebiocalculator.neb.com/#!/ligation>) was used to calculate the required mass ratios of vector:insert (typically 1:3 or 1:5). Ligations reactions were set up as follows.

2.0	μ L T4 Ligase Buffer (10X)
X	μ L Vector DNA (50.0 ng)
X	μ L Insert DNA (variable amount)
X	μ L Nuclease Free H ₂ O
1.0	μ L T4 DNA Ligase
<hr/>	
20.0	μ L Total volume

Samples were incubated at room temperature for 10 minutes or overnight in thawing ice. For ligations involving annealed oligos, a 1:3 dilution of the original annealing reaction (ca. 500 ng/ μ L) was used. Ligation reactions were typically used for bacterial transformation (SEE REF) or stored at -20°C.

1.5.3. Dephosphorylation

Some cloning required dephosphorylating of the digested vector to ensure no self-re-ligation. This was performed using Fast AP (Thermo Fisher), following the manufacturers protocol or sequentially after the restriction enzyme treatment to avoid double purification been needed.

1.5.4. Oligo Annealing

For some of the required DNA fragments i.e. miRNA-based toehold switches or miRNA sponges, individual single stranded and complementary oligonucleotides (50-100 bp) had to be designed ordered on IDT. Each pair of oligos was then annealed in house. For that, lyophilized oligos were re-suspended in Nuclease free water for a stock final concentration of 100 μ M, mixed and spun down. All prepared oligos were stored at -20°C. Annealing reaction were prepared as follows:

1.0	μ L Oligo 1 (100 μ M)
1.0	μ L Oligo 2 (100 μ M)
1.0	T4 Ligase buffer (10X) (Roche)
6.5	μ L Nuclease Free H ₂ O
0.5	T4 PNK (NEB)
<hr/>	
10.0	μ L Total volume

Reaction was incubated at 37°C for 30 min, for the T4 PNK to phosphorylate the 5' ends. Samples were then incubated at 100°C in a heat block for 5', then spun down and replaced in the heat block, that was switched off to allow for a slow cooling down along 2-3h, until the temperature dropped below 40°C. Samples were then stored at -20°C until required.

1.5.5. miRNA sponge construction

miRNA sponges were constructed based on the method published by Kluiver *et al.* (2007) and optimized *in house* (Costello *et al.*, 2017). In short, complementary DNA oligonucleotides containing two miRNA binding sites separated by a AATT (5'->3') spacer and flanked by cut *KflI* (*SanDI*) restriction sites (GGGTCCC) were ordered to Integrated DNA Technologies Ltd (<https://eu.idtdna.com/>) and annealed as per Methods 1.5.4. Complementarity of the miRNA binding site to the target miRNA varied depending on the experiment, with fully-complementary binding sites, in which all nucleotides in the binding site match the sequence of the miRNA, and bulged binding sites, in which the

nucleotides 10-11' of the binding sites are changed to create a mismatch with the miRNA in that region, aiming to impair DICER nuclease activity.

A backbone vector (CMV-de2GFP or EF1a-de2GFP depending on the experiment) was digested with *Kfl*I (*San*DI) as per Methods 1.5.1 and ligated with 1 μ L of a 1:3 dilution from the annealed oligo stock as per Methods 1.5.2. The resulting ligation mix was transformed in *E. coli* as per Methods 1.1.3. to obtain clones containing sponges of different sizes.

1.5.6. Colony PCR for miRNA sponge screening

In order to pre-screen transformed *E. coli* clones and select miRNA sponges of the desired size while excluding possible empty vector transformants, colony PCR was used. In short, 5-10 colonies per sponge depending on the needs were picked and each re-suspended in 10 μ L of nuclease-free water. PCR reactions were prepared as per Methods 1.1.1 using 1 μ L of bacterial suspension as template. The following primers: FW: GGGATCACTCTCGGCATGGA and RV: GGCATGCATCGAGAATTCCG were designed to bind the flanking region of the *Kfl*I (*San*DI) site in the vector and were used to amplify the sponge insertion region using the standard PCR cycling conditions (Methods 1.1.1). PCR reactions were run in a 2% agarose gel as per Methods 1.4.2 in order to visualize bands which size depended on the size of the ligated sponge (**Figure 4**).

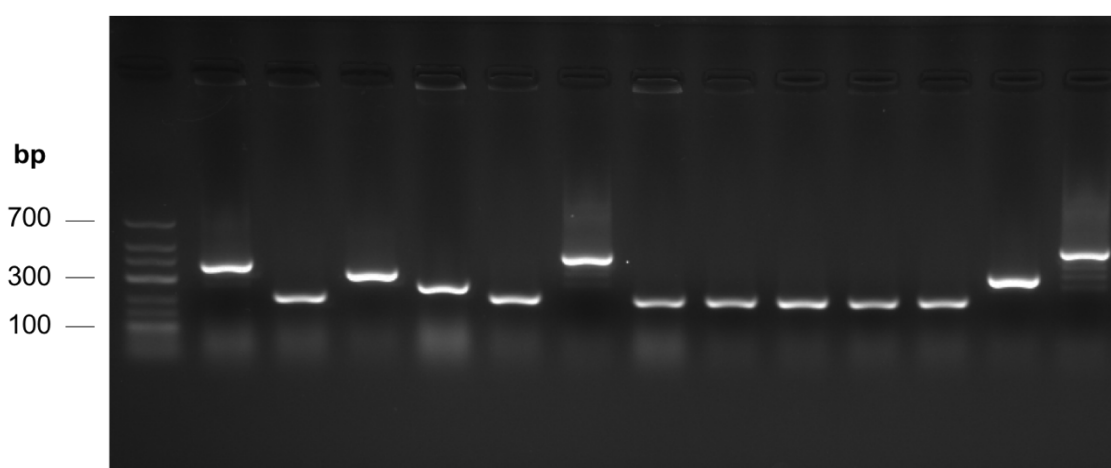


Figure 4 - Example of PCR colony screening of *E. coli* transformed with a miRNA sponge ligation in a 2% agarose gel

2. Cell Culture Techniques

2.1. General Cell Culture Good Laboratory Practices

All cell culture work, including reagent preparation that required sterility, was performed under a class II Laminar air-flow cabinet (LFC) (Holten). Before use, the LFC was let stabilise for 15 minutes and wiped using 70% Industrial Methylated Spirits (IMS) for sterilization. Anything entering the LFC was sprayed with the same 70% IMS, including glove protected hands. When possible, mainly when handling clonal stable cell lines or stock seed cultures, cell culture work was performed with a single cell type at a time, to avoid cross-contamination. Regular monthly maintenance, including full clean and disinfection of the LFC using Virkon® and 70% IMS and yearly certification were performed to ensure proper function and sterility inside the LFC.

2.2. Cell Line Sub-Culturing

Routinely, cells were grown in adherent culture or suspension culture.

For adherent culture, cells were grown in Tissue-culture Flasks of different sizes (25 cm² – 75 cm²) or Tissue culture plates (6 well plates) in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12) with 1-5% Foetal Calve Serum (FCS), and incubated in a static incubator set at 37°C, 5% CO₂ and 80% humidity. Cells were allowed to grow to ca. 80-90% confluence and were typically passaged twice a week (every ca. 3.5 days), by trypsinizing the flask using a Trypsin-EDTA solution, 5-10 minutes at 37°C. Fresh media containing serum was added to inhibit the trypsin and cells were centrifuged at 1,000 x g for 5 minutes. Supernatant containing trypsin was discarded and cell pellet were re-suspended in of fresh medium for cell density determination. New flasks were typically seeded mixing 9:10 parts in volume of fresh media with 1:10 parts in volume of harvested cells.

For suspension culture, cells were grown in either, 24-well suspension plates for 1 mL cultures (Corning) and transfections, 50 mL vented cap spin tubes (TPP) for up to 5 mL cultures or in 250 mL vented cap shake flask (Corning) for cultures ranging from 10 to 60 mL.

Volumes for the different flask were as follows:

Flask Size	Culture media	PBS	Trypsin-EDTA	Stop trypsin	Re-suspension media
24-well	0.5 mL	1 mL	0.5 mL	0.5-1 mL	0.5-1 mL
6-well	1-2 mL	1 mL	0.5 mL	1-2 mL	1-2 mL
T25	3 mL	3 mL	1 mL	3 mL	1-3 mL
T75	9 mL	10 mL	3 mL	9 mL	1-5 mL

All cultures were maintained in a Kühner Climo-Shaker ISF1-X incubator set up at 37°C, 5% CO₂, 80% humidity and 170 rpm. Routinely, cells were grown in serum-free media. Typically CHO-S SFMII (Gibco) or BalanCD CHO Growth A (Irvine Scientific) with or without 1% Poly-vinyl alcohol (PVA) as anti-clumping agent were used. Cells were seeded at 2×10^5 cells/mL and passaged every 3-4 days to maintain them growing and in mid-exponential phase.

2.3. Cell Density and Viability Monitoring

2.3.1. Trypan blue

Cell counts with viable/dead cells were obtained using a haemocytometer and using trypan blue staining. An aliquot of the original cell suspension was diluted to a known dilution factor using PBS and mixed 1:1 with trypan blue (Gibco). Cells were counted by loading a 10-20 μ L of cell dilution/trypan mix into the haemocytometer chamber, and counting the cells present in the 4 big squares (16 small squares) of the grid. The following formula was then used to calculate viable cell density and percentage of viable cells:

Viable Cell Density = Average N° of cells per big square * dilution factor / Volume of small square (mL)

Volume of small square = Width x Height x Depth (typically: Width = Height = 1 mm depth = 0.1 or 0.2 mm)

Viability (%) = $[1 - (\text{N}^\circ \text{ of blue cells} / \text{N}^\circ \text{ of white cells})] * 100$

2.3.2. Flow cytometry

For large amounts of samples, cell density and viability were assed using a GUAVA EasyCyte® bench-top flow-cytometer (Merck) and the Viacount assay. The Viacount dye was prepared in house using LDS-751 (Invitrogen), a membrane permeant fluorescent dye to stain all cell nuclei for a total cell count and Propidium Iodide (PI) (Invitrogen), a membrane-impermeant fluorescent dye to specifically stain membrane compromised cells to assess non-viable cell number. The Viacount reagent was prepared as follows:

For 100 mL of Viacount Dye:

400 µL of Propidium Iodide

200 µL of LDS-751

100 mL of PBS

* Store at 4°C and protected from light.

Samples for the Viacount assay are prepared in U-shaped bottom 96 well plates, were cell culture aliquots are diluted in PBS and mix 1:1 with Viacount dye (i.e. for a 1:4 final dilution, 50 µL cells, 50 µL PBS and 100 µL of Viacount). Cell density in the measured samples needs to fit between 10-500 cells/µL for accurate cell number determination by the instrument.

2.4. Cryopreserved Cell Lines Revival

Cells were stored at -80°C for mid-term storage and in Liquid Nitrogen for long-term storage. In order to bring back cells to culture, a cryovial was taken from -80°C (or liquid Nitrogen) and immediately placed on ice. Cells were quickly thawed by adding drop wise pre-warmed (37°C) medium on the frozen cells and transferring it to 5 mL of pre-warm media using a Pasteur pipette. Thawed re-suspended cells were then centrifuged at 1,000 x g for 5 minutes, supernatant containing cell debris and DMSO discarded, and the cell pellet was re-suspended in 5 mL fresh culture media and transferred into the desired culture vessel.

2.5. Cell Line Cryopreservation

For the cryopreservation of both parental and generated transgenic cell lines, the following protocol was used. Cells to be frozen were passaged, either split 1:2, or centrifuged at 1,000 x g for 5 minutes and re-suspended in fresh medium). On the day

of freezing, freezing media containing 20% DMSO (Sigma-Aldrich) and no anti-clumping agent (+ 20% FCS for adherent cell lines) was prepared beforehand and kept in ice until used. Typically, $2.5\text{--}10 \times 10^6$ cells/cryovial were harvested, centrifuged at $1,000 \times g$ for 5' minutes and re-suspended in 0.5 mL/cryovial ice cold conditioned media (or fresh media). 0.5 mL of cell suspension was placed in each cryovial. 0.5 mL of ice-cold freezing media was added on top of the cells. Final concentrations in the cryovial were: 10% DMSO (+10% FCS for adherent cells). Cryovials were stored in a polystyrene box and placed in a -80°C freezer, to ensure a slow freezing down process (ideally $-1^\circ\text{C} / \text{min}$). In the absence of polystyrene boxes, cryovials were placed for 2 hours in a -20°C freezer prior to be transferred to the -80°C .

2.6. Transient Transfection

Transient transfections were performed using the TransIT-X2 Dynamic Delivery System (Mirus Bio). Transfections were performed in 24-well suspension plates. Cells to be transfected were seeded in fresh media 24 hours before transfection at a cell density of $0.5\text{--}1 \times 10^6$ cells/mL and incubated overnight. On the day of transfection, the TransIT-X2 reagent and the DNA were pre-complexed in pre-warmed media (no anti-clumping agent!) at concentration suited for the experiment (typically 2 μL of TransIT reagent per 1 μg of DNA and 250-1000 μg DNA / 10^6 cells). Cells were counted and cell concentration adjusted using fresh to the desired amount (typically 1×10^6 cells/mL). Cells were then plated adding 900 μL of cell suspension per well. 100 μL of pre-complexed TransIT-DNA mix was added drop wise on top of the cells and carefully mixing the well by swirling the pipette tip (no pipetting up/down!). Plates were incubated under standard cell culture conditions (Section 2.2). Transfection experiments involving GFP expressing plasmids were analysed by flow cytometry typically 24-48 hours after transfection. **Modified transfection protocol for actinomycin D treatment:** Cells were seeded at 2.5×10^5 cells per well in 6-well tissue culture treated plates in 2 mL of SFM II media + 5% FBS and incubated overnight in a static incubator at 37°C , 5% CO_2 and 80% Humidity. The next day, cells were washed twice with pre-warmed up sterile PBS buffer to remove traces of conditioned media (negatively affect transfection) and 1 mL of pre-warmed SFM II + 5% FBS media per well was added. Complexes were formed as for regular transient transfections (see above) in SFM II media **without FBS** using a total DNA amount of 1 μg / well with 50:50 GFP:Empty Vector.

2.7. Stable Cell Line Generation

Stable cell lines were generated from some of the transiently transfected cell lines generated as described in Section 2.6. To do that, the contents of replicate transfections (typically 2-4 wells and only when low variability was seen across all replicates) were transferred 24 hours after transfection to a 6-well plate or a T25 flask containing fresh media with 5% FCS and appropriate amount of selective antibiotic. Cells were expanded under selection pressure to a T75 scale. At this stage, a cryo-bank was made (see Cell Line Cryopreservation). Selection in adherent culture was typically performed for a period of 2+ weeks. A new cryo-bank of selected stable cell lines was made at that point. If cells were to be used for suspension culture, an aliquot was transferred to a 50 ml vented cap spin tube in the appropriate suspension media, and adapted to grow under standard suspension culture conditions (Methods 2.2) for at least 3 passages (or up to a consistent cell growth pattern was observed) before using them.

2.8. GFP detection and quantification

In order to evaluate transfection efficiencies (% of fluorescent cells) or transgene expression (Mean Fluorescence Intensity – MFI) when using the GFP reporter gene, the *Express Plus* program of the GUAVA EasyCyte® bench top flow cytometer was used. Gates for dark and bright cells were defined using dark untransfected cells and bright positive control cell populations. All samples were analysed using the same settings for comparability of the obtained results. Cells to analyse were diluted in PBS and plated in U-shaped bottom 96-well plates to a concentration between 10-500 cells/ μ L when possible. Typically, 3000-10000 cells were analysed per sample.

2.9. RNA decay assay – Actinomycin D treatment

In order to determine RNA stability, cells were treated with actinomycin D (Gibco), a transcriptional blocker, and samples were taken at different times between 0 and 24h. In order to determine the stability of transcripts expressed from transiently transfected plasmids, the transfection protocol was adapted as described in Methods 2.6. as, in order to minimize centrifugation steps with actinomycin D contaminated samples, cells were transfected in adherent cultures. For the actinomycin D treatment, actinomycin D containing media was freshly prepared on the day by diluting the stock of actinomycin D (1 mg/mL) 1:200 to a final concentration of 5 μ g/mL in SFM II media + 5%FBS. All work

with concentrated actinomycin D drug was carried out in a CytoGuard laminar flow safety cabinet. Typically, 1.5 mL of actinomycin D media per well plus the amount for an extra well was prepared. At time 0 (typically 24h post-transfections), old media was removed from the cultures and 1.5 mL/well of actinomycin D media was carefully added. For optimization experiments, not treated wells were included. In order to take down cells after the different treatment time, supernatant was removed and disposed in a drug waste container, and 1 mL of Trizol® reagent was directly added to each well. Cells were incubated on Trizol for 1-2 minutes at room temperature before thoroughly mixing up and down the content of each well in order to ensure appropriate cell lysis before transferring it to a 1.5 mL micro centrifuge tube and store at -80°C until processed as per Methods 1.2.2.

3. In silico Tools

3.1. Nucleic Acid Secondary Structure Prediction

In order to predict secondary structures of nucleic acid molecules (i.e. miRNA-toehold switches), different online tools were used and are listed below:

- **RNAstructure** – <http://rna.urmc.rochester.edu/RNAstructureWeb/index.html>
- **IDT Oligo Analyser Tool** - <https://eu.idtdna.com/calc/analyzer>

3.2. Plasmid Cloning/ Maps

In silico maps of the different constructed plasmids were created using two license-free software:

- **SnapGene Viewer** – used to generate the plasmids maps
- **Serial Cloner** (http://serialbasics.free.fr/Serial_Cloner.html). This software was used to edit sequence in plasmid maps, to simulate cloning procedure before performing the actual experiments in the lab as well to predict expected band pattern following restriction digestion (using the Virtual Cutter function).

3.3. Differential Analysis of miRNA expression

Differential expression analysis was performed using raw count data imported into the DeSeq2 software package within the R statistical framework. Scripts are available in the digital version of this thesis and upon request.

3.4. Unsupervised hierarchical clustering

Unsupervised hierarchical clustering was as well performed from raw count data by using the DEGReport software package within the R statistical framework. Scripts are available in the digital version of this thesis and upon request.

3.5. miRNA Target Prediction and Pathway Analysis Tools

Target prediction and functional pathway enrichment analysis were performed using the miRPath3.0 algorithm from the Diana Tools suite. In short, the shortlisted miRNAs were uploaded to the web-based platform which searches the TarBase 7.0 for predicted and validated targets for each miRNA.

The output of the miRPATH algorithm is shown in consists a list of the significantly enriched pathways arising from predicted targets for the given list of miRNAs.

Link: ***<http://snf-515788.vm.okeanos.grnet.gr/index.php?r=mirpath>***

CHAPTER 3

Identification of endogenous miRNA as
potential gene switches in CHO cells

1. Introduction

Over the last decade miRNAs have been extensively profiled in CHO cells using different technologies such as TILDA cards, microarrays and RNASeq and several CHO microRNA expression datasets generated under different conditions have been published. To date, most of the studies involving CHO miRNA profiling have shared the same main focus with regard to studying miRNA expression: the aim is the identification of miRNAs that could be engineered to promote a particular industry relevant phenotype, namely an increase in growth, productivity, product quality and/or a reduction of by-product formation and cell death. Traditional approaches using differential expression analysis on microarray or RNASeq datasets comparing fast and slow cell lines or low and high producers have a remarkable track record in identifying suitable miRNA candidates that have been then engineered and have resulted in some cases in improved cellular phenotypes (N. Barron *et al.*, 2011; Druz *et al.*, 2013; Fischer, Mathias, *et al.*, 2015; Fischer, Paul, *et al.*, 2015; Kelly *et al.*, 2015).

As the toolbox of molecular methods using miRNAs as effector molecules for the control of transgene expression keeps growing, the need to identify miRNA candidates with robust, predictable and even controllable expression profiles is becoming more and more apparent (Valdés-Bango Curell and Barron, 2018). Endogenous changes in cell metabolism are tightly regulated events which encompass synchronized action of various regulatory effectors at a transcription level, involving transcription and epigenetic factors, as well as effectors at a translational level, involving all the different elements of the translational machinery which include elongation factors, ribosome binding proteins among others and microRNAs, which play a role in the regulation of translational events as previously described. Therefore, it is reasonable to hypothesize that in this context, it should be possible to identify microRNAs whose expression changes with the depletion of key metabolites or the build-up of certain by-products over time. In this first chapter, two miRNA expression datasets were analysed with the aim to identify and potentially validate miRNAs exhibiting expression profiles that might be used as effector molecules for transgene expression control using miRNA-based genetic control elements in the context of biopharmaceutical manufacturing and investigate the factors driving their expression pattern.

Three potential control factors or triggers were investigated in this work – cell metabolites (glutamine/ammonia concentrations), cell culture stage (time/growth phase), which will be presented in this chapter, and temperature, a central bioprocess parameter, that has previously been modified in order to improve culture outputs, and has been investigated in the work included in Chapter 4.

For the work presented in this chapter two miRNA expression datasets were analysed. The first one was generated as part of a larger study involving microarray profiling of various CHO cells lines and was kindly provided by Prof. Nicole Borth's group in BOKU, (Vienna, Austria) (Klanert *et al.*, 2016). The second dataset was generated as part of the consortium eCHO Systems and consists of miRNA expression over the course of a fed-batch culture of a CHO producer cell line under industrially relevant conditions.

2. Results

2.1. Glutamine metabolism as guiding a differential factor for miRNA profiling

The investigated dataset comprises time course microarray expression data for 4 CHO cell lines grown in 2-L stirred tank fermenters and serum free media. This dataset is a subset from data generated as part of a larger CHO miRNA profiling study published by Klanert *et al.* in 2016 and from which a robust 12 miRNA signature correlating with growth rate was identified (Klanert *et al.*, 2016). In addition to miRNA expression data, the dataset contains process data included: viable cell density, viability, growth rate, and concentration for the main bioprocess relevant metabolites: glucose, lactate, glutamine, glutamate and ammonia. Looking at the data on the production and consumption of key metabolites (Figure 5), the presence/absence of L-glutamine in the culture media was identified as one of the main differential factors between the analysed cell lines.

Two of the cell lines were parental CHO-K1, one glutamine-dependent (K1-FERM-8-SF-HOST) and the other adapted to glutamine-free conditions (K1-FERM-0-SF-HOST). These cell lines were generated as part of a study on adapting CHO cells to glutamine –free conditions using a FACS-based evolution strategy (Bort, Stern and Borth, 2010).

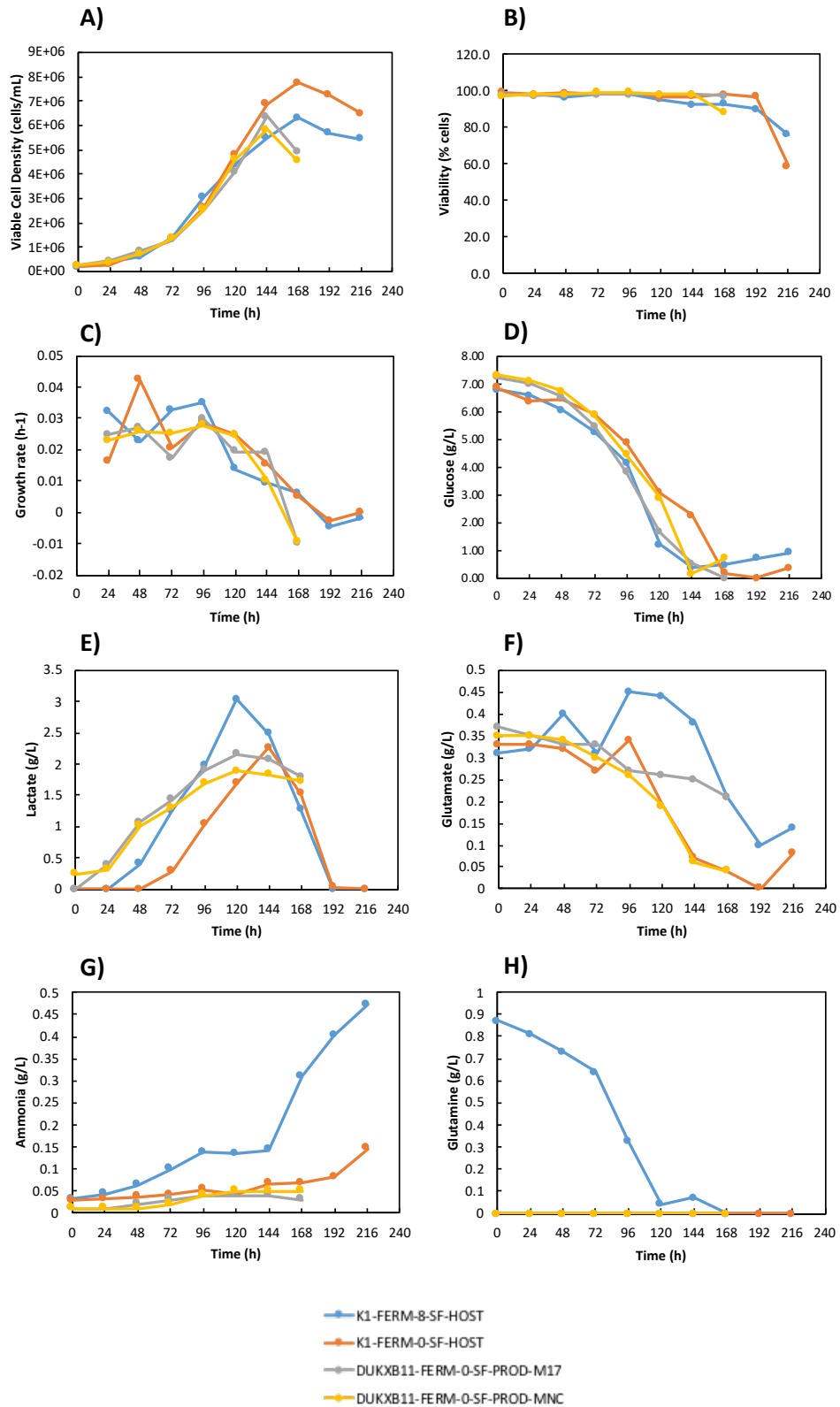


Figure 5 - Growth profiles and culture parameters for the 2 CHO-K1 and the 2 DUKXB11 bioreactor runs included in the analysed microarray dataset: A) Viable Cell Density, B) Viability, C) growth rates (h⁻¹), D) Glucose concentration (g/L), E) Lactate concentration (g/L), F) Glutamate concentration (g/L), G) Ammonia concentration (g/L) and H) Glutamine concentration (g/L) (Klanert et al., 2016). All cells were grown in the 1L STR in either serum free media containing 8 mM glutamine (8-SF) or serum free media containing 8 mM glutamine (0-SF).

The other two cell lines were transgenic DUXB11-derived cell lines, originated from a DUXB11 derived cell line producing recombinant EPO-Fc and adapted to growth in glutamine-free media using the same strategy published by Bort *et al.* (Bort, Stern and Borth, 2010; Taschwer *et al.*, 2012). In particular, these cell lines, named DUXB11-FERM-0-SF-PROD-M17 and DUXB11-FERM-0-SF-PROD-MNC, were generated for a study on the effect of overexpressing miR-17 (Jadhav *et al.*, 2014). All cell lines were grown in stirred tank reactors (STR) and exhibited very similar growth profiles in terms of cell density (Figure 5 – A), viability (Figure 5 – B), growth rates (Figure 5 – C), glucose consumption (Figure 5 – D) and lactate (Figure 5 – E). In contrast, major differences were observed in glutamine and glutamate consumption and ammonia production between cell lines grown in glutamine-free conditions and the cell line grown in 8 mM Glutamine (Figure 5 – F, G and H).

2.1.1. Correlation analysis allows the identification of a subset of miRNAs exhibiting high correlation with ammonia concentration

Expression data for all miRNA in all cell lines was compiled for analysis. Pearson's correlation coefficients (PCC) between miRNA expression (Log2 fold-changes) and each different process parameter were calculated. miRNA exhibiting a Pearson's correlation coefficient of ± 0.6 were counted for each process parameter and counts are shown in **Figure 6**.

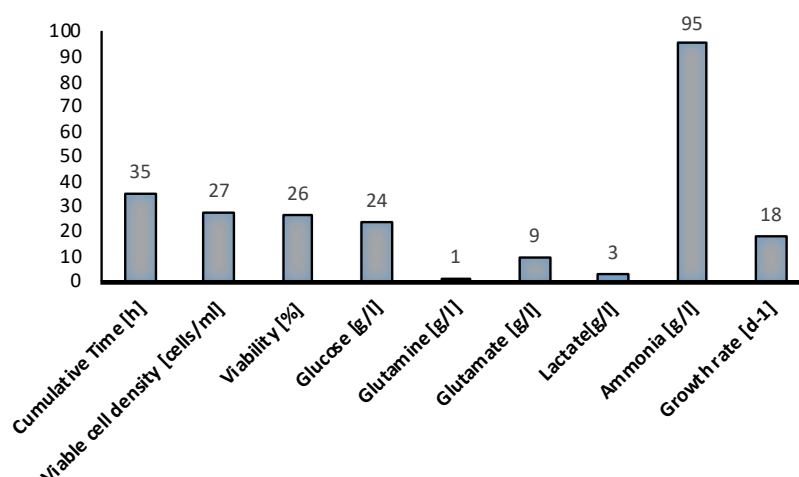


Figure 6 – Number of miRNA targets above the ± 0.6 Pearson's Correlation coefficient threshold for each correlated cell culture parameter.

Ammonia concentration in the media appeared as the parameter with the most highly correlating miRNAs, with 95 out of the 277 probed targets exhibiting Pearson's

correlation coefficients above the cut-off. Time was the second most correlated parameter, with 35 out of the 277 miRNA showing correlation coefficients above the cut-off, followed by viable cell density, viability and glucose concentration with 27, 26 and 24 highly correlated miRNAs respectively. miRNAs exhibiting a PCC > 0.75 with ammonia concentration were considered to have a high correlation with ammonia concentration and were selected. To further filter down the list of miRNA candidates, difference between the maximum fold-change and the minimum fold-change for each miRNA were considered as a measure of the amplitude of the expression change. A cut-off of $[\text{Log2FC}_{\text{Max}} - \text{Log2FC}_{\text{Min}}] > 1.5$ was applied to select the top miRNA candidates. Originally, all the selected miRNA candidates were to undergo RT-qPCR validation. However, as explained in the sections below, only a few could finally be tested and are highlighted (Table 2).

Table 2 - Selected miRNA strongly correlating with ammonia concentration across the 4 analysed cell lines. A cut-off of PCC > 0.75 and Log2FC difference > 1.5 were applied. miRNAs that were selected for validation by RT-qPCR are highlighted in grey.

Probe ID/Array ID	Pearson's Correlation Coefficient	P-value
rno-miR-187*	0.95	2.32E-10
hsa-miR-187*/mmu-miR-187*	0.93	1.14E-08
rno-miR-344b-1-3p	0.90	1.72E-07
mmu-miR-1952	0.89	2.88E-07
mmu-miR-3103	0.87	1.41E-06
hsa-miR-483-3p	0.86	3.37E-06
hsa-miR-3685	0.82	1.80E-05
hsa-miR-638	0.81	2.97E-05
hsa-miR-1275	0.81	3.06E-05
hsa-miR-4290	0.80	4.58E-05
mmu-miR-489	-0.77	1.31E-04
hsa-miR-183/mmuc-miR-183/rno-miR-183	-0.80	4.86E-05
mmu-miR-709	-0.82	2.09E-05
hsa-miR-138-1*/mmu-miR-138-1*/rno-miR-138-1*	-0.85	4.54E-06
hsa-miR-3613-3p	-0.86	2.36E-06
hsa-miR-302c*	-0.89	4.64E-07

2.1.1.2. Removing glutamine from the media slows down cell growth but does not impact cell viability in CHO-K1 cells

As mentioned before, the original glutamine-free adapted cell line had been adapted by selecting surviving cells by FACS after progressively reducing L-glutamine concentration in the media (Bort, Stern and Borth, 2010). In the absence of the original cells used to generate the above-mentioned dataset, an in-house CHO-K1 cell line, adapted to glutamine-free conditions using a chemically defined media had to be generated. In order to assess the impact on cell growth and viability of a sudden reduction of L-glutamine concentration in the media, CHO-K1 cells were seeded in media containing a range of L-glutamine concentrations, decreasing from 8 to 0 mM and cell density and viability were monitored over time. As a preliminary test, no biological replicate experiments were performed and only observational trends are described.

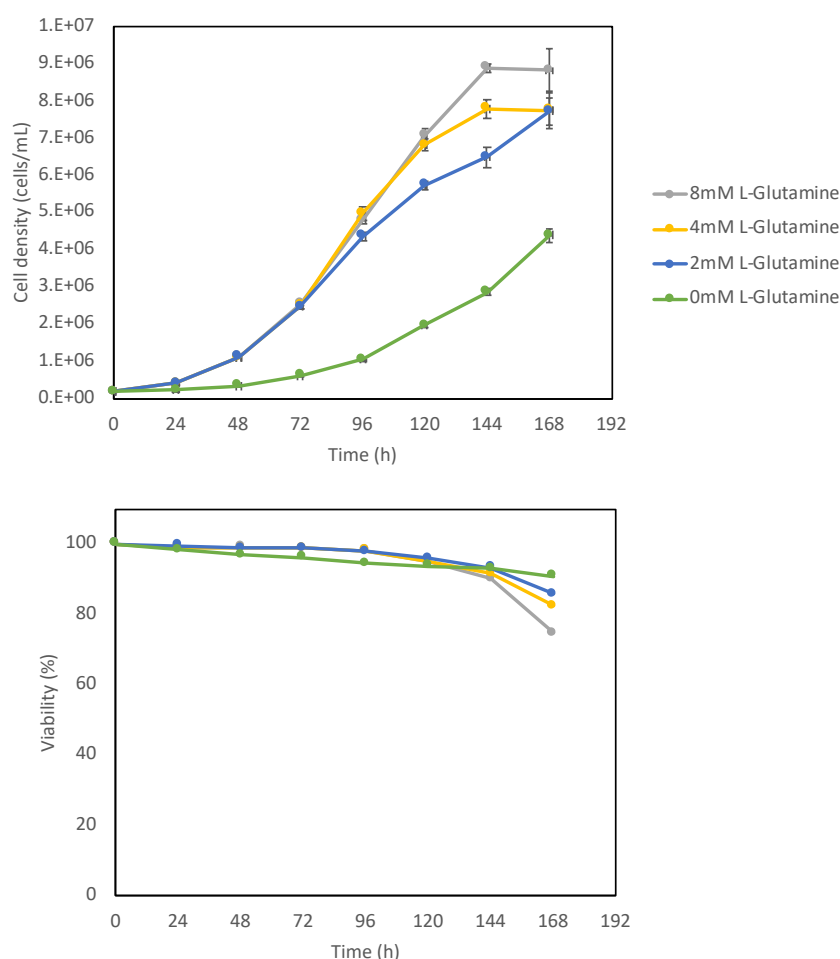


Figure 7 – Effect on cell density and viability of glutamine-reduction in BalanCD culture media on CHO-K1 cells. BalanCD media containing different L-glutamine concentrations from 0 to 8 mM were used. Viable Cell Density (top) and viability (bottom) measurements were taken every 24h. Each point represents the average of triplicate technical replicates (N=1) and error bars represent standard deviation.

As shown in Figure 7, cells growing in media containing down to 2 mM of L-glutamine did not show differences in growth during the first 72 hours of culture and then reached higher cell densities, the higher the initial concentration of L-glutamine was. Cells seeded in media without L-glutamine exhibited a slower growth profile and reached lower cell densities. Interestingly, cell viability remained above 90% for all conditions until the late stages of the culture, even in the cultures without L-glutamine.

Considering cells did not die in L-glutamine-free media but instead drastically slowed growth, it was decided to attempt the adaption process by using a single glutamine reduction step followed by continuous sub-culturing.

2.1.3. A one-step glutamine reduction approach followed by continuous sub-culturing generates glutamine-free adapted CHO cell line with minimal impact on cell viability

Having observed that a sudden reduction of L-glutamine from 8 mM to 0 mM did not result in a decrease of cell viability, a one-step reduction strategy was attempted to adapt CHO-K1 cells to glutamine-free conditions. The aim was to obtain adapted CHO-K1 cells that able to grow in glutamine-free conditions at a comparable rate to the one observed in glutamine containing media. CHO-K1 cells were seeded at a standard cell density (2×10^5 cells/mL) in BalanCD media with 2% PVA without glutamine and were sub-cultured in that media for up to 17 passages, every 3-5 days and with cells being kept as much as possible in exponential growth phase. The original CHO-K1 population was maintained in BalanCD media containing 8 mM L-glutamine and was sub-cultured and sampled at the same time points as a control. When possible, before each passage, cultures were sampled for cell density and viability measurements.

Figure 8 shows growth rate, doubling time and viability data for the CHO-K1 cells maintained in glutamine-free media and the CHO-K1 cells maintained in glutamine-containing media over the course of 18 passages. Viable cell density data has been used to calculate growth rates and doubling times.

CHO-K1 cells grown in the absence of glutamine exhibited lower growth rates and higher doubling times during the first 2-3 passages of the adaptation process. Interestingly, from passage 4, growth rates progressively increased, reaching values in a 10% range of the ones observed for CHO-K1 cells growing in 8 mM L-glutamine after 14 passages

(Figure 8). After 17 passages cells were considered adapted and a growth curve experiment was set to compare growth profiles, maximum cell density, viability and ammonia production for both cell lines (Figure 9).

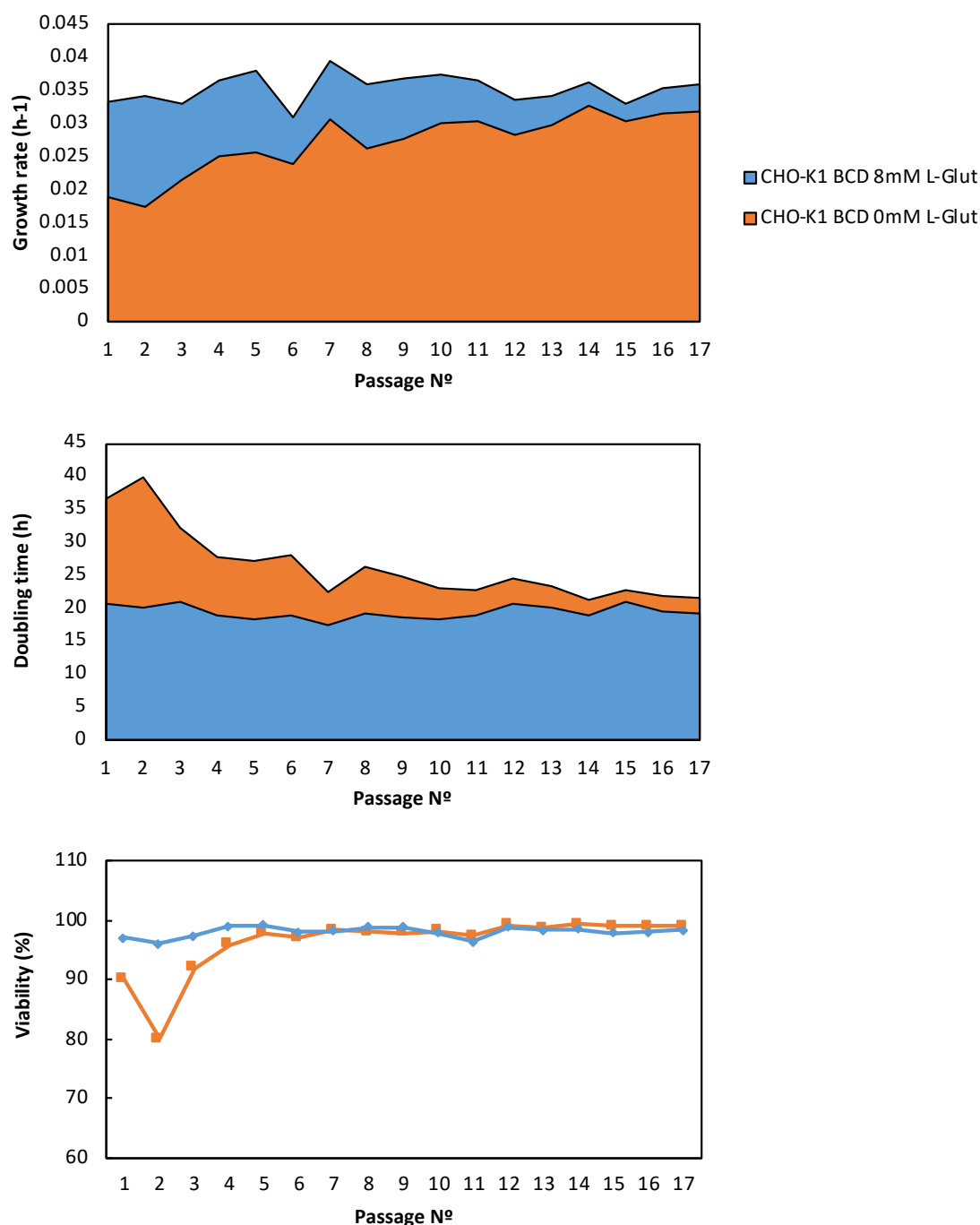


Figure 8 – CHO-K1 cells adapt to glutamine-free conditions. Growth rates (top), doubling times (middle) and cell viability (bottom) for the two CHO-K1 cell lines grown in 8 mM L-Glutamine (blue) and 0 mM L-Glutamine (orange) calculated over the adaptation period (17 passages). Each growth rate/doubling time was calculated from day 0 to the last day of each passage. N=1.

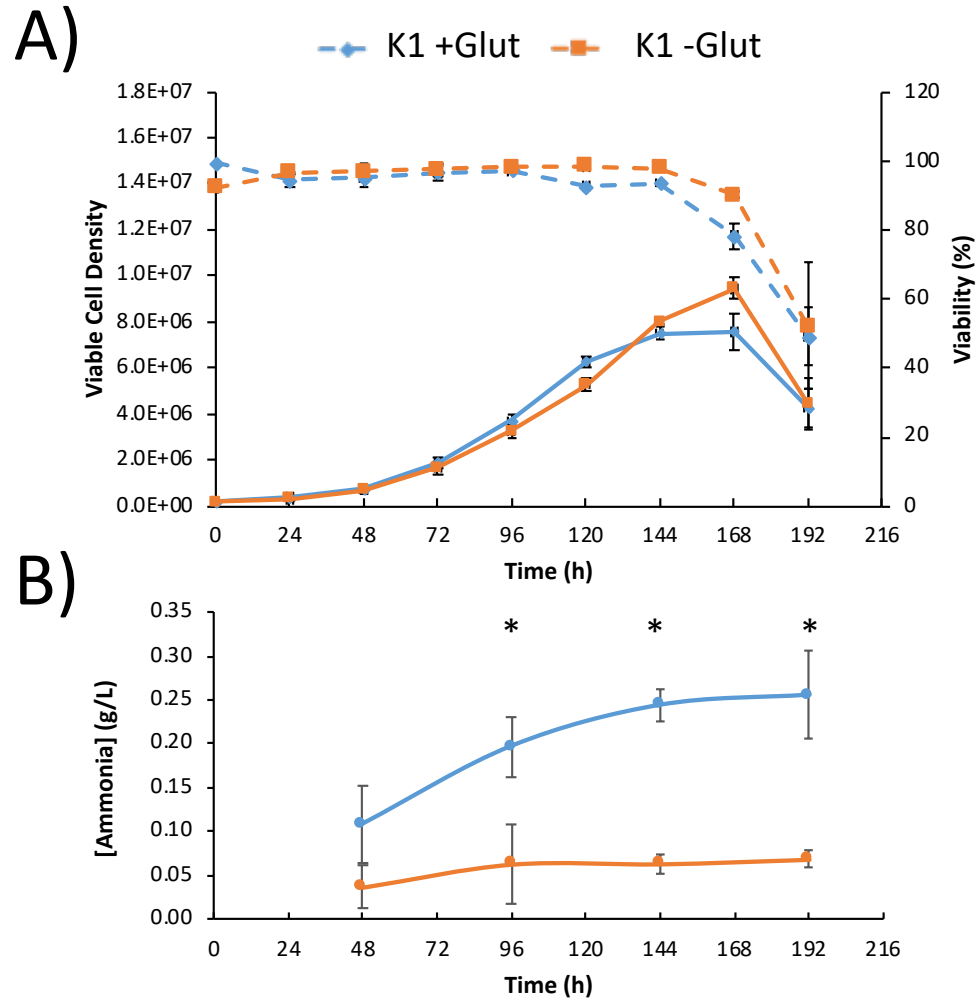


Figure 9 – Growth profile and ammonia concentration overtime for in-house CHO-K1 grown in BalanCD 8 mM L-Glutamine (K1 +Glut) and adapted CHO-K1 grown in BalanCD Glutamine-free (K1 –Glut) after 17 passages. A) Viable Cell Density (full lines and Viability (dashed lines) and B) Ammonia concentration measurements were taken every 24h. Each point corresponds the average of 3 biological replicates (N=3) and error bars represent standard deviations. Differences were evaluated using a two-tailed homoscedastic Student T-test. * = p-value < 0.05

CHO-K1 cells adapted to glutamine free conditions exhibited a very similar growth profile to the original glutamine dependent K1 cell line, with comparable growth rates and pick cell densities.

In addition, a lower ammonia production over time could be observed for the glutamine-free adapted K1 cell line, with ammonia concentration remaining constant around 0.05 g/L for the whole duration of the culture, in comparison with the CHO-K1 cell line growing in L-glutamine containing media, which showed increasing ammonia concentrations over the time of the culture ranging from 0.1 g/L after 48h to >0.25 g/L at the end of the culture, 192h) (Figure 9 – C).

2.1.4. qPCR validation of selected miRNAs correlating with ammonia concentration

In order to validate the selected targets TaqMan® microRNA assays were used. A preliminary study using different concentrations of the RNA on the RT step was carried out, aiming to determine the optimal concentration of template RNA to obtain a specific amplification signal for each target with a good the PCR efficiency (ideally a 10-fold difference should lead to a different of 3.3 on the Ct values). This was done considering that the microarray probes were not specific for CHO and that the available microarray data did not inform on the expression levels for the detected miRNAs. In short, reverse-transcriptase reactions were set with different amounts of RNA template (Table 3) and the correspondent RT primers per assay. Reversed transcribed samples were then analysed by qPCR.

Table 3 - Tested RNA concentrations for the optimization of miRNA TaqMan Assays. Samples were named based on the final amount of reverse transcribed RNA template that would be used in the qPCR step, from 2 to 200 ng.

Sample Name*	Template RNA amount for RT step (ng)
2	10
20	100
100	500
200	1000
NTC	0

***Note:** sample names are used in the amplification curves shown in Appendix A and correspond to the RNA concentration in the stock sample used for each RT reactions.

Samples were adjusted in order to obtain the indicated RNA amounts in the 5 µL of sample going into each RT reaction.

A no template control (NTC), comprising only TaqMan® master mix and qPCR assay primers was included in each plate to account for potential contamination of the reagents and test assay integrity. Amplification curves corresponding to each assay for the different concentrations of RNA template are shown in Appendix A.

U6-snRNA and hsa-miR-378 assays were used as positive control assays due to their previously demonstrated specificity in CHO cells (Costello *et al.*, 2018). For both miRNAs, specific amplification signals were observed with template amounts of 10 ng, 100 ng and 500 ng.

Some samples with high amount of template exhibited amplification signals at high Ct values. This was likely due to reverse-transcriptase inhibition and subsequent poor amplification at the qPCR stage. For assays hsa-miR-638, rno-miR-187*, hsa-miR-3163 and mmu-miR-1275 high background amplification in the H₂O and NTC controls was detected. The causes for these unspecific signals were investigated together with the technical team of the manufacturer and assays were deemed unusable. These assays and targets were discarded. Table 4 shows the Ct values obtained for the remaining assays at the different template concentrations.

Table 4 - Ct values determined using miRNA TaqMan assays for a small subset of the selected ammonia correlating miRNA targets. U6-snRNA and hsa-miR-378 were used as control assays

RNA amount in RT (ng)	U6-snRNA	hsa-miR-378	hsa-miR-409	hsa-miR-1275	mmu-miR-1195
10	21.4	25.5	N/A	33.1	38.1
100	17.9	22.1	30.8	30.6	33.7
500	15.6	20.5	28.7	31.6*	33.1*
1000	Undet.	31.9*	N/A	Undet.	Undet.
0	36.0	Undet.	Undet.	37.5	Undet.
Efficiency (%)	95.97	118.64		112.66	

*Marked values were flagged due reverse-transcriptase inhibition and have been excluded for efficiency calculations in Figure 10.

After plotting Ct values versus the logarithm (LOG10) of the concentration, the slope was used to calculate primer efficiency and to detect potential inhibition at high amount of template (500-1000 ng).

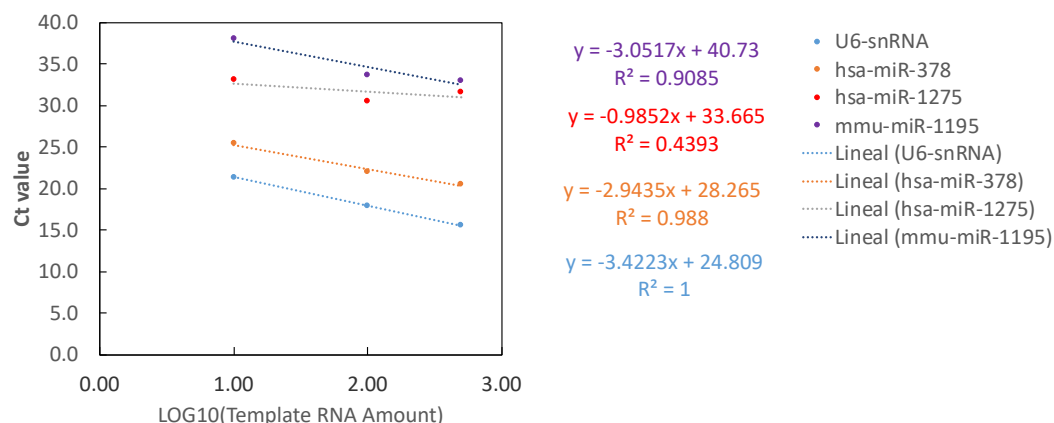


Figure 10 – Primer efficiency determination for miRNA assays U6-snRNA, hsa-miR-378, hsa-miR-1275 and mmu-miR-1195. A linear regression fit was used to calculate the primer efficiency for each assay (R^2). This could only be done with assays for which amplification at 3 or more concentrations was obtained.

This was observed for assays hsa-miR-1275 and mmu-miR-1195, with low R^2 coefficients and slopes very different to the expected 3.3 Ct difference with a 10-fold change in template amount (Bustin *et al.*, 2009). Based on that, PCR efficiencies could only be calculated for the U6-snRNA and hsa-miR-378 have been included in Table 4 for reference. The U6-snRNA assay showed an efficiency of 95.97%, whereas assay for hsa-miR-378 showed an amplification efficiency of 118.64%.

Based on the above results, validation was carried out for cgr-miR-1275 and cgr-miR-1195. Having verified the glutamine-free phenotype on the above described glutamine-free adapted cell line, cell pellets obtained in the time course experiment described above, from both cell lines CHO-K1 growing in 8 mM L-glutamine containing media and CHO-K1 growing in L-glutamine-free media every 48 hours were used to generate RNA samples for RT-qPCR validation using TaqMan® assays. Growth profile and viability correspond with those displayed in Figure 9.

hsa-miR-1275 was detected at Ct values between 28-30 and cgr-miR-1195 was detected at Ct values between 31 and 33. More detailed Ct data is included in Appendix A.

Figure 11 shows the relative quantification result for the two selected miRNAs. It can be observed that similarly to the microarray dataset, cgr-miR-1275 showed higher levels of expression in cells grown in glutamine containing media than in cells adapted to grow in glutamine-free conditions, with expression levels increasing over time. For cgr-miR-1195, expression profiles determined by microarray profiling and by RT-qPCR also matched. Cgr-miR-1195 expression decreased over time in glutamine-dependent CHO-K1 cells over time while expression was fairly constant in glutamine-free cells.

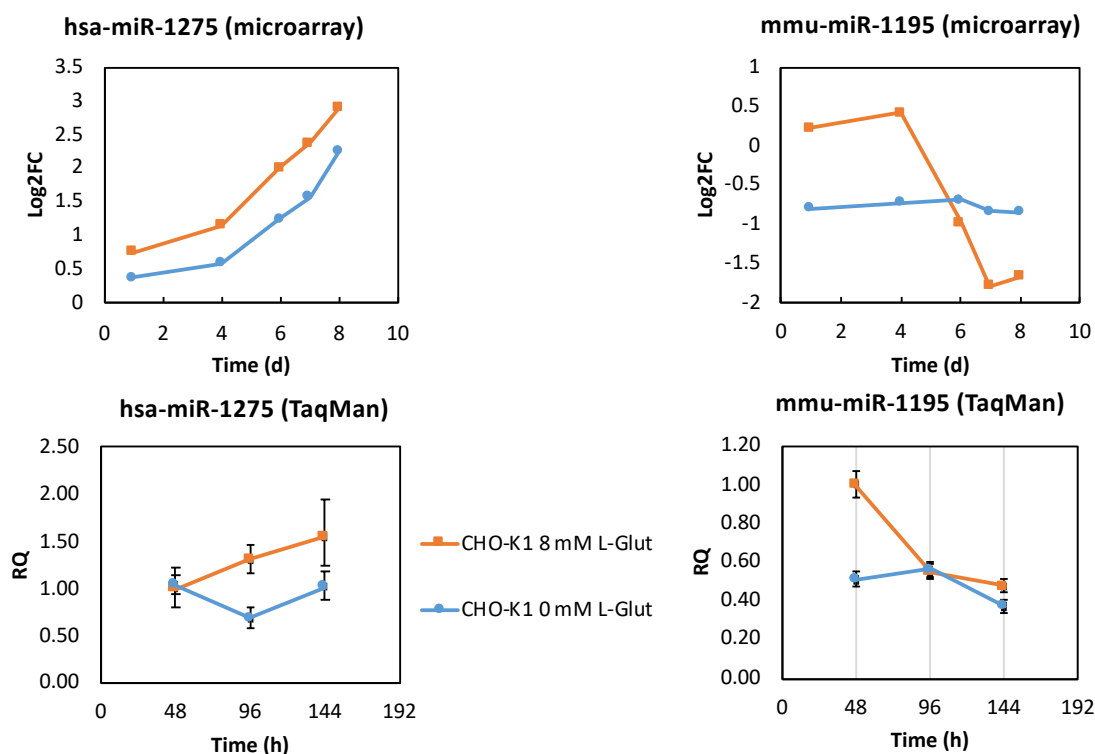


Figure 11 – RT-qPCR validation of expression profiles of cgr-miR-1275 and cgr-miR-1195 in glutamine dependent versus glutamine-free adapted CHO-K1 cells using TaqMan miRNA assays. RQ was calculated using U6-snRNA as endogenous reference and 8 mM L-glutamine as reference condition. Plotted Log2FC values (Top plots – microarray) correspond to expression values at each time point a pooled reference. Plotted RQ values (Bottom plots - TaqMan) correspond to average RQ values of 3 technical replicates (N=1) and error bars correspond to standard deviation.

None of these two miRNAs was moved forward due to their low expression levels (Ct > 30 eve with large amounts of template on the qPCR). The wet lab work for this project was stopped here. The next section describes some additional bioinformatics analysis performed improve the correlation analysis and explore the biological context of the miRNAs identified from it.

2.1.5. Incorporating annotation data to the initial selection of miRNA profiles correlated with Glutamine/Ammonia metabolism leads to new candidate miRNAs

The fact that the validation study carried out was not able to specifically detect some of the selected targets led to redesigning the target selection approach. It must be noted that the probes spotted on the microarrays corresponded to human (*hsa*), mouse (*mmu*) and rat (*rno*) miRNA sequences and those are not necessarily perfect matches to the CHO sequences. Therefore, each candidate miRNA of the obtained sub-list after filtering by Pearson's correlation coefficient was manually investigated by searching for:

- Annotation and sequence in miRBase for the human/mouse/rat probe
- Annotation in CHO in miRBase or in CHO published datasets for the homologous CHO miRNA (Table 5)
- Evidence of expression of the miRNA in published and unpublished CHO miRNA expression datasets (Table 5)

As previously described, expression data for all miRNA from all cell lines was compiled and Pearson's correlation coefficients (PCC) with ammonia concentration were calculated. A cut-off of $PCC > 0.6$ was applied to eliminate miRNAs showing low correlation to ammonia. A shorter list of miRNA candidates was obtained by selecting miRNA, which are annotated in CHO and whose expression has been shown in at least two published datasets (Table 5). It is interesting to note that the previously validated miRNA cgr-miR-1195 was found in this filtered list.

Further curation of the list was done checking for an approximation of the abundance of the selected miRNAs by looking at the observed reads for each candidate on the different analysed datasets. Only miRNAs with reported abundances above 100 counts in at least two datasets were considered. The average number of reads per identified miRNA and the maximum number of reads found for the most abundant target in the dataset are also provided as reference (Table 6).

The only miRNAs fulfilling all the above mentioned criteria were cgr-miR-222-3p and cgr-miR-93-5p. Figure 12 shows correlation plots between miRNA expression and ammonia concentration for these two miRNAs.

Table 5 – Final selected miRNAs candidates using a cut-off of 0.6 for their Pearson's Correlation Coefficient with ammonia concentration and after performing a manual search for their annotation in different miRBase.org and different available public and in-house miRNA expression datasets

Probe ID/Array ID	miRBase	Sequence miRbase	Published CHO miRNA datasets							Unpublished	CHO annotation	Pearson
			Hackl et al. 2011	Hackl et al. 2012 (hairpins)	Johnson et al. 2011	Hammond et al. 2012	Maccanti et al. 2014	Clarke et al. 2012 (hairpins)	In-house Fedbatch RNASeq/Temp Shift Dataset			
rno-miR-187*	YES	ggcuacaacaacaggaacccgggc	YES	YES	NO	YES	YES	YES	NO	cgr-miR-187-5p		0.95
hsa-miR-187*/mmu-miR-187*	YES	aggcuacacacaggaacccggg	YES	YES	NO	YES	YES	YES	NO	cgr-miR-187-5p		0.93
rno-miR-344b-1-3p	YES	gaauuaaccaaagcccgacugu	YES	YES	NO	YES	NO	YES	YES	cgr-miR-344-3p		0.90
hsa-miR-203/mmu-miR-203/rno-miR-203	NO	guagaauquuuaggaacacuaag	YES	NO	NO	YES	NO	NO	YES	cgr-miR-203-3p		-0.85
hsa-miR-138-1*/mmu-miR-138-1*/rno-miR-138-1*	YES	agcuguguuuguagaucagcgcg	YES	YES	YES	YES	NO	YES	NO	cgr-miR-138-5p		-0.85
mmu-miR-138-2*/rno-miR-138-2*	YES	agcuguguuuguagaucagcgcg	YES	YES	YES	YES	NO	YES	NO	cgr-miR-138-5p		-0.84
hsa-miR-222/mmu-miR-222/rno-miR-222	YES	agcuacaucucugcuacucgggu	YES	YES	YES	YES	NO	YES	YES	cgr-miR-222-3p		-0.82
hsa-miR-183/mmu-miR-183/rno-miR-183	YES	uauggcacucuguaaauucaau	YES	YES	NO	NO	NO	YES	YES	cgr-miR-183-5p		-0.80
hsa-miR-93/mmu-miR-93/rno-miR-93	YES	caaaugucuguuucguagcagguag	YES	YES	YES	YES	NO	YES	YES	cgr-miR-93-5p		-0.73
mmu-miR-1195	NO	ugaguuucgaagccagccugcuca	YES	NO	NO	NO	NO	NO	YES	cgr-miR-1195-5p		-0.69

Table 6 – Abundance data (counts from different RNASeq published experiments) for the shortlisted miRNA. For each target, expression data was searched in different available published and in-house datasets. Average reads per miRNA and max reads obtained for the highest expressed miRNA on the dataset are shown for reference.

Probe ID/Array ID	CHO annotation (Hackl.2011)	Hackl <i>et al.</i> 2011	Johnson <i>et al.</i> 2011	Hammond <i>et al.</i> 2012	UCB/eCHO dataset
rno-miR-187*	cgr-miR-187-5p	17	n/a	3	15
hsa-miR-187*/mmu-miR-187*	cgr-miR-187-5p	17	n/a	3	15
rno-miR-344b-1-3p	cgr-miR-344-3p	134	n/a	1	1450
mmu-miR-1195	cgr-miR-1195-5p	139	n/a		
hsa-miR-93/mmuc-miR-93/rno-miR-93	cgr-miR-93-5p	101273	6827	4328	64880
mmu-miR-489	n/a				
hsa-miR-183/mmuc-miR-183/rno-miR-183	cgr-miR-183-5p	35	n/a		55
hsa-miR-222/mmuc-miR-222/rno-miR-222	cgr-miR-222-3p	11445	920	826	2231
mmu-miR-138-2*/rno-miR-138-2*	cgr-miR-138-5p	34	4	12	4
hsa-miR-138-1*/mmu-miR-138-1*/rno-miR-138-1*	cgr-miR-138-5p	34	4	12	4
hsa-miR-203/mmuc-miR-203/rno-miR-203	cgr-miR-203-3p	3263	n/a	20	

REFERENCES				
AVERAGE READS	46214	4900	5024	27621
MAX.	1341517	858424	140587	888908

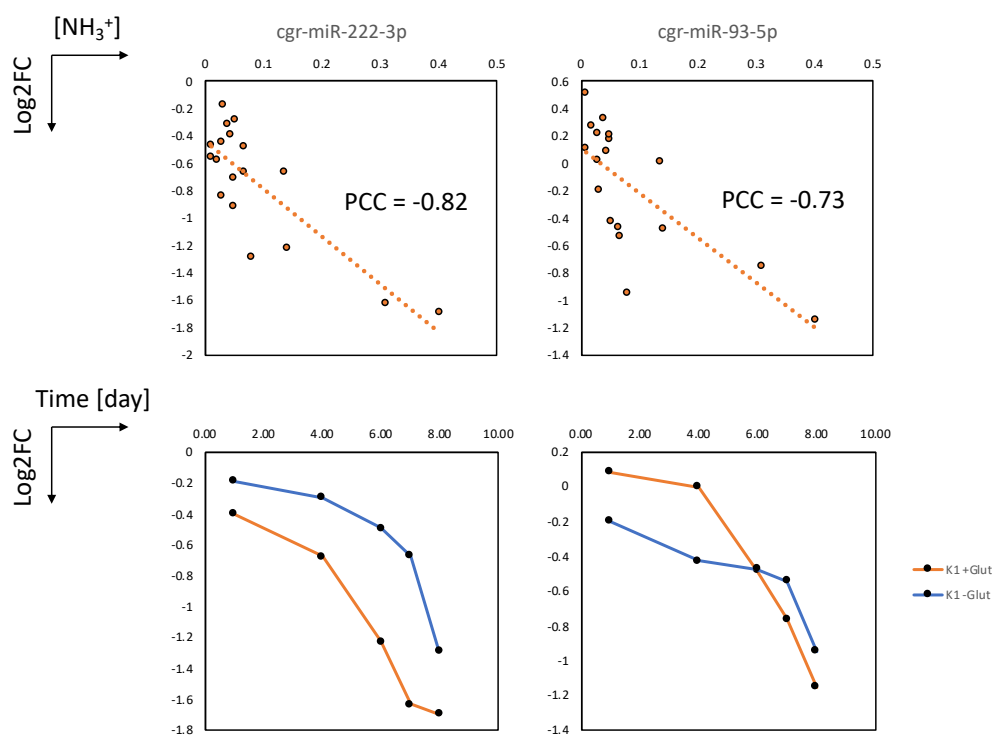


Figure 12 – Correlation plots between miRNA expression and ammonia concentration for the selected miRNA targets (top) and expression profiles over time (bottom). Targets were selected after applying the Pearson's Correlation Analysis and annotation/expression filters.

These two miRNAs exhibited very similar expression profiles that seemed independent of the presence or absence of glutamine, suggesting their expression is independent from glutamine metabolism. In order to gain a broader understanding of the roles that these two miRNAs have in terms of regulating cellular functions, target prediction and pathway analysis were performed.

2.1.6. Pathway Analysis of selected miRNAs – assessing potential off target effects

In order to further investigate if there is a common regulatory function of the two selected miRNAs, target prediction and pathway enrichment analysis were performed as described in Methods Section 3.5.

Table 7 – Enriched pathways identified based on the predicted list of targets for the mouse orthologs of two selected miRNAs: *cgr-miR-93* and *cgr-miR-222*. Target prediction and pathway enrichment analysis were performed using the miRPath3.0 (DIANA Tools) software

#	KEGG pathway	p-value	#genes	#miRNAs
1	Hippo signaling pathway (mmu04390)	1.29E-05	19	2
	mmu-miR-222-3p Tarbase		8	
	mmu-miR-93-5p Tarbase		13	
2	Lysine degradation (mmu00310)	9.08E-05	11	2
	mmu-miR-222-3p Tarbase		5	
	mmu-miR-93-5p Tarbase		6	
3	Hepatitis B (mmu05161)	1.89E-03	21	2
	mmu-miR-93-5p Tarbase		20	
	mmu-miR-222-3p Tarbase		4	
4	Thyroid hormone signaling pathway (mmu04919)	2.64E-03	23	2
	mmu-miR-222-3p Tarbase		10	
	mmu-miR-93-5p Tarbase		15	
5	MAPK signaling pathway (mmu04010)	1.09E-02	38	2
	mmu-miR-93-5p Tarbase		27	
	mmu-miR-222-3p Tarbase		17	
6	FoxO signaling pathway (mmu04068)	1.51E-02	24	2
	mmu-miR-93-5p Tarbase		21	
	mmu-miR-222-3p Tarbase		5	

(*) Continues next page

Table 7 (continuation)

7	<u>Osteoclast differentiation</u> (mmu04380)	1.78E-02	23	2
	mmu-miR-93-5p Tarbase		17	
	mmu-miR-222-3p Tarbase		8	
8	<u>Proteoglycans in cancer</u> (mmu05205)	1.78E-02	27	2
	mmu-miR-93-5p Tarbase		15	
	mmu-miR-222-3p Tarbase		15	
9	<u>Renal cell carcinoma</u> (mmu05211)	1.78E-02	14	2
	mmu-miR-93-5p Tarbase		9	
	mmu-miR-222-3p Tarbase		7	
10	<u>Fatty acid degradation</u> (mmu00071)	2.06E-02	5	2
	mmu-miR-93-5p Tarbase		4	
	mmu-miR-222-3p Tarbase		1	
11	<u>Valine, leucine and isoleucine degradation</u> (mmu00280)	2.06E-02	8	2
	mmu-miR-222-3p Tarbase		4	
	mmu-miR-93-5p Tarbase		5	
12	<u>Prolactin signaling pathway</u> (mmu04917)	2.06E-02	18	2
	mmu-miR-93-5p Tarbase		13	
	mmu-miR-222-3p Tarbase		8	
13	<u>Estrogen signaling pathway</u> (mmu04915)	3.16E-02	13	2
	mmu-miR-93-5p Tarbase		9	
	mmu-miR-222-3p Tarbase		7	
14	<u>Prostate cancer</u> (mmu05215)	3.16E-02	18	2
	mmu-miR-93-5p Tarbase		17	
	mmu-miR-222-3p Tarbase		3	

In short, the miRPATH3.0 from the online suite Diana Tools was used to predict targets for those two miRNAs. Following that, the obtained lists of targets were used as input for a pathway enrichment analysis resulting in a list of pathways in which cgr-miR-93 and cgr-miR-222 might play a regulatory function (Table 7). Interestingly, a number of pathways related with signal transduction (KEGG ID 09132) were identified: the Hippo signalling pathway (mmu04390), the MAPK signalling pathway (mmu04010) (also included in the Thyroid hormone signalling pathway (mmu04919), Proteoglycans in cancer (mmu05205), Prolactin signalling pathway (mmu04917), Estrogen signalling

pathway (mmu04915)), and the FoxO signalling pathway (mmu04068). In addition, catabolic pathways related to amino acid and lipid catabolism (Lysine degradation (mmu00310) and Valine, leucine and isoleucine degradation (mmu00280) and Fatty acid degradation (mmu00071) were identified.

In summary, both cgr-miR-93 and cgr-miR-222 have predicted targets that belong to pathways related with important catabolic processes of lipids and amino acids as well as pathways related to phosphorylation signalling cascades. Together with the fact that these two miRNAs showed decreased expression over time, it is possible to hypothesize that both belong to one or several regulatory networks responsible to activate these pathways during the late stages of the culture, when cells are more and more under nutrient deprivation conditions. While the two identified miRNAs would not be suitable switches under the scope of this thesis, they should be considered as interesting engineering targets.

2.2. RNASeq microRNA profiling reveals expression patterns correlated with cell culture stage

As part of the effort to identify miRNAs expression profiles that correlate with culture conditions or cell metabolic states, a collaborative effort in the eCHO Systems network led to the generation of a multi-omics dataset of a CHO-DG44 derived producer cell line expressing a monoclonal antibody grown in fed-batch under industrially relevant conditions (10 L scale). Here, the preliminary analysis of the miRNA expression dataset after quality control and read count after alignment to the CHO genome is presented.

2.2.1. Differential expression analysis and unsupervised clustering reveal three distinct miRNA expression profiles over the fed batch culture.

In order to identify miRNAs with expression patterns changing over time, a Likelihood Ratio Test was performed on the data using the DESeq2 software package within the R statistical framework. Principal component analysis (PCA) was used as a quality control step (Figure 13). In short, PCA is a multivariate analysis that aims to explain the variance of the data by transforming it to a new coordinate system (each coordinate called a principal component) usually after a normalisation step. Practically, by visualising the first few principal components, a lower-dimensional picture of the data can be obtained. This simpler picture often allows to assess the quality of the data (i.e. replicate samples

should cluster together) and in some cases can help in the interpretation of the data if a particular variable can be matched to one of the principal components. The first observation from this analysis was that samples aligned on the first component axis (PC1) chronologically, from day 0 to day 12, indicating that variation explained by this first component can be explained by changes related to culture time. In addition, duplicate samples clustered close to each other for all time points except for samples on day 10 (10d). Samples corresponding to lag phase and early exponential phase (0d, 3d, 4d) and death phase (10d, 11d, 12d) were grouped in two distinct clusters while samples corresponding to growth phase and early stationary phase time were separated from each other. This indicates that changes in miRNA expression are more temporally localised during these two culture stages and are more constant during early and late phases of the culture.

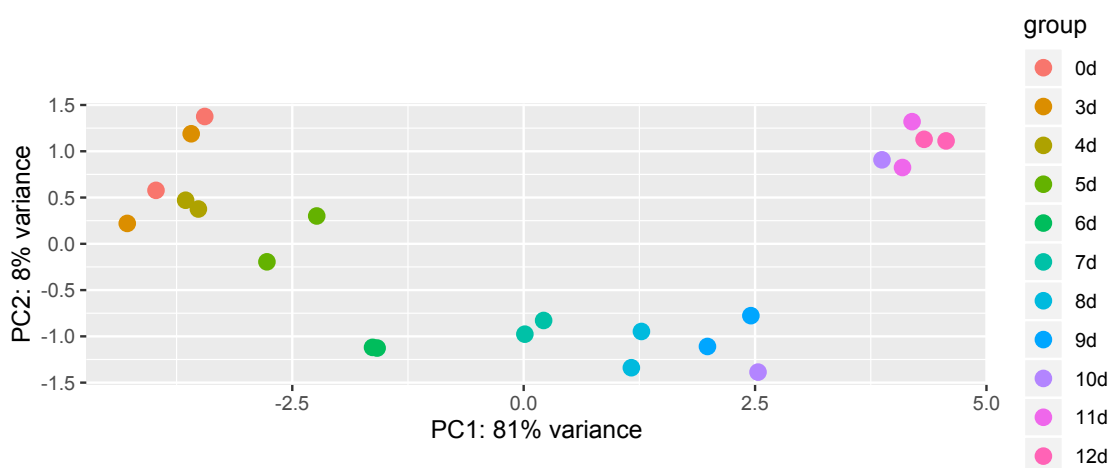


Figure 13 – Principal Component Analysis clustering of miRNA expression per sample (group). PCA analysis was performed as a quality control step during data analysis to assess clustering of duplicate samples. Each colour corresponds to a sampling time point and duplicate samples are shown (N=2).

Following that, the differential expression analysis revealed 141 miRNAs differentially expressed over time after applying an average read cut off > 100 counts and a p-value > 0.05. A filter was also applied to only select miRNAs currently annotated in the latest version of miRBase. A comprehensive list of the DE targets including BaseMean counts, Log2 fold change and statistical analysis is included in Appendix A

Having identified miRNAs whose expression changes over time, the next step aimed to identify common expression patterns among those and map the to the different growth

phases. Using the DEReports R package, unsupervised hierarchical clustering was used to group the different DE miRNAs in clusters according to their expression profile. Three main clusters, containing 19, 47 and 27 miRNAs respectively, were identified. A comprehensive list of miRNAs for each cluster is included in Appendix A - Supplementary Table 1. Expression profiles for each cluster are depicted in Figure 14 - A. Figure 14 - B shows the growth and viability data corresponding to the profiled experiment. It can be observed how each identified cluster shows a distinctive expression profile in relation to the different growth phases.

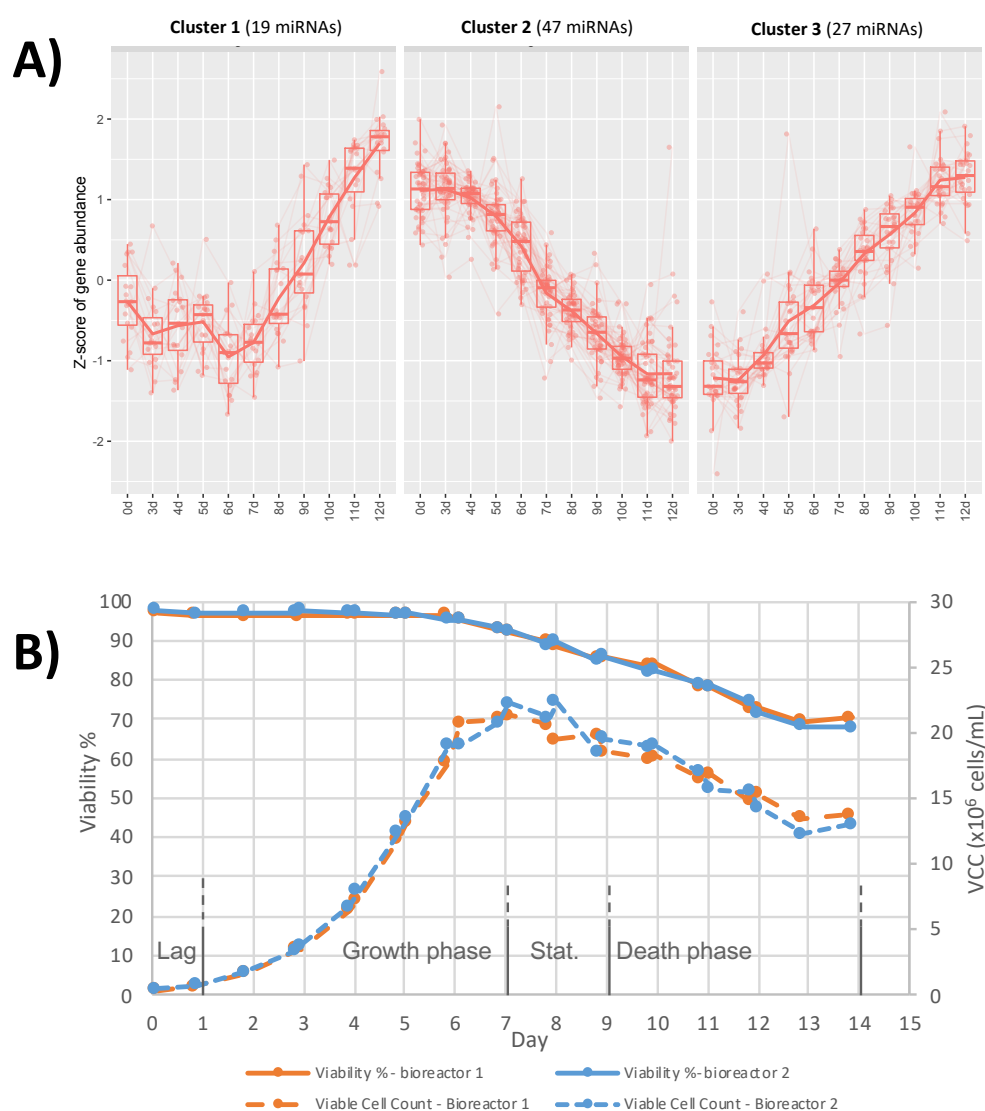


Figure 14 – A) Graphical representation of the expression patterns observed for the main three expression clusters identified by unsupervised hierarchical clustering on the DE miRNA subset. B) Viable cell concentration (VCC) and Cell Viability (%) for the two bioreactor runs performed. The different growths phases are indicated next to the X axis.

Cluster 1 comprises miRNAs which show stable expression during the lag and exponential growth phases, followed by an increase in expression during the stationary phase and death phase. Cluster 2 comprises miRNAs whose expression decreases over time independently of the growth phase. Finally, cluster 3 includes miRNAs that show increased in expression over the duration of the culture.

2.2.2. Pathway Enrichment Analysis indicated redundant regulation of fatty acid metabolism by the different miRNA clusters

As before, it was considered important to investigate the biological context related to the different expression profile clusters and the miRNAs included in each of them. For that, functional analysis of the different clusters was performed as outlined in Methods using the DIANA Tools suite / miRPath3.0 algorithm. In this case, each individual cluster was analysed separately by using the list of miRNAs belonging to it as input (miRNAs belonging to each cluster can be found in Appendix A). Due to the large number of predicted targets for each miRNA, the Pathway Union a posteriori method was used. IN this mode, the server identifies pathways enriched predicted targets of at least one miRNA. Figure 15 shows a breakdown of the identified pathways for each cluster and the specific miRNAs which have predicted targets included in each pathway.

The most interesting result of this analysis was the presence of the lipid biosynthesis and lipid degradation pathways among the enriched pathways identified. More in detail, miRNAs whose expression increased right after the end of the exponential growth phase (Cluster 1) such as miR-30a-5p, or overall during the whole culture period (Cluster 3) such as miR-27b-3p, miR-34c-5p, miR-338-3p and miR-320-3p showed predicted targets belonging to the fatty acid biosynthesis pathway. Some miRNA included in Cluster 2, whose expression decreases over time, such as miR-18a-5p, miR-17-5p, miR-7a-5p, miR-30e-5p or miR-107-3p also had predicted targets belonging to the same pathways.

It is worth noting that similarly as for the previously analysed glutamine miRNA profiling dataset, a second group of enriched pathways containing pathways related to signal transduction by phosphorylation. This group included the Hippo signalling pathway, the FoxO signalling pathway, as well as some of the pathways related to the MAPK signalling pathway such as the Thyroid hormone signalling pathway or the Proteoglycans in cancer pathway.

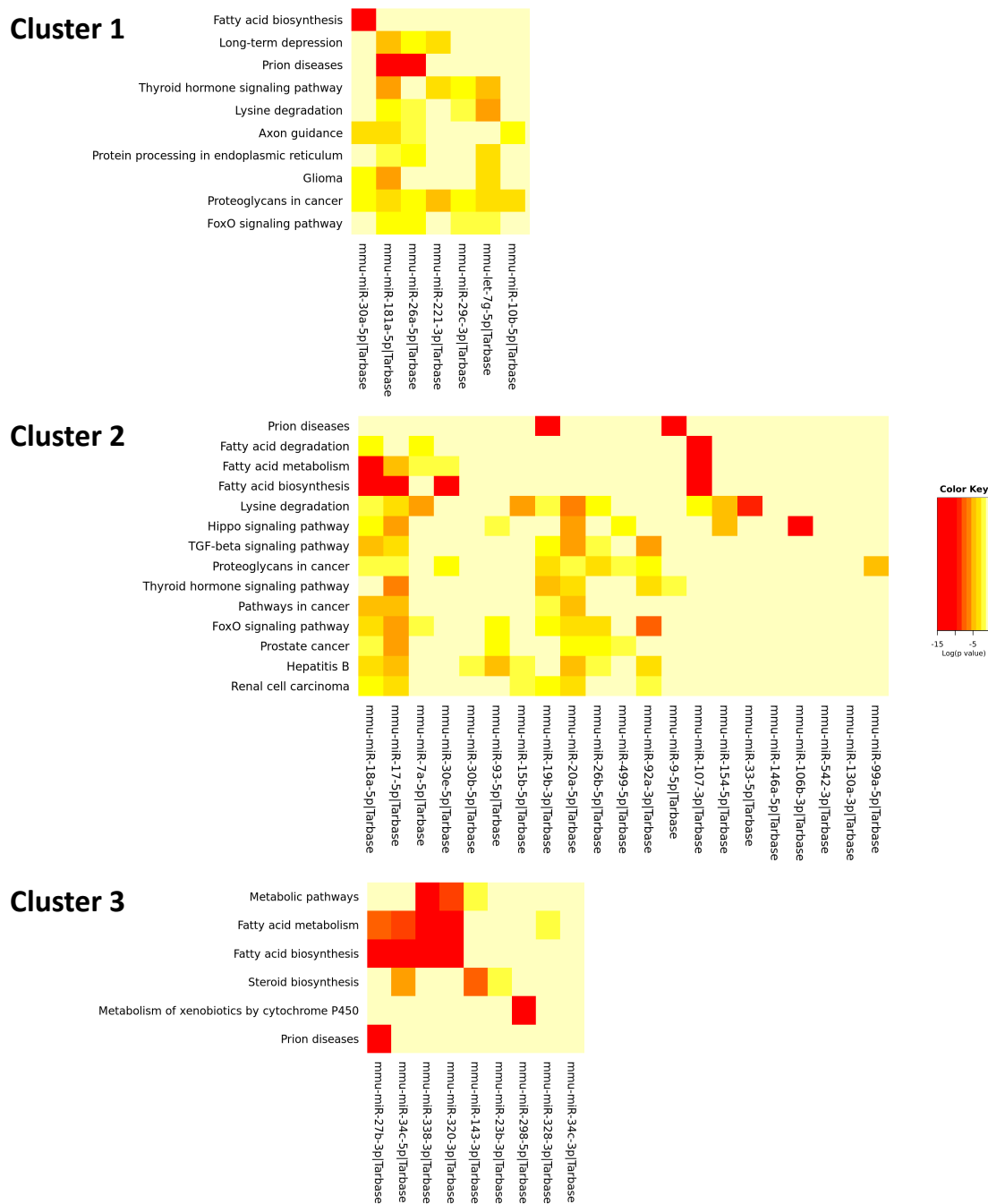


Figure 15 – Heat maps showing enriched pathways after performing enrichment analysis on each miRNA cluster. Heat maps were generated using the Diana Tools Suite miRPath3.0 / Pathway union algorithm with a cut off at FDR < 0.05 degrees and conservative stats for the analysis.

Due to the late arrival of the dataset described above, this project was stopped here. As discussed below, the most interesting cluster from the 3 identified was Cluster 1. miRNAs included in this cluster showed and increased expression coinciding with the end of the exponential phase and the beginning of the stationary phase. From the perspective of building miRNA-based switches, these candidates would be the most

interesting as they could potentially be used to modulate transgene expression at this crucial time point during the cell culture.

3. Discussion

3.1. Adapting cells to glutamine free conditions – one step is enough

Since the early days in mammalian cell culture directed to biotechnological production, the role of glutamine as a major source of carbon and nitrogen for the cell has been acknowledged and studied. Several studies have shown that despite it being a source of carbon and nitrogen for the cell, glutamine as such is not essential for cell growth or recombinant protein production. There have been several published studies reporting the benefits of reducing or substituting glutamine with other amino acids or TCA intermediary metabolites as a way to reducing unwanted side effects of glutamine metabolism such as ammonia build up, while maintaining and even improving growth and productive characteristics of the tested cell lines (Altamirano *et al.*, 2000; Genzel *et al.*, 2005; Hong, Cho and Yoon, 2010; Rajendra *et al.*, 2012; Ha and Lee, 2014). In addition, other studies have adapted cell lines to grow under glutamine-free conditions. Besides, the CHO-K1 cell line adapted by Bort *et al.* mentioned above, another CHO-DUKXB11 producer cell line has been adapted to glutamine-free growth by FACS-directed evolution (Bort, Stern and Borth, 2010; Taschwer *et al.*, 2012) demonstrating the feasibility of recombinant protein production in such conditions. An interesting part of this study was the fact that it was possible to generate glutamine-free CHO-K1 cells using a simpler approach than previous studies. The original publication from Bort *et al.* on adapting CHO-K1 cells to glutamine free conditions used an adaptation strategy involving several rounds of glutamine concentration reduction, seeding cells at high density and FACS sorting the surviving cells once viability was close to 10% in order to obtain an adapted population able to grow in glutamine free media (Bort, Stern and Borth, 2010). The overall process is described as taking no more than 15 weeks and the resulting cells are shown to perform similarly, if not better, than the original glutamine dependent CHO-K1 cell line, exhibiting similar cumulative viable cell densities, with higher peak cell densities, although at a slightly slower growth rate. In this study, cells were adapted using a one single glutamine reduction step followed by regular sub-culturing every 3 to 5 days. The whole process lasted less than 10 weeks and the obtained adapted population performed similarly to the original CHO-K1 growing in

glutamine containing media. Surprisingly, while Bort *et al.* reported a strong reduction in cell viability after the sudden reduction of glutamine concentration in the media, the results presented in this study show that cell viability was not compromised. A recent study by Maralingannavar *et al.* used a similar approach to study CHO producing cells under glutamine deprivation conditions (Maralingannavar *et al.*, 2019). The authors there described a selection of cells under glutamine free conditions that consisted of seeding CHO-DG44 at high cell density (1×10^6 cells/mL) and passage them every two days. Compared with control cultures grown in 5 mM L-glutamine containing media, cells grew to lower cell densities but survived up to 28 passages. Although the authors do not report cell viabilities and there are differences in terms of the cell lines and seeding densities used, this study is consistent with the results reported in this work that a sudden reduction in glutamine concentration does not necessarily lead to cell death.

In this work ammonia concentration in the media was also measured during a batch culture comparing the original CHO-K1 cells grown in 8 mM L-glutamine with the glutamine-free adapted CHO-K1 cells in shake flasks. Ammonia concentration remained very low during the whole duration of the culture, staying at similar levels to the ones previously observed for the glutamine free cells used in the microarray profiling experiment, which were grown in stirred tank reactors under tightly controlled oxygen and pH conditions. It is interesting to note that Bort *et al.* report in their study that their cells adapted to glutamine-free conditions using the FACS strategy reached similar ammonia concentrations to the original cells grown in glutamine containing conditions suggesting that the adaptation strategy has an impact on the metabolic characteristics of the resulting cells, which might be different depending on the strategy. It is reasonable to think that a selection strategy involving the sorting of cells surviving in high cell density cultures, despite a stepwise reduction in glutamine concentration at each step, does not necessarily imply that the selected cells are going to be the ones that better adapted to glutamine-free conditions. Nutrient depletion and accumulation of toxic by-products, together with flow-cytometry cell sorting, are two examples of factors that are likely to generate high levels of stress for cells in addition to glutamine removal and therefore, assuming that the selected characteristics will be only linked with glutamine removal might be incorrect. An interesting study would be to compare

glutamine-free adapted cell lines by using the two different strategies in order to be able to determine the metabolic changes underlying the observed adaption in each case.

3.2. Glutamine metabolism and culture time as a guiding differential factors for miRNA profiling

In this study, the aim was to use available CHO miRNA expression data to identify miRNA profiles related to specific cell metabolic requirements or cell culture stage. The first studied dataset presented the opportunity to explore a model in which glutamine/ammonia could be used in order to switch ON/OFF the expression of a particular miRNA. Due to the nature of the compared cell lines, the most differential trait was the fact that some of the profiled cells lines had been adapted to glutamine-free media. Ammonia being the direct by-product of glutamine metabolism, it was not a surprise to find it as the variable with the highest number of correlating miRNAs. In terms of applicability in an industrial setting, the design and engineering of a CHO host in which miRNA expression could be controlled by addition of glutamine to the media or synchronized with the depletion of glutamine after the growth phase, would be desirable. The second dataset investigated represented a model system for a typical recombinant protein production process and provided the opportunity to investigate miRNA expression changes occurring over time. This scenario would also be suited for the development of gene expression control systems that would allow to coordinate expression of a particular transgene, e.g. product gene, to a particular phase of the culture, e.g. stationary phase, thus reducing metabolic burden during growth phase and limiting the amount of time that recombinant product remains in the reactor before harvest.

It must be acknowledged that in both cases, the experiments which led to the generation of that data were not specifically designed towards extracting robust relationships between miRNA expression and cell culture parameters but more with a more traditional focus to unveil the expression landscape during a typical bioprocess run. Due to the dynamic nature of CHO cell cultures and the always changing environment in the bioreactor, where cells are subject to multiple stresses related to nutrient deprivation and by-product accumulation, looking for specific relationships between targets and particular metabolic processes is challenging. Time course miRNA expression data most probably reflects the mechanisms that the cells use to maintain

homeostasis, while growing and producing recombinant protein. The complexity of the biological processes behind that are only beginning to be unveiled in the last years and therefore, an attempt to find specific links between miRNA expression and a particular metabolic event or cellular state in that dynamic context may be like looking for a needle in the haystack on a windy day. Of course, one can argue that complexity does not matter as soon as robust and significant changes in miRNA expression can be identified if the purpose is to use them as miRNA switches. However, better experimental setting will have to be designed and more importantly robust validation post-profiling will have to be performed in order to find these robust switches.

3.3. Hsa-miR-1275 and cgr-miR-1195 as ammonia sensitive miRNAs

Despite the above-mentioned difficulties, two miRNA candidates were identified from the glutamine / glutamine-free dataset and were validated. RT-qPCR results showed similar expression patterns as the ones observed in the original microarray profiling experiment. Considering culture conditions were drastically different in terms of reactor (controlled fermenter versus shaken 50mL spin tube), cells and media, these results represent a good starting point for further investigation of these two candidates. It must be noted that both miRNAs were detected at low levels in the RT-qPCR experiments and no strong evidence of robust expression for any of them could be found in other published RNASeq datasets. Hsa-miR-1275 has no annotation for CHO in miRBase and has only been reported by Maccani *et al.* where the authors showed differential expression of this miRNA between two producer cell lines, although no validation was pursued (Maccani *et al.*, 2014). Cgr-miR-1195 is the CHO annotation for mmu-miR-1195 and has been identified by Hackl *et al.* in 2011, without a sequence being reported, and has also been found in two in house CHO-specific datasets (Hackl *et al.*, 2011). In future work, it will be important to perform further work to establish if the correlations observed are reflecting a specific involvement of these miRNAs in the regulation of the glutamine metabolism, or if the observed changes in miRNA expression are a consequence the metabolic changes but play no role in regulating them. Pinpointing the actual role that miRNA changes play in the response of the cell to a changing environment, and establishing strong cause/consequence relationship between miRNA expression and changes in culture conditions will be critical information to have in order to build these miRNAs into a downstream control system and avoid the disruption of

essential metabolic functions by introducing miRNA targets that could compete with endogenous ones.

In the context of biopharmaceutical production, for example, this would open the door for feeding strategies to play a role in the control of transgene expression. In particular, during a fed-batch culture, it would be interesting to keep the product transgene in an OFF state during the growth phase and coordinate its expression with the depletion of certain metabolites like glucose or glutamine, known to promote growth and which typically are kept at a low concentration during the stationary production phase. If found, a miRNA or subset of miRNAs whose expression would be progressively reduced with glucose or glutamine depletion, could be the effector molecules controlling the switch as it would allow for a conditional silencing of the targeted transgene during high glucose and high glutamine growth phases, and switch it on with its levels of expression decreasing.

3.4. miRNAs correlated with ammonia and/or culture time are linked to amino acid and lipid catabolism as well as to membrane-nucleus phosphorylation signalling pathways.

Due to the complex nature of the regulatory events attributed by miRNA in mammalian cells (Bartel, 2018), the identification of useful miRNA targets to be integrated in artificially created regulatory networks needs to go hand in hand with research on the actual function that these miRNAs play in cell metabolism in order to anticipate and if possible avoid any potential interference with essential cell metabolic pathways when introducing artificial gene networks. In this work, pathway enrichment analysis was used as a way to unveil the biological context in which selected miRNA targets lie and as a way to predict, or at least inform the effect that manipulation of those could have.

Despite differences in the experimental setting, cell lines, data collection and analysis performed, it is interesting to see that using two independent datasets and two different methodologies (correlation analysis versus DE analysis followed by clustering), both pathway enrichment analysis undertaken in this study led to the identification of a very similar subset of biological pathways. In fact, it is possible to draw strong parallels between the two analysed experiments. On the one hand, the glutamine/glutamine free dataset consists of a batch culture and using ammonia concentration. As in the presented process ammonia concentration increases over time, correlating miRNA

expression with ammonia concentration implies correlating it with time, as am. On the other hand, in the fed-batch, where time was the variable used for the analysis, one cannot forget that, as in any model cell culture, ammonia accumulates over time and cells transition from a high-glutamine scenario at the beginning of the culture, to a glutamine-depleted scenario towards the end of the culture. Taking these two things into account, it is not surprising that the identified pathways underlying the observed changes in miRNA expression are not that different.

While the investigation of the particular metabolic pathways identified was always out of the scope of this thesis, it is worth to broadly discuss them in the context of the two analysed experiments as well as to outline the implication of potentially affecting their regulation with a hypothetical cell engineering intervention targeting any of the identified miRNAs. The identified pathways from both datasets can be grouped in three according the biological processes they are involved in. Pathways related to diseases (e.g. Hepatitis B) or specific to a particular cell type/tissue (e.g. Axon guidance) will be left outside of this discussion as they usually consist of a compendium of pathways related to a particular condition/cell type and are not relevant in the context of CHO.

As outlined in the results section, several pathways involved in Signal transduction (KEGG ID 09132) were identified as differentially regulated over time in both datasets. Just to recapitulate, these pathways are the Hippo signalling pathway (mmu04390), FoxO signalling pathway (mmu04068), the MAPK signalling pathway (mmu04010) and TGF-beta signalling pathway (mmu04350). The common feature between all these pathways is that they all involve phosphorylation signalling cascades from the cell membrane to the nucleus and have been widely studied in cancer and are linked with the regulation of cell proliferation, survival and stress response as well as cytoskeleton remodelling to promote adhesion and invasion (Herzig and Shaw, 2018). In the CHO field, signal transduction by protein phosphorylation has been studied since early days, and recently more and more studies are investigating these pathways (Dahodwala *et al.*, 2019). Maybe the most studied pathway involving phosphorylation signalling in CHO cells has been the mammalian target of rapamycin or mTOR, which has been shown to act as a signalling hub and coordinates cellular responses to nutrients in the environment as well as to regulate cell growth (Kim and Guan, 2019). Different studies have manipulated this central pathway in order improve cell phenotypes. For example,

overexpression of mTOR was shown to increase cell proliferation and productivity in CHO cells (Dreesen and Fussenegger, 2011). Another study looking at the effects of treating CHO cultures with rapamycin, a known mTOR inhibitor, showed that blocking mTOR increases viability towards the end of the culture by promoting autophagy (Lee and Lee, 2012). Transcriptomic analysis of high and low producing CHO cell lines has revealed up-regulation of several mTOR signalling pathways genes in cells showing higher productivity (Edros, McDonnell and Al-Rubeai, 2014). In the context of amino acid metabolism, a study in 2013 by Fomina-Yadlin *et al.* on the effects of individual amino acid depletion on CHO cell cultures, highlights that glutamine-deprivation does not lead to comprehensive downregulation of mTOR effectors (Fomina-Yadlin *et al.*, 2014). It is worth noting that in a follow-up proteomic and phospho-proteomic study carried out *in house* on the same model system, CHO-K1 glutamine-dependent versus CHO-K1 glutamine-free adapted, revealed differential expression of proteins belonging to several pathways linked to ECM-receptor interactions, actin remodelling and membrane trafficking as well as differential phosphorylation of several critical proteins belonging to the above discussed pathways, providing further evidence of the relevance of these pathways as drivers for the observed glutamine-free adaptation (Kaushik *et al.* 2019 - manuscript submitted). Another recent study has linked increased productivity in a monoclonal antibody producing CHO cell line with differential phosphorylation of CREB1, a transcription factor downstream of the cAMP and MAPK signalling pathways (Dahodwala *et al.*, 2019). The results obtained in this study using miRNA expression data show evidence that alternative phosphorylation signalling pathways such as MAPK should be further investigated for cell engineering in the future. Interestingly, very little literature is available about MAPK signalling in CHO cells, except a few references where CHO cells have been used to express druggable targets related to that pathway. However, as discussed below, several reports have linked nutrient deprivation to ER-stress response and autophagy as a cell survival strategy, and the MAPK pathway has been shown to connect these two cellular processes (Rashid *et al.*, 2015).

The second of group of pathways identified from both datasets comprises catabolic reactions of amino acids and lipids and included Lysine degradation (mmu00310), Fatty acid degradation (mmu00071), Valine, leucine and isoleucine degradation (mmu00280) pathways. In the context of glutamine-deprivation (or general nutrient deprivation towards the end of a batch culture), the emergence of pathways related to amino acid

degradation is not so surprising. In fact, a study by Wahrheit *et al* used dynamic metabolic flux analysis to dissect the impact of the availability of glutamine in the physiology of CHO cells, focusing on central carbon metabolism (Wahrheit, Nicolae and Heinzle, 2014). High-glutamine conditions were found to promote high flux toward the generation of TCA intermediary metabolites, such as aspartate and pyruvate, which are themselves substrates for the synthesis of lysine, valine, leucine and isoleucine. The same study found that glutamine-deprivation led to flux rearrangements in pyruvate and amino acid metabolism facilitating glutamine synthesis from glutamate by the enzyme glutamine synthetase. Another study by Duarte *et al.* looked at amino acid metabolism in GS-CHO cells, producer lines generated using the GS selection system and cultivated in glutamine-free media. This study shows that asparagine depletion leads to growth arrest in glutamine-free conditions and suggests compensatory mechanisms with higher consumption of other amino acids (Duarte *et al.*, 2014). A more recent study by Pavlova *et al.* shows how asparagine was able to rescue growth and survival of glutamine deprived cells (Pavlova *et al.*, 2018). Interestingly, the same study shows that asparagine is not used a substrate for glutamine biosynthesis, nor restores aspartate or ammonia levels in a panel of cells. These finding also support the hypothesis that other amino acids are being used to replace the lack of glutamine. In fact, a study by Hosios *et al.* showed the contribution of glutamine to total cell carbon and total cell nitrogen to be not more than 10% and that the importance of glutamine catabolism, glutaminolysis, for proliferating cells lies in an ability to generate metabolic products beyond biomass carbon (Hosios *et al.*, 2016). The same study shows how carbon originated from glutamine is mainly incorporated in other amino acids and how a large proportion of the carbon cells incorporate comes from amino acids other than glutamine. Regarding lipid metabolism, Cabodevilla *et al.* reported that beta-oxidation of fatty acids sustains cell viability in CHO cells (Cabodevilla *et al.*, 2013). Moreover, a recent study has linked nutrient deprivation to global changes of the lipidomic profile of cancer cells (Lisec *et al.*, 2019).

The third and last group of pathways enriched in both datasets includes the protein processing in endoplasmic reticulum (mmu04141) pathway. Predicted targets included in this pathway are related to processes such as ER-associated degradation and the ubiquitin ligase complex, as well as proteins involved in ER-nucleus signalling as a result of ER stress. A recent study on cancer cells reported downregulation of translation

machinery upon nutrient deprivation, namely glutamine and branched-chain amino acids, which is consistent with the observation ER stress pathways (Gameiro and Struhl, 2018). Earlier studies have also reported evidence of autophagy in CHO cells as a result of nutrient deprivation (Hwang and Lee, 2008; Han *et al.*, 2011), or in particular as a result of glutamine deprivation (Jardon *et al.*, 2012). As mentioned above, ER-stress can also regulate autophagy in a variety of contexts (Rashid *et al.*, 2015).

Taken altogether, our analysis resulted in the identification of pathways that, in a way or another, allow CHO cells to maintain energetic homeostasis and survive and grow despite the lack of nutrients. Our data provides further evidence that both, signal transduction pathways and lipid metabolism, two areas in which more research is needed in the CHO field, have the potential to originate targets for novel engineering approaches to further improvement of CHO cell lines for recombinant protein production.

3.5. Towards CHO-specific and robust miRNA profiling

When analysing pre-existing CHO profiling datasets, the technology used to generate the data must be taken into account. Microarrays are an established cost-effective profiling technology and require a short timeline in order to obtain robust and reproducible profiling results (Inwood, Betenbaugh and Shiloach, 2018). However, it is also well established that they have certain limitations, such as the need to have specific transcript sequence knowledge for the probe design stage, the risk of potential cross-hybridization of similar sequences due to homogenous hybridization conditions leading to false positive signal and their limitation in terms of absolute quantification of the detected transcripts. In addition, when profiling miRNAs, these limitations are exacerbated by the short sequence length of these non-coding RNAs, larger differences in GC content in between them and the difficulty to be able to distinguish between members of the same family, which differ only in a couple of nucleotides (Pritchard, Cheng and Tewari, 2012). In the CHO field, microarrays have extensively been used to profile both, coding and non-coding transcripts (Lin *et al.*, 2011; Bort *et al.*, 2012; Klanert *et al.*, 2016). However, early efforts to profile miRNA expression using microarrays in CHO relied on genome sequence and annotation from evolutionary related species such as human, mouse or rat, with early RNASeq data showing good sequence conservation for miRNA between CHO and closely related species (Bort, Stern and Borth, 2010; Hackl

et al., 2011). Altogether, the difficulties encountered to validate some of the targets outlined in this work should not be seen as an unexpected result. Indeed, since the first CHO-K1 genome in 2011 further efforts have been done to improve annotation of miRNAs capitalising on the more and more available genomic and transcriptomic studies in Chinese Hamster and Chinese Hamster Ovary cell lines (Diendorfer *et al.*, 2015). Maccani *et al.* in 2014 published their work in the identification of miRNA specific for a high producer phenotype in CHO cell and used a CHO specific microarray designed based on the CHO-K1 2011 genomic sequence (Maccani *et al.*, 2014). Recently, improvements in our knowledge of the CHO genome and transcriptome have also led to the design and validation of the first CHO-specific microarray for genome-wide mRNA profiling (Shridhar *et al.*, 2017). Looking forward, relying on CHO specific profiling technologies such as the above-mentioned microarrays will surely lead to more robust datasets to be available in the future.

From this study, the benefits of developing combined approaches using microarray and RNASeq data are apparent as well as the need to further improve the CHO specific annotation. In line with a recent meta-analysis of CHO transcriptomic data (Tamošaitis and Smales, 2018) there is a clear need for further validation of CHO datasets and standardized resources to compile currently available miRNAs expression datasets in CHO and develop bioinformatics pipelines that allow for a rapid analysis and target mining, thus providing a more effective and efficient platform to support cell engineering efforts in the future.

4. Conclusion

In this work two datasets were analysed in order to identify miRNAs with expression profiles that could be related to specific culture conditions or phases in order to use them as effector molecules in a transgene expression control system. In the first dataset, generated by microarray-based profiling of several cell lines with different glutamine requirements, a number of miRNA candidates with expression profiles correlating with ammonia build up in the media were identified. From these, hsa-miR-1275 and cgr-miR-1195 were validated by RT-qPCR. However, strong evidence of expression in other CHO published data could not be found and the low levels of expression detected should be taken with caution and make them unattractive candidates for effective transgene suppression. From the second dataset, the result of a collaborative effort to profile CHO

cells during a fed-batch under industrial conditions using RNA-Seq, 3 main clusters of miRNAs with distinct expression patterns were identified. Two of the clusters showed continuous increased or decreased expression over time, while the third one showed expression activation from the end of the exponential phase up to the death phase. Unfortunately, due to the timing of the dataset being available for analysis, none of the identified candidates was pursued for validation and tool development. In addition, the biological context behind the observed miRNA expression changes observed was investigated by pathway enrichment analysis, identifying relationships between changes in miRNA expression in both datasets and cell signalling by phosphorylation as well as lipid metabolism. These links are only hypothetical and based on predictions but represent a starting point for further investigation. Last but not least, a glutamine-free CHO-K1 cell line was generated by using a one single glutamine reduction step and was used for the validation work undertaken in this study. This work was also part of a phospho-proteomic study investigating glutamine deprivation (Kaushik et al. 2019, manuscript submitted). In order to further investigate the use of miRNAs as genetic switches in CHO cells, two different approaches were pursued in the following chapter. First, the use of miRNA-sponges as tools for transgene expression modulation in a semi-artificial system, using sponges for endogenous miRNAs coupled with the use of miRNA mimics to simulate controlled changes in miRNA expression was investigated. Following that, the use of an external intervention such as shifting culture temperature to generate temperature-induced changes in miRNA expression and the use of sponges specific to temperature sensitive miRNA to transgene expression were investigated.

CHAPTER 4

Temperature-induced control of transgene expression using miRNA sponges

1. Introduction

1.1. Temperature shift as a way to control transgene expression

Shifting the bioreactor temperature from the optimal 37°C to sub-physiological lower temperatures (31-33°C) has been a widely used strategy to improve recombinant protein productivity at the expense of cell growth, both in mammalian systems (Masterton and Smales, 2014; Xu *et al.*, 2019). In particular, the effect of reducing culture temperature has been extensively studied in CHO cells over the past decade. The main phenotypic changes are related to cell growth and viability, with cells being arrested in G1 phase and prolonged together with increased productivities and quality attributes. At a molecular level, gene expression in mammalian cells has been shown to be affected by temperature downshift at a transcriptional and post-transcriptional level by various mechanisms. These include: RNA remodelling by modulation of splicing efficiency or alternative splicing (Gotic *et al.*, 2016; Tzani *et al.*, 2020), alternative promoters activated under low temperature conditions (Al-Fageeh and Smales, 2009, 2013; Sumitomo *et al.*, 2012), temperature sensitive differential expression of ncRNAs (miRNAs and lncRNAs) (Emmerling *et al.*, 2016; Motheramgari *et al.*, 2019).

Temperature is also a very easy parameter to change in comparison with other bioprocess relevant parameters such as pH or the concentration of metabolites in the media such as carbon or nitrogen sources. In the context of this study, the aim was to capitalise on the effects of shifting culture temperature on microRNA expression in order to develop a temperature-dependent transgene control system.

1.2. miRNA-sponges as molecular tools for transgene expression modulation

miRNA sponges, also known as decoy vectors were proposed almost a decade ago as an alternative method to miRNA inhibitors to perform loss-of-function studies on miRNA targets (Ebert, Neilson and Sharp, 2007). The idea behind it was to express synthetic UTRs downstream of a reporter gene, containing multiple bulged miRNA binding sites. The incorporation of that bulge aimed to inhibit the action of the endonuclease Slicer following binding of the miRISC complex to the target mRNA transcript, thus impairing the degradation of the decoy transcripts. The result of overexpressing these synthetic 3'-UTR sequences was to sequester available miRNAs, titrating them away from their endogenous cellular targets. To date, several studies in CHO have used this technology

to transiently and stably knock-down specific miRNAs to positively affect cell phenotype (Sanchez *et al.*, 2013; Kelly *et al.*, 2014, 2015; Costello *et al.*, 2018). There have also been efforts toward conditionally modulating miRNAs by using inducible sponges (Costello, Lao, Gallagher, *et al.*, 2019). In this chapter, the use of sponges in CHO was investigated from a different angle. Instead of using synthetic UTRs as decoy vectors for miRNA depletion, we explored the possibility to obtain modulated transgene expression in CHO cells by means of having it under the control of a synthetic 3'-UTR targeted by endogenous miRNAs. As outlined in the introduction, this approach has been successfully applied in other systems.

In a first stage, the aim was to establish the experimental framework for the implementation and testing of miRNA-sponges as sensors that can react to changing miRNA concentrations in CHO cells. The possibility to establish a link between endogenous miRNA levels and transgene expression as well as the level of transgene expression control that can be obtained by using miRNA-sponges / synthetic 5'-UTR in CHO cells were investigated. Following that, specific miRNAs exhibiting differential expression upon temperature shift were identified using RNASeq data and validated in a robust temperature-shift cell culture model. Following that, expression vectors carrying miRNA sponges targeting specific for the validated targets were constructed and tested to evaluate their ability to modulate reporter gene expression upon temperature shift.

2. Results

2.1. Construction of miRNA-sponges targeting endogenous CHO miRNAs

Initially, two model endogenous CHO miRNA were used in this work: cgr-miR-204-5p and cgr-miR-409-3p. For cgr-miR-204-5p, two types of binding sites were used: bulged binding sites, with a mismatch sequence around the 10/11 nucleotides to avoid endonuclease cleavage by the miRNA induced silencing complex (miRISC), and fully complementary binding sites, expected to drive transcript degradation upon miRNA binding (Figure 16). For cgr-miR-409-5p only bulged binding sites were tested. As negative control sponge, an A-rich binding site was used, named from here on NC-sponge. Sequences for the different binding site oligos used are detailed in Table 8. Sponge vectors were constructed as described in Methods (Section 1.5.5). Sponges of

various sizes ranging from 1 to 7 inserts (2 to 14 binding sites) were obtained for each sequence. A sample of them per sequence after plasmid isolation and PCR amplification of the sponge region is shown in Figure 17.

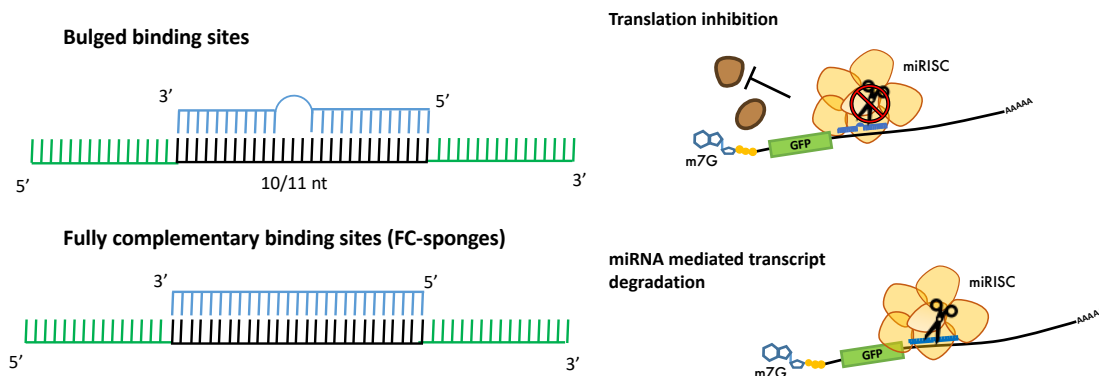


Figure 16 - Schematic representation of the two types of miRNA binding sites used: bulged and fully complementary, and their associated mRNA transcript interaction regulatory mechanisms

Table 8 - Oligo sequences used to construct miRNA sponges. For cgr-miR-204 two types of binding site - bulged and fully complementary (FC) were used. Only bulged binding sites were designed for cgr-miR-409-3p. A negative control sponge was constructing using A-rich binding sites (NC-sponge). The overhanging ends for cloning using KflI (SanDI) site and spacers are marked in bold. miRNA binding sites regions are underlined.

Sequence	
NC-spg	
Forward	GTCCC <u>AAGTTTTCAGAAAGCTAAC</u> CCGGAAGTTTTCAGAAAGCTAAC AGG
Reverse	GACCTGTTAGCTTTCTGAAA ACTTCGGGTTAGCTTTCTGAAA ACTTGG
204-spg	
Forward	GTCCC <u>AGGCATAGGACTGAAAGGGAA</u> AATTAGGCATAGGACTGAAAGGGAA AGG
Reverse	GACCTTCCCTTTCAGTCCTATGCCT AATTTCCCTTTCAGTCCTATGCCT GG
204-FC-spg	
Forward	GTCCC <u>AGGCATAGGATGACAAAGGGAA</u> AATTAGGCATAGGATGACAAAGGGAA AGG
Reverse	GACCTTCCCTTTCAGATCCTATGCCT AATTTCCCTTTCAGATCCTATGCCT GG
409-spg	
Forward	GTCCC <u>AGGGGTTCACTTTCAACATT</u> CAATTAGGGGTTCACTTTCAACATT CGG
Reverse	GACCCGAATGTTGAAAGTGAACCCCT AATTGAATGTTGAAAGTGAACCCCT GG

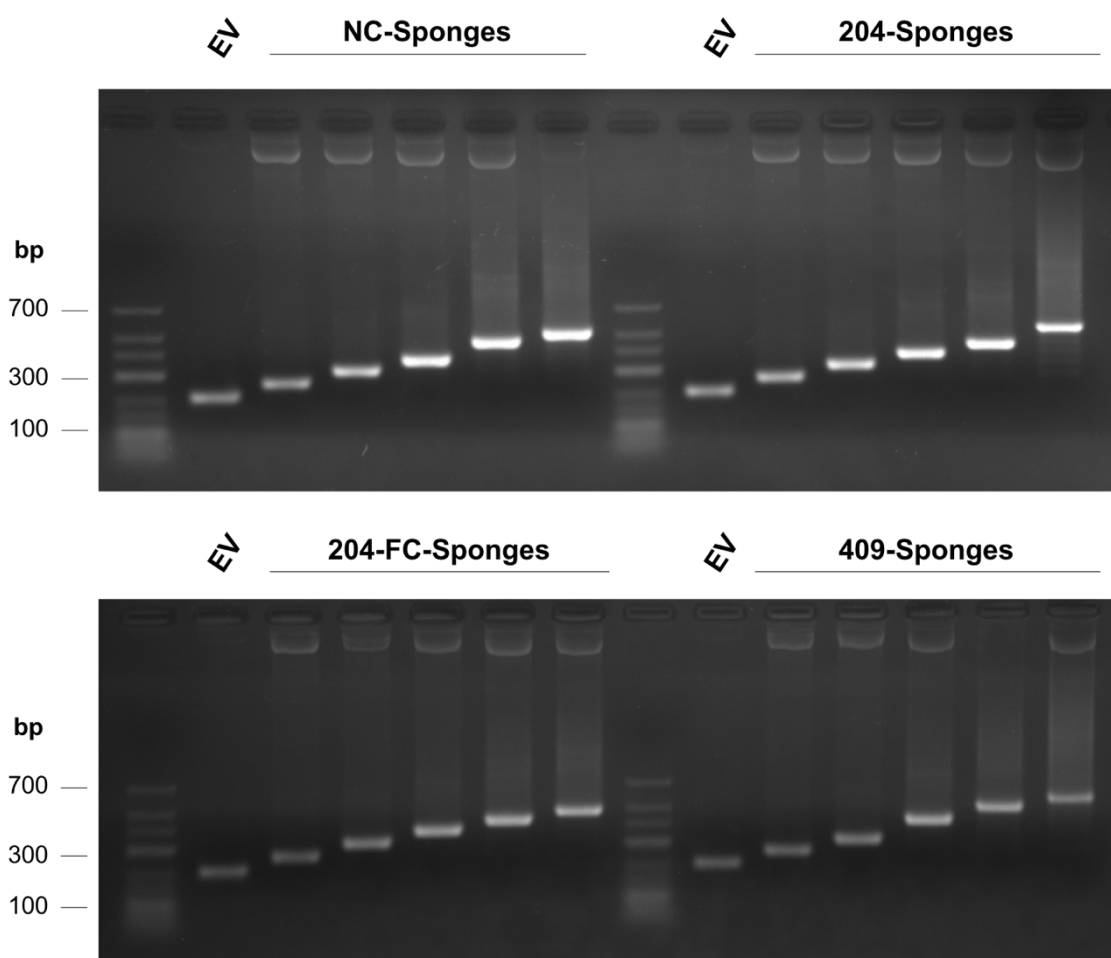


Figure 17 - Agarose gel showing different sized sponges after PCR using flanking Tet9 primers. The PCR product obtained from an empty vector (EV) control is shown as reference.

2.2. Linking transgene expression and miRNA expression using miRNA sponges

As a first step, the possibility of correlating expression levels of a particular microRNA with the expression levels of a stably integrated reporter gene under the control of a sponge sensor was investigated. In house generated CHO DP12 cells lines constitutively expressing GFP-miRNA sponge constructs, either for the negative control mock sequence sponge (NC) or for cgr-miR-204 were used in this preliminary study. The rationale behind using 204-spg cell line which contained a GFP-transcript with 12 binding sites targeted by cgr-miR-204-5p was that expression of this miRNA had been previously correlated with growth rate (Costello *et al.*, 2018). Hence, miRNA-204 was hypothesised to change over time with cells transitioning from fast growth rate during exponential phase to a slow growth rate during stationary phase, making it interesting to investigate the possibility of correlating these changes to changes in GFP protein levels.

Batch cultures for both cell lines, NC-sponge and 204-sponge, were set up and samples were taken at different time points in order to monitor the expression levels of the GFP reporter gene, at both protein and transcript level, as well as their expression levels of the targeted miRNA. Cells were seeded at 0.2×10^6 cells/mL and grown under standard conditions for a period of 10 days. Every 48 hours, cell density and GFP expression were measured using flow cytometry and samples for RNA extraction were collected. Levels of GFP mRNA transcripts and cgr-miR-204 were analysed by RT-qPCR as outlined in Methods Section 1.3.2. This experiment did not allow biological replication as cells slowly lost GFP expression. However, observational differences are reported below.

In terms of growth, both cell lines exhibited similar rates and cell viability (Figure 18 – A and B). In terms of reporter gene expression, the DP12-NC-Spg cell line showed higher levels of GFP fluorescence throughout the culture when compared with the DP12-204-Spg cell line (Figure 18 – C and D). This was expected and was a good indication of specific binding of endogenous miRNA driving downregulation of the reporter gene.

In terms of GFP mRNA transcript levels, in the DP12-NC-cells levels of GFP transcripts fluctuated throughout the culture (Figure 18 – E – Blue) without any particular pattern. However, in the DP12-204-Spg cells, a progressive increase of the de2GFP mRNA transcript abundance, was observed over time (Figure 18 – E – Red). When comparing the levels of GFP mRNA transcript between the two cell lines, lower mRNA levels were observed in the DP12-204-Spg cell line in comparison with the DP12-NC-Spg for the first two time points, while towards the end of the culture, both cell lines exhibit similar levels of GFP mRNA transcript (Figure 18 – F).

As hypothesised, cgr-miR-204 expression levels increase over time, negatively correlating with growth rate. Expression profiles followed a similar pattern in both cell lines, with expression low during the early days of culture, and increasing over time. Also as expected, due to the effect of the miRNA-sponge, cgr-miR-204 was down-regulated in the DP12-204-Spg cell line in comparison with the DP12-NC-Spg cell line for all the analysed time points (Figure 18 – H).

Correlations between growth, gene expression and miRNA were investigated for the two cell lines separately as well as between the two cell lines by calculating Person's Correlation Coefficient (PCC) matrices on the different datasets.

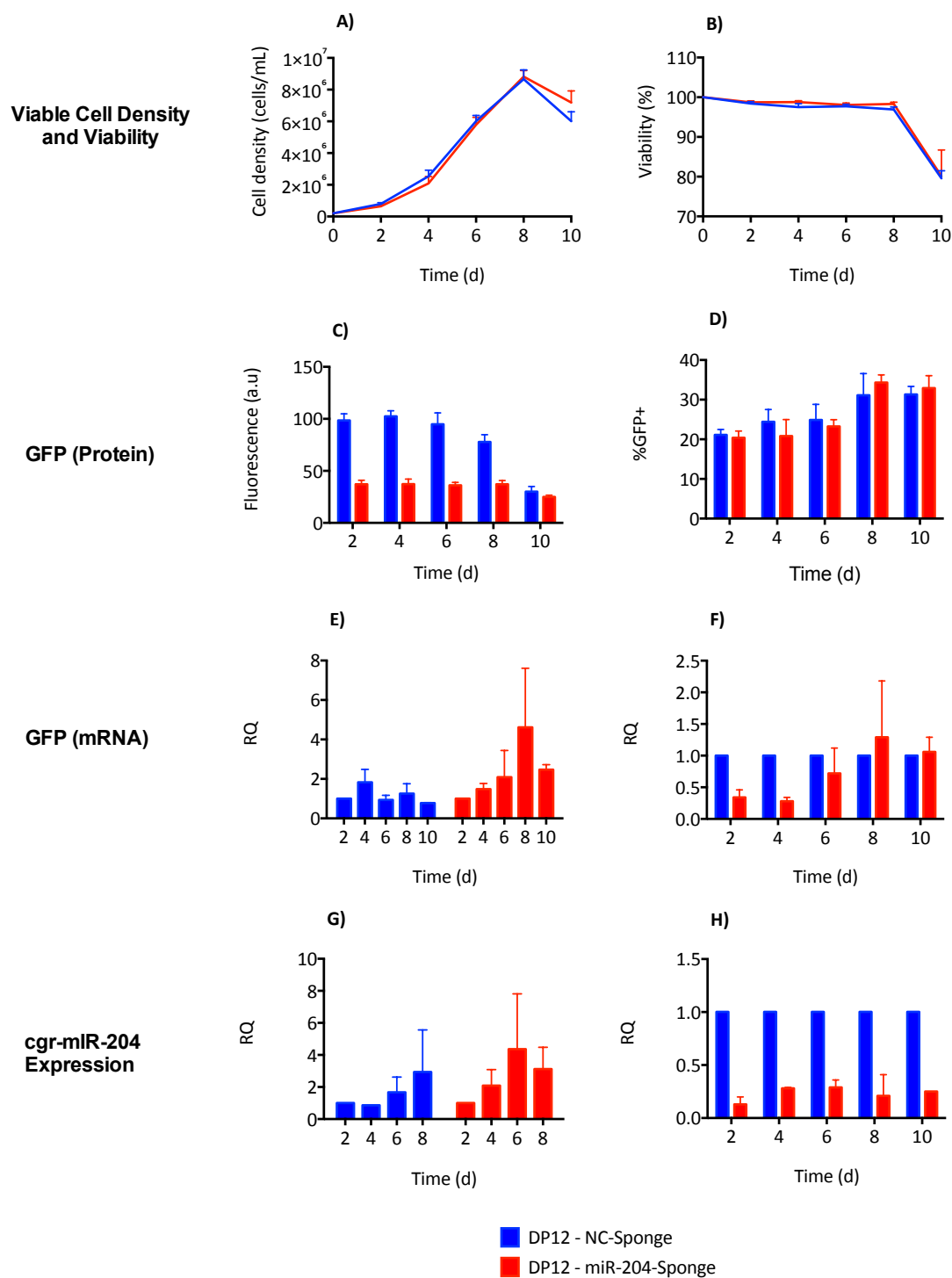


Figure 18 - Evaluation of the expression of the de2GFP reporter gene in NC-sponge and 204-sponge DP12 cell lines. A) Viable cell density, B) Cell Viability. GFP expression: C) Mean Fluorescence Intensity and D) Percentage of GFP+ cell, relative quantification of GFP mRNA transcripts and relative quantification of cgr-miR-204 expression: E) and G) show expression of GFP and cgr-miR-204-3p normalised versus day 2 expression for each cell line F) and H) show relative expression of GFP and cgr-miR-204 in the DP12-204-spg cell line versus expression in the DP12-NC-spg cell line at each time point. Each point represents the average of 3 technical replicate (N=1) GFP measurements and error bars represent their standard deviation.

Looking at the data for the NC-sponge cell line, strong positive correlations were found between cgr-miR-204 and culture time (PCC = 0.92) or viable cell density (VCD) (PCC = 0.88). GFP protein expression (fluorescence) was negatively correlated with Time (PCC = -0.86) but no correlation between GFP miRNA transcript levels and Time/VCD was found (Table 9 – A).

Table 9 - Correlation analysis between culture time, growth, GFP expression (mRNA and protein) and miR-204 for the DP12-NC-Spg cell line (A), the DP12-204-spg cell line (B). Pearson's correlation coefficients > 0.6 and < -0.6 are highlighted and color coded: green for positive correlations and orange for negative correlations.

A)						B)					
	Time	VCD	GFP (protein)	GFP (mRNA)	miR-204		Time	VCD	GFP (protein)	GFP (mRNA)	miR-204
Time	1.00	0.84	-0.86	-0.32	0.92	Time	1.00	0.91	-0.72	0.72	0.88
VCD		1.00	-0.47	-0.13	0.88	VCD		1.00	-0.39	0.90	0.83
GFP (protein)			1.00	0.53	-0.79	GFP (protein)			1.00	-0.11	-0.62
GFP (mRNA)				1.00	-0.37	GFP (mRNA)				1.00	0.51
miR-204					1.00	miR-204					1.00

Similarly, for the 204-spg cell line, strong positive correlations was found between cgr-miR-204 expression and culture time or viable cell density (VCD) and GFP expression at protein (fluorescence) was negatively correlated with Time / VCD (Table 9 – B). Interestingly, in this cell line, a strong positive correlation between mRNA levels and time (PCC = 0.72) and VCD (PCC = 0.90) was found as well as a positive correlation between cgr-miR-204 and GFP mRNA transcript levels was observed.

2.3. Transient expression of miRNA-sponge sensors - transfection optimisation

Despite observing changes in the expression of miRNA cgr-miR-204-3p over time in the previously described stable cell lines, no direct link between these and GFP expression could be established due to the strong inhibitory effect driven by the used sponge constructs. In order to have better control of the amount of sponge present in the cell and in an attempt to find a range in which links GFP changes and miRNA changes could be established, these were further evaluated in a transient expression model.

In order to define optimal conditions for transient expression experiments, transfections were carried out under two different conditions: constant total DNA amount of 1000 ng /10⁶ cells and variable amounts of GFP plasmid from 0 to 1000 ng with the corresponding amount of empty vector DNA to maintain the total amount constant and variable DNA amount / 10⁶ cells as shown in Figures 4 to 7.

The rational for that was to evaluate the effect of varying the amount of GFP plasmid used on the readable outputs of GFP expression (%GFP positive cells – Figure 19, Mean Fluorescence Intensity and Median Fluorescence Expression – Figure 20) without varying the total DNA amount used for transfection as well as to determine the effects of the plasmid dose on cell viability (Figure 21).

As expected, reducing the amount of total DNA plasmid led to a reduction in transfection efficiency ranging from 60% GFP+ for 1000ng of transfected plasmid to below 5% of detectable transfected cells for 62.5 ng of DNA (Figure 19 – right bar plot). Similarly, lower amounts of transfected DNA resulted in lower amounts of fluorescent signal detected, with low amount of GFP plasmid DNA resulting in fluorescence levels undistinguishable from the background observed in the no GFP plasmid control (Figure 20 – right bar plot).

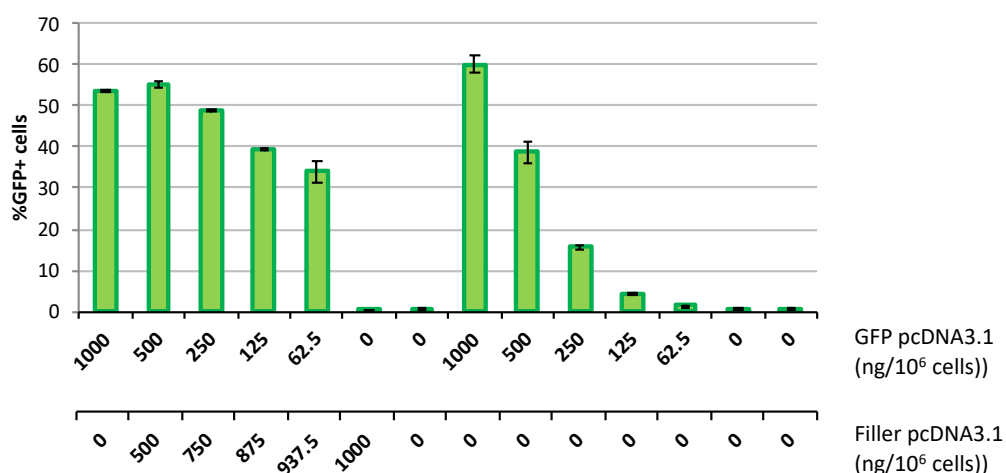


Figure 19 - Percentage of GFP+ cells when transfected with a different amount of GFP plasmid / filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

By maintaining the total amount of DNA constant at 1000 ng, and changing the proportion of GFP plasmid: empty vector plasmid, it was possible to detect substantial fluorescence signal even with low amount of GFP plasmid while maintaining a transfection efficiency above 30% (Figure 19 and Figure 20 left bar plots). However, when maintaining a constant amount of DNA of 1000 ng, amounts of GFP between 250 and 1000 ng led to a significant cytotoxic effect as observed from viability measurements (Figure 21). Having observed that transfecting cells with a total amount of 500 ng of GFP did not cause observable cytotoxic effects, the same experiment was repeated by using a constant total amount of DNA of 500 ng. Using this total amount of DNA plasmid, it was possible to transfect even lower amounts of GFP plasmid without sacrificing transfection efficiency (Figure 22). It was also possible to reduce the amount of GFP plasmid transfected to very low levels while maintaining a readable output in terms of GFP fluorescence intensity (Figure 23). Detrimental effects on cell viability were again observed for the 500 ng transfections, while good viabilities were obtained for the remaining DNA amounts tested (**Figure 24**). A total amount of 500 ng of DNA with 250 ng of GFP plasmid and 250 ng of empty vector (1:1 ratio) were selected as optimal and used in all further experiments.

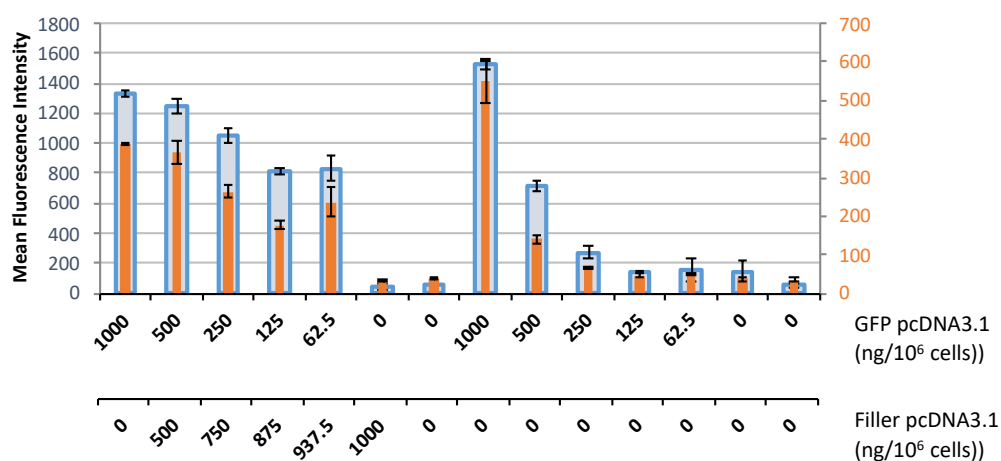


Figure 20 – Mean fluorescence intensity (blue bars) and median fluorescence intensity (Red bars) representing GFP expression in cells 24h after transfection with different amounts of GFP plasmid + filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

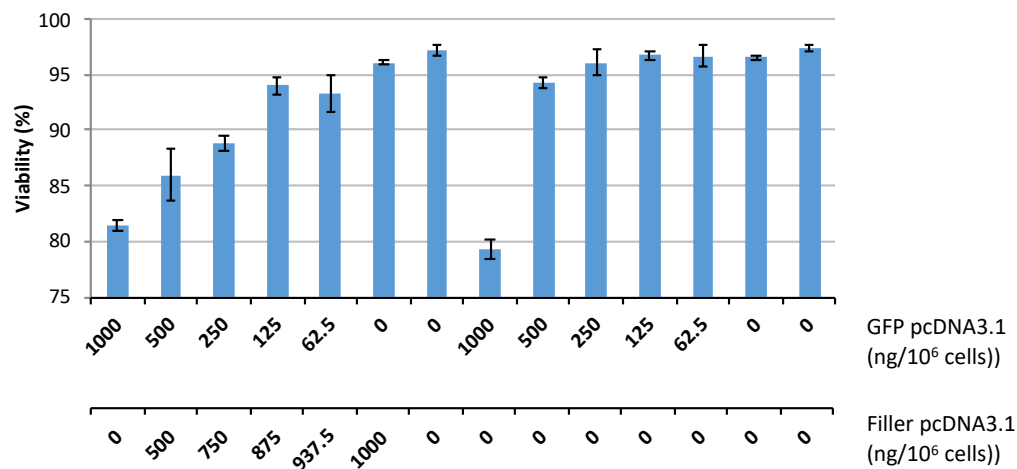


Figure 21 – Cell viability 24h after transfection with different amounts of GFP plasmid + filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

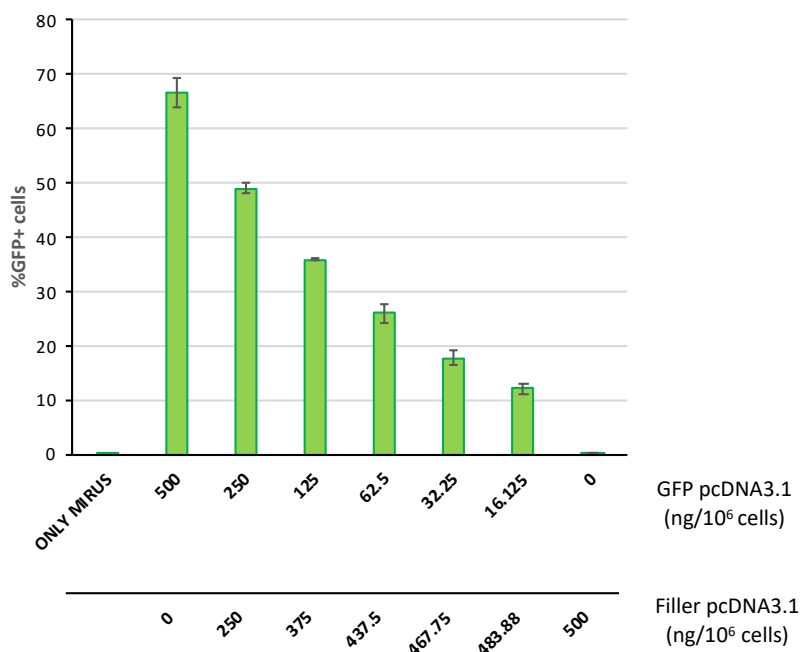


Figure 22 - Percentage of GFP+ cells when transfected with different amounts of GFP plasmid / filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

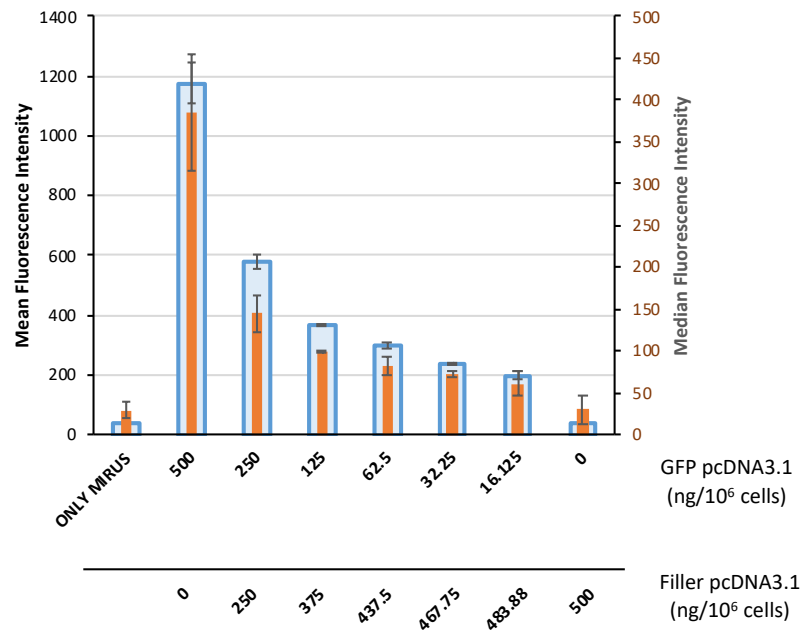


Figure 23 – Mean fluorescence intensity (blue bars) and median fluorescence intensity (Red bars) representing GFP expression in cells 24h after transfection with different amounts of GFP plasmid + filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

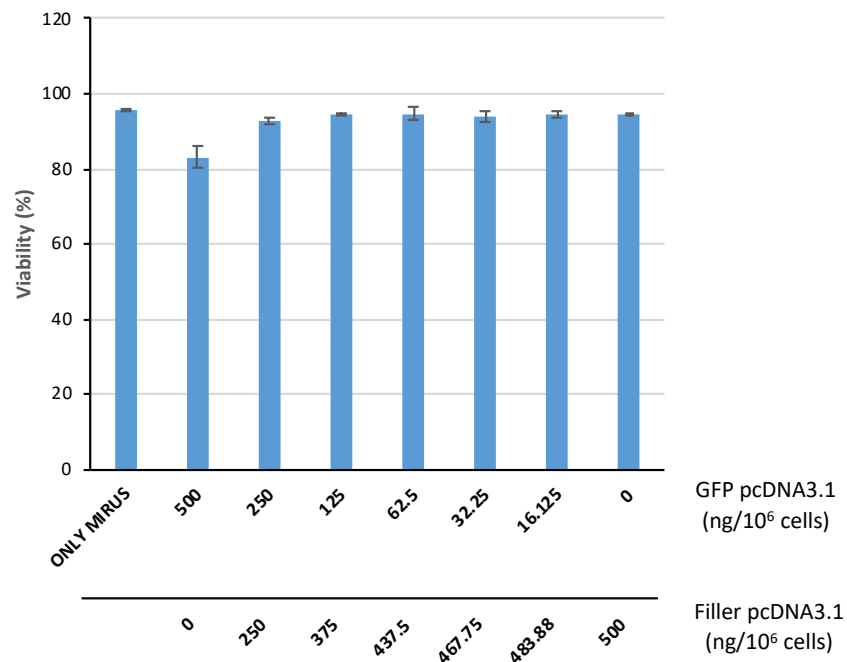


Figure 24 – Cell viability 24h after transfection with different amounts of GFP plasmid + filler DNA / 10⁶ cells. Each bar represents the average of 3 technical replicates and error bars represent standard deviation.

2.4. Effect of binding site number and DNA amount on transgene expression and sponge/sensor performance

The aim of these series of experiments was to investigate:

- the effect of the number of binding sites present in a miRNA sponge / synthetic 3'-UTR
- the effect of varying the degree of complementarity between the binding site and the targeted miRNA
- the effect of the plasmid dose on the ability of endogenous miRNAs to modulate expression of the reporter transgene.

In order to define the appropriate GFP plasmid dose for the evaluation of the different miRNA sponges, cells were transfected with varying amounts of GFP sponge plasmids while keeping the total amount of DNA constant as previously described with empty vector used as filler DNA. GFP expression evaluated 24h after transfection. Sponges for two different sequences, NC and cgr-miR-409-3p, and of two different sizes (2 and 10 binding sites) were evaluated.

Table 10 - Effect of sponge UTR on GFP expression as percentage of change in GFP Mean fluorescence intensity for each tested sponge plasmid. Coloured cells mean significant increase (green) or decrease (red) in GFP signal when compared with the GFP pcDNA3.1 no sponge control. White cells indicate no significant differences. Data was calculated using the ratio sponge/control from the average of 3 biological replicate transfections and error bars represent their standard deviation. Significant differences were evaluated $p\text{-value} < 0.05$ using a homoscedastic 2-tailed T-Test.

		GFP pcDNA3.1	NC-spg		409-spg	
	Nº BS →	0	2	10	2	11
DNA amount	500	n/a	10%	-15%	23%	-63%
	250	n/a	0%	-20%	13%	-58%
	125	n/a	-9%	-33%	10%	-56%
	62	n/a	3%	-15%	14%	-48%
	31	n/a	8%	-28%	9%	-77%
	EV	n/a	1%	2%	0%	-4%

A “no sponge” GFP plasmid (0 binding sites) was always transfected as no-3'-UTR positive control and used as reference for the normalisation. Figure 25 shows the percentage of GFP positive cells measured for each conditions and Figure 26 shows mean fluorescence intensity for each whole population (GFP+ and dark events) for each sequence and condition. Table 10 shows the calculated GFP signal difference for each

sponge when compared with the GFP pcDNA3.1 control. Adding a long 3'UTR significantly decreased GFP signal in most conditions for the NC-sponge with 10 binding sites (from -15% to -33%) and the 409 sponge with 11 binding sites (from -48% to -77%).

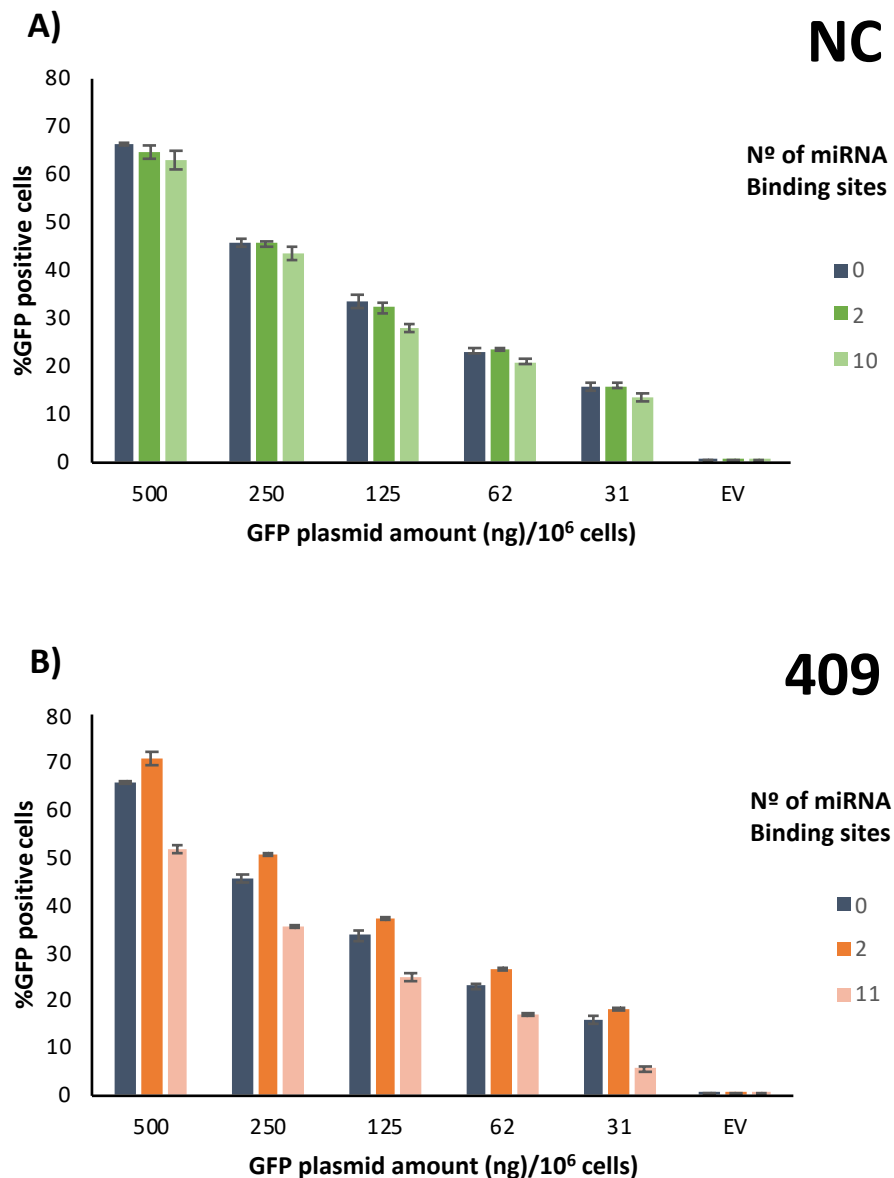


Figure 25 – Effect of gene dose (GFP plasmid amount) and number of miRNA-binding sites on the 3'-UTR sensors:: a negative control NC (A), cgr-miR-409-5p (B). Percentage of GFP positive cells for each transfected population. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation. (*) Significant differences with p -value<0.05 using a homoscedastic 2-tailed T-Test.

Surprisingly, a short 3'UTR (2 binding sites) resulted in an increase of GFP signal at the maximum tested DNA amount for the NC-sponge and for three of the five tested DNA amounts for the 409-sponge. Using 250 ng of GFP DNA and 250 ng of filler vector

allowed good transfection efficiencies and with the low impact on cell viability reported before and therefore, those conditions were used in subsequent experiments.

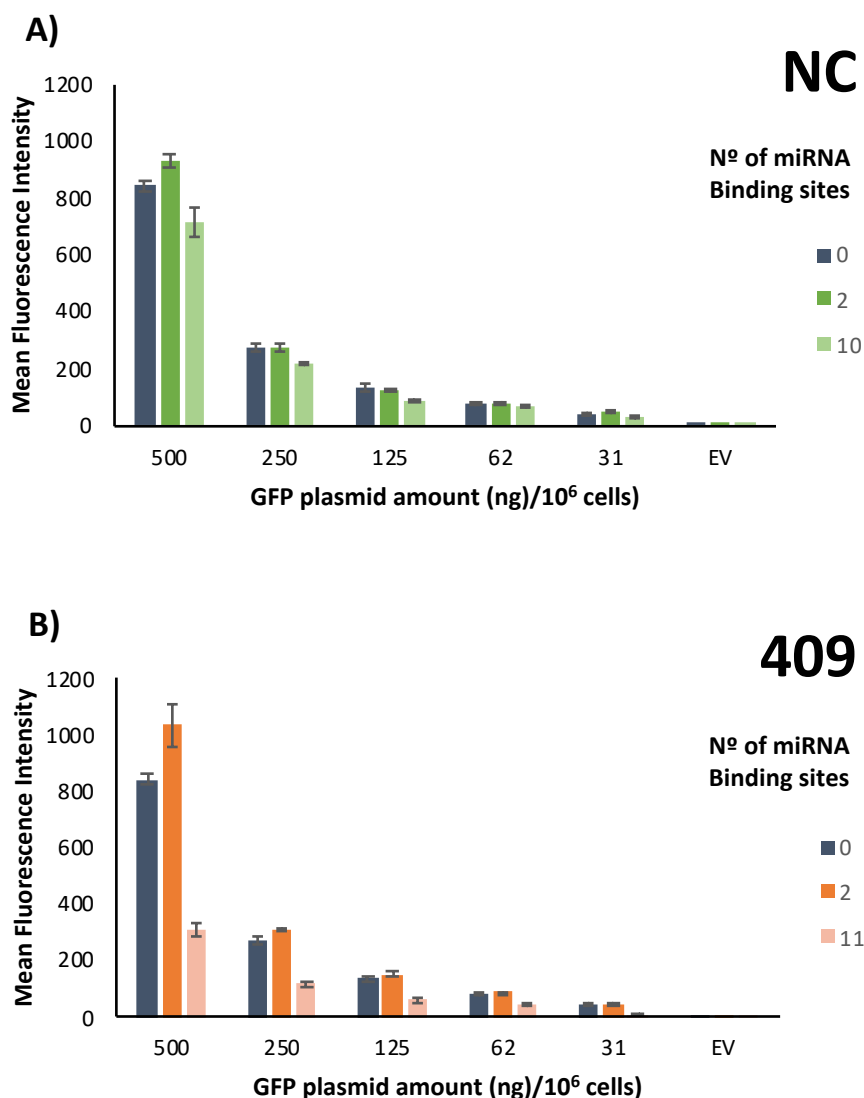


Figure 26 – Effect of gene dose (GFP plasmid amount) and number of miRNA-binding sites on the 3'-UTR sensors:: a negative control NC (A), cgr-miR-409-5p (B). Mean Fluorescence Intensity for all measured events represents GFP expression in the transfected population. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation. (*) Significant differences with $p\text{-value} < 0.05$ using a homoscedastic 2-tailed T-Test.

In order to further explore the possibility of obtaining gradual levels of modulation by modifying the number of binding sites contained in the sponge, 409-sponge plasmids with binding sites ranging from 2 to 13 were tested and compared with NC-sponges of similar sizes. Figure 27 shows the GFP expression data (percentage of GFP positive cells

in the top plot and Mean fluorescence intensity in the bottom plot) obtained for this experiment 24h after transfection.

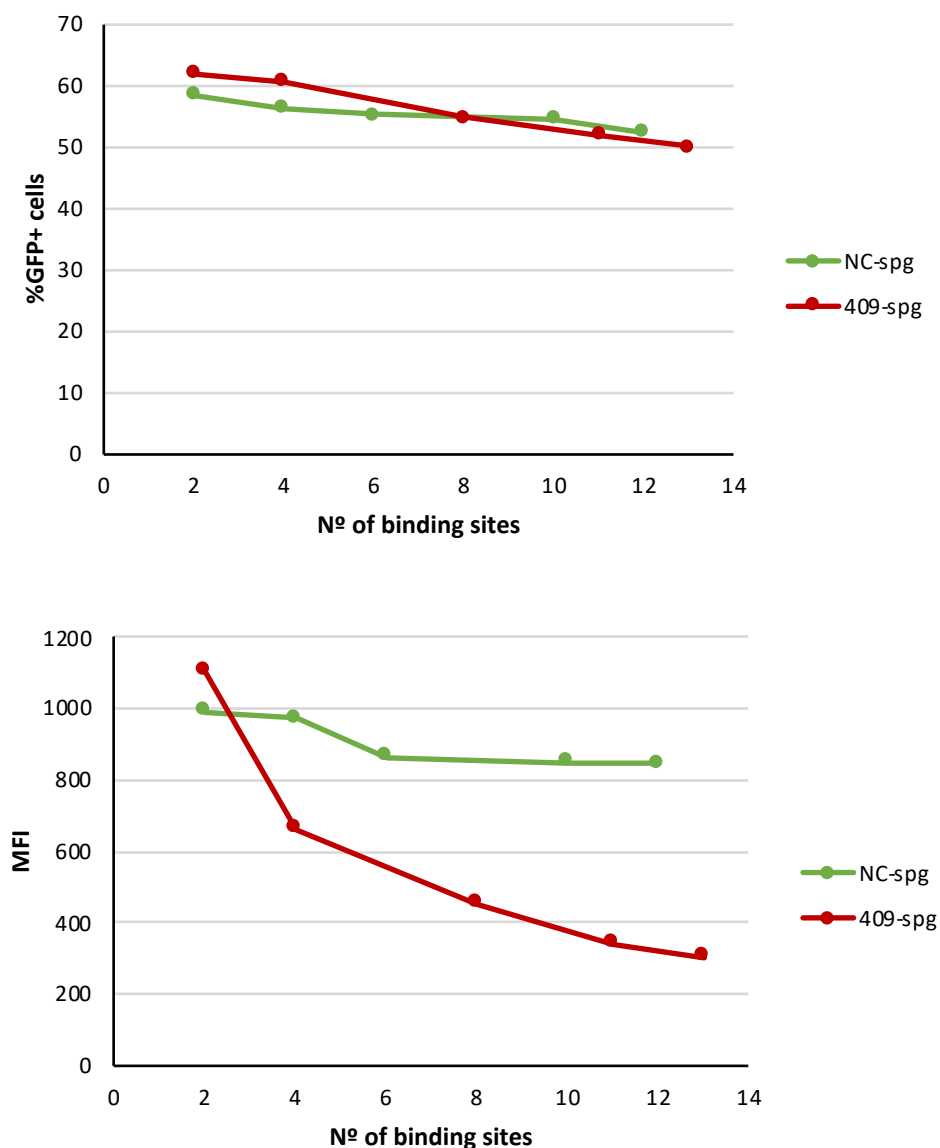


Figure 27 - Effect of the number of 3'UTR miRNA binding on GFP output. Sponges targeting *cgr-miR-409-3p* and NC-sponges with varying numbers of binding sites were transfected in CHO-K1 cells. **Top** – Percentage of GFP positive cells 24h after transfection. **Bottom** – Mean Fluorescence Intensity (MFI) for the GFP+ population 24h after transfection. Each point represents the average of 3 biological replicate transfections and error bars represent their standard deviation.

As observed before, increasing the number of binding sites in the sponge had a negative impact on GFP expression, with a similar reduction as observed before for the long sponges: ca. 20% drop in GFP fluorescence between the short NC-sponge (2 binding sites) and long NC-sponge (12 binding sites) and ca. 60-70% drop in GFP fluorescence between the short 409-sponge (2 binding sites) and long 409-sponge (13 binding sites)

for the long 409-sponges. Interestingly, in the NC-sponges, 6 binding sites were enough to cause the observed reduction in GFP signal (ca. -20%), with only marginal reduction observed for 2 and 4 binding sites. For the sponges targeting cgr-miR-409-3p, it can be observed how increasing the number of available miRNA binding sites increases the inhibitory effect of the sponge on the measured GFP output, except for the sponge with 2 binding sites, which again showed increased GFP signal. For the other sponges, increasing the length of the 3'-UTR did not proportionally impact GFP levels, with saturation observed beyond 8 binding sites.

2.5. Effect of binding site number and complementarity on sponge/sensor performance with changing concentrations of miRNA

In order to evaluate the effect of having bulged or fully complementary binding sites on the ability of miRNAs to modulate the expression of the transgene (GFP), sponges of different length for the negative control sequence and for cgr-miR-204-5p with bulged binding sites or fully complementary binding sites were constructed and transfected in CHO cells. In order to simulate changes in miRNA expression, synthetic miRNA mimics for each sequence (NC or cgr-miR-204) were subsequently transfected and GFP expression measured after 24h.

As observed previously, the GFP signal from sponges targeting a specific miRNA (cgr-miR-204-5p in this case) was 20-40% lower than the GFP signal measured from the NC-sponge, indicating specific binding of the miRNA down-regulating GFP levels (Figure 28 – UTF samples). As previously observed, increasing the number of binding sites led to a decrease in GFP expression for the three tested sponge sequences. This is visible when comparing the measurements done in cells transfected with sponge but not transfected with mimic (Figure 28 – UTF samples). Interestingly, NC-sponge with only 2 binding sites showed a different repression profile with increasing concentration of mimic, being less reactive at low mimic concentrations than NC-sponge with 4 or more binding sites, which all behave similarly. All 3 tested 204-sponges, with 4 to 8 binding sites responded similarly to increasing concentrations of mimic by repressing GFP to similar levels.

A stronger inhibition effect was observed from fully complementary binding sites when compared with the bulged binding sites. A similar percentage of GFP positive cells (30-40%) for the untransfected and the negative control conditions but with lower fluorescence intensity values for measurement done on the GFP+ populations as well as

on the whole population (Figure 28 – 204-spg versus 204-FC-spg). Differences between differently sized 204-FC sponges became minimal when increasing the amount for transfected mimic to concentrations ranging from 2 to 50 nM..

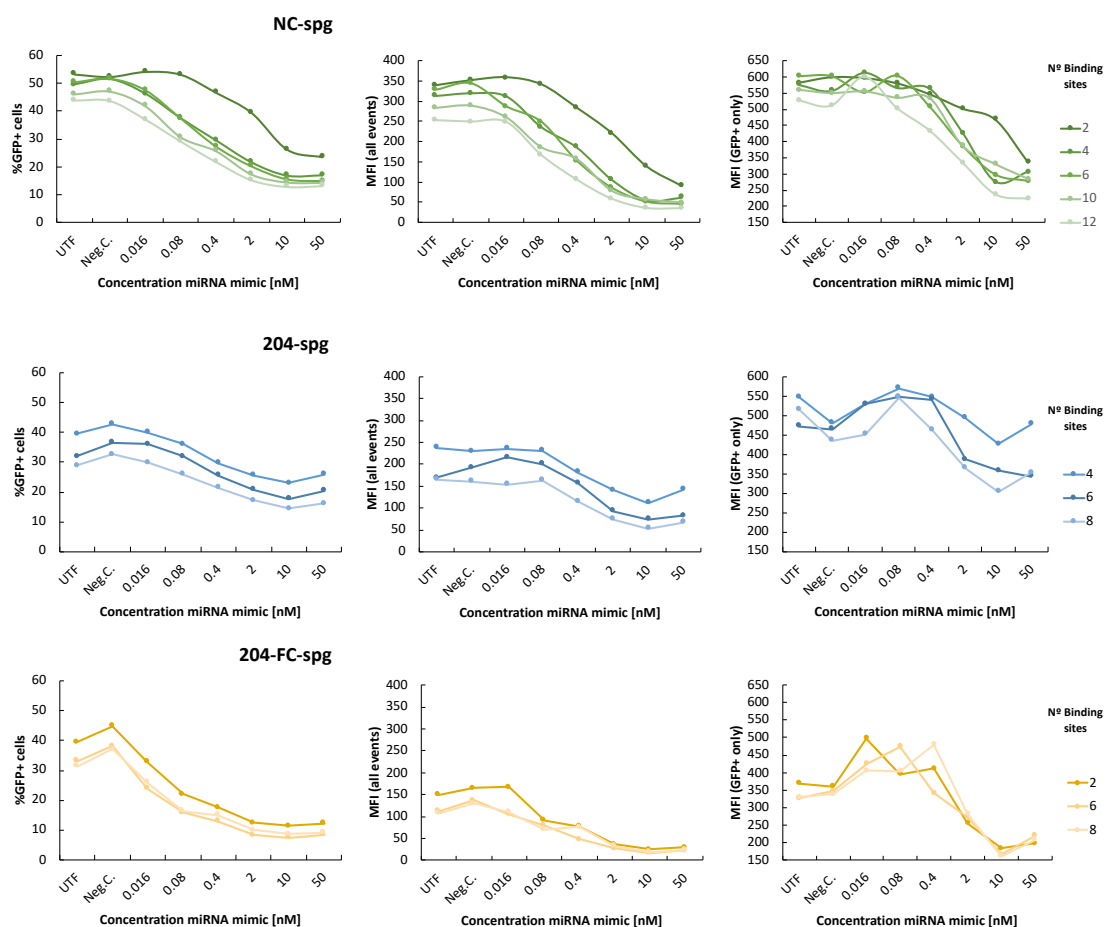


Figure 28 – Effect of varying concentrations of miRNA mimics on GFP expression modulation by sponges with different number of binding sites. Percentage of GFP + cells (left), mean fluorescence intensity for all events (middle) and mean fluorescence intensity for only GFP+ cells. A negative control sequence (NC) or *cgr-miR-204* at variable concentrations were tested. As a negative control mimic for each condition (Neg.C.) a non-complementary mimic was used. Untransfected samples (UTF) correspond to cells transfected with sponge but not transfected with mimic. Each point represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

This was not observed for bulged 204-sponges, with different sized sponges still showing different levels of repression depending of the number of binding sites under varying concentrations of mimics. Sponges with fully complementary binding sites (204-FC-spg) were also more reactive to varying concentrations of miRNA mimics, even when low concentrations of mimics (0.08 nM and 0.016 nM) were used. Strong decreases of %GFP+ cells, with up -75-80% when compared with the negative control condition, and MFI values, with up to 66% reduction when compared. These were less pronounced in

sponges with bulged binding sites (204-spg), which only showed comparable reductions of GFP upon transfection with higher concentrations of miRNA mimics .

Increasing the amount of transfected miRNA mimic led to an increased repression of GFP expression. This was more apparent for the NC-sponges than for the 204-sponges and 204-FC-sponges, presumably due to the presence of the endogenous miR-204 that was already repression GFP expression from these constructs.

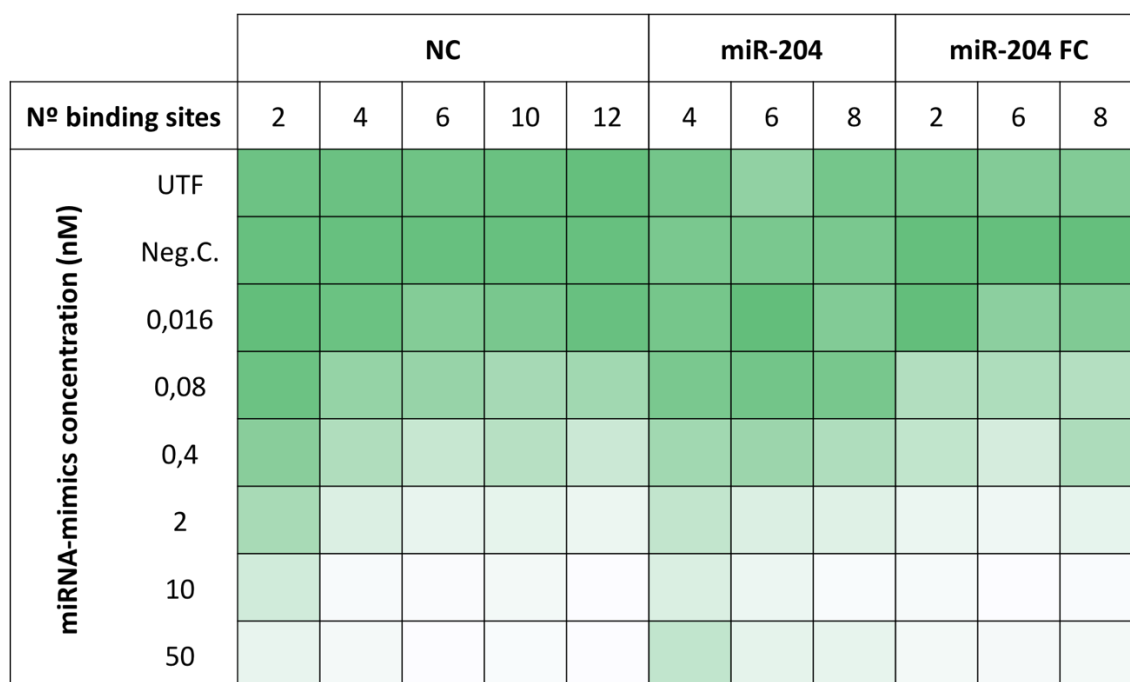


Figure 29 – Heat map representing the expression of GFP under the control of a miRNA-sponge sensor specific to a synthetic negative control sequence (NC) or cgr-miR-204 at variable concentrations. As a negative control mimic for each condition (Neg.C.) a non-complementary mimic was used.

Figure 29 provides a heat map overview of the GFP expression levels obtained by placing the reporter gene under the control of the different sponges and mimic concentrations. Data was normalized for each plasmid using its negative control as a reference for visualization. It can be seen that by using miRNA sponge, either targeted by a synthetic miRNA (NC) or by endogenous miRNA (204 and 204-FC), transgene expression levels could be adjusted at different levels depending on the number of binding sites used and the concentration of specific miRNA.

2.6. RNA-Seq profiling of CHO cells exposed with mild-hypothermia reveals temperature sensitive microRNAs

The data collection part of this project up to the selection of miRNA targets was not performed by the author of this thesis. Laboratory work and analysis were performed as outlined in as described in the publication by Tzani et al. (Tzani *et al.*, 2020). The miRNA dataset originating from that study was analysed by Dr. Colin Clarke (NIBRT/UCD) and miRNA targets were provided as a starting point for the work described in this section. For reference a brief summary of that work is provided.

In order to identify miRNAs, whose expression was influenced by subjecting the cells to temperature shift, two sets of quadruplicate 60-mL shake flasks containing CHO-16F cells were grown for 48 hours at 37°C. Half of the flasks were then shifted to 31°C for 24 hours, keeping the other four at 37°C as controls.

Shifting temperature to 31°C slowed down cell growth, resulting in a cell density 24% lower in the cultures shifted to 31°C in comparison with cells grown at 37°C (Figure 30).

Total RNA was extracted using the mirVana™ miRNA Isolation kit, with phenol (Thermo Fisher) and rRNA depleted using the Ribo-Zero Gold rRNA Removal Kit. Library preparation and sequencing were performed as described in Tzani *et al.* (Tzani *et al.*, 2020).

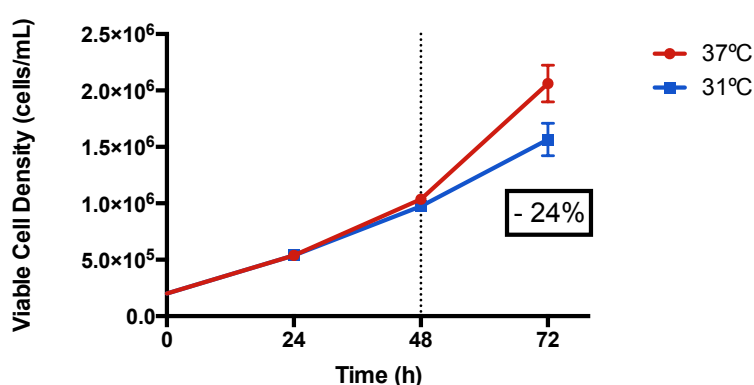


Figure 30 –Viable cell density over time for the CHO-16F cultures. Cells were grown for 48h at 37°C. At this point, half of the cultures were shifted to 31°C for 24h. Cells were harvested after 72h total culture time for RNA extraction (N=4)

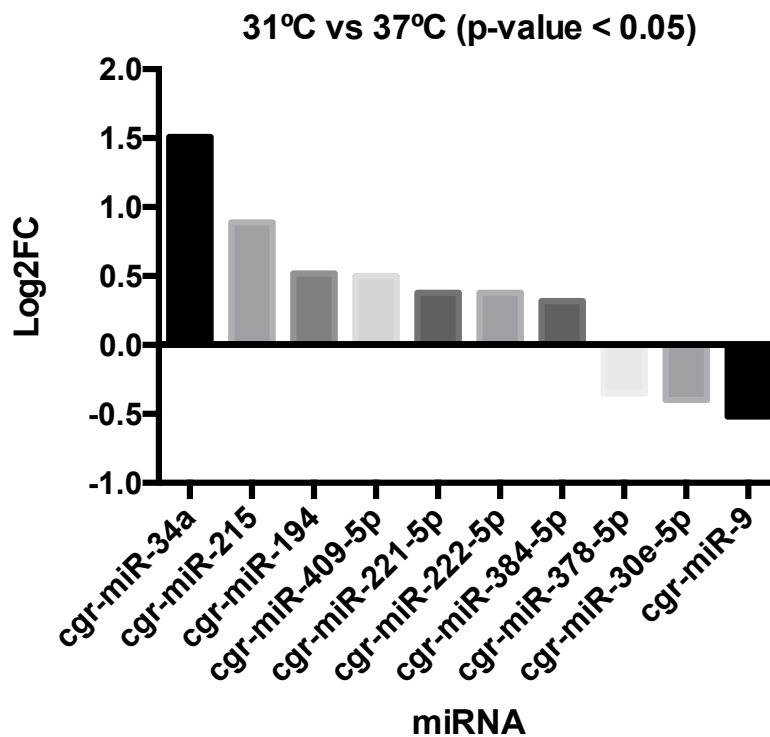


Figure 31 – Relative expression levels (Log2FC) of microRNA candidates selected from the RNA-Seq data, based on differential expression between 37°C and 31°C (p-value < 0.05).

After performing the data quality analysis, and the differential expression analysis using R Bioconductor/DESeq2 software, 10 microRNA showing differential expression between 37°C and 31°C were found (Figure 31).

2.7. Independent temperature-shift experiments to validate the identified temperature sensitive miRNAs

In order to validate the identified microRNAs, the original temperature shift experiment was replicated. Cells were seeded at 2×10^5 cells/mL and grown for 48h at 37°C in shake flasks with volumes ranging from 25-50 mL in SFM II + 5% PVA media. Half of the cultures were then shifted at 31°C for 24h. At this point, cell pellets of ca. $1-5 \times 10^6$ cells were sampled, spun down and re-suspended in Trizol® reagent for RNA extraction. Figure 32 shows final cell densities obtained for 3 independent temperature shift experiments at 37°C and 31°C. A reduction of cell density of 27-33% was observed after 24h at reduced temperature.

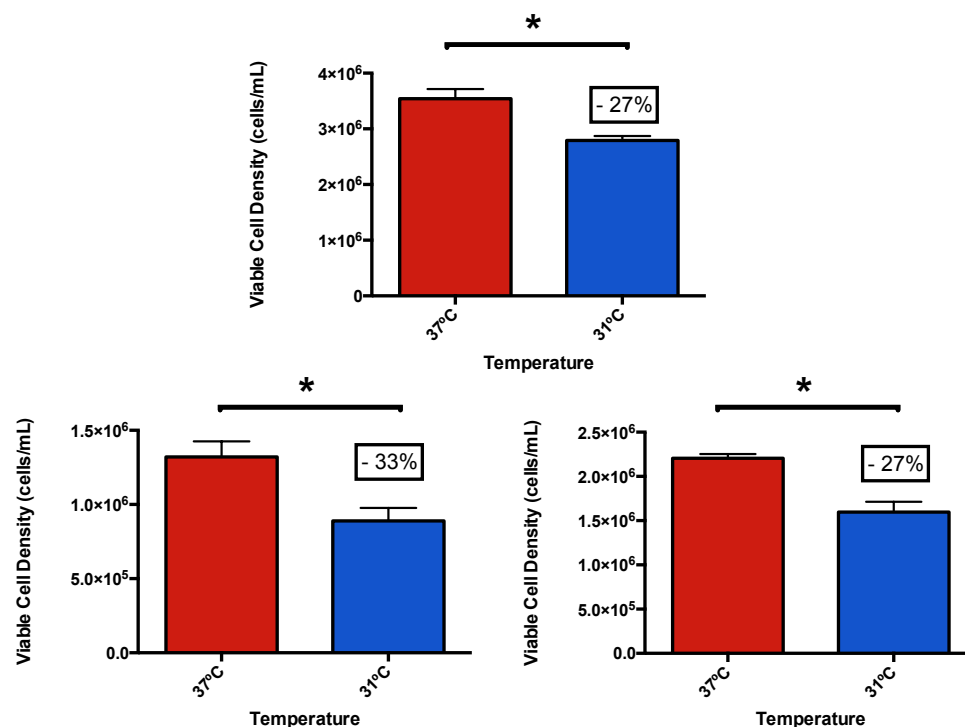


Figure 32 – Cell density measured 24h after the temperature shift during the temperature shift of CHO-16F cultures. Cells were grown in 20-mL SFMII media in 250 mL shake flasks for 48 hours at 37°C, then half of the cultures were shifted to 31°C for 24h. Cell density was measured at this point and samples ($1-5 \times 10^6$ cells) were taken for RNA extraction. Each bar represents the average of 3 technical replicates and error bars represent the standard deviation. Each plot represents 1 biological replicate (N=3). Differences for each experiment were evaluated using a two-tailed homoscedastic Student T-test. * = p-value < 0.05.

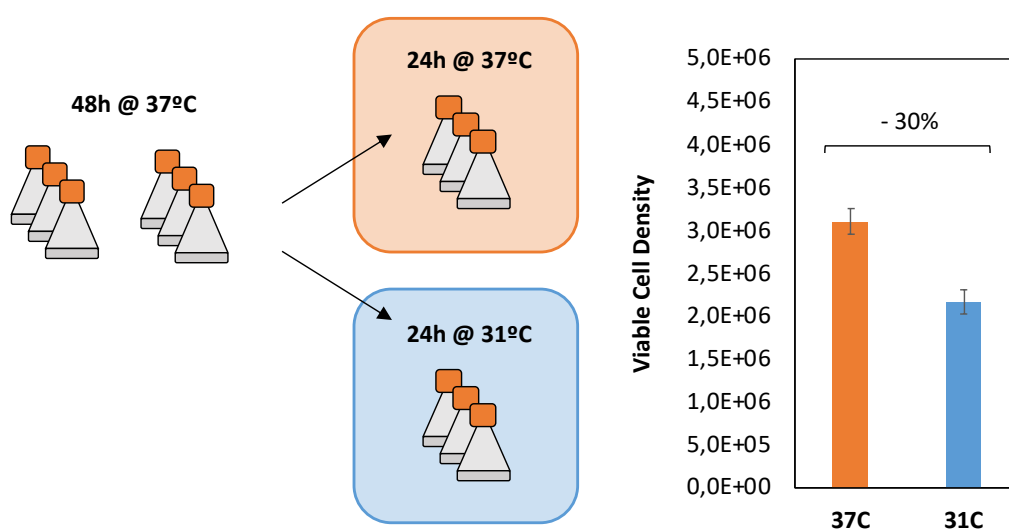


Figure 33 – (Left) Experimental design and (right) cell density measured 24h after the temperature shift during the temperature shift of CHO-K1-16F cultures. Cells were grown in 20 mL SFMII media in 250 mL shake flasks for 48 hours at 37°C, then half of the cultures were shifted to 31°C for 24h. Triplicate culture per conditions were used (N=3). Cell density was measured at this point and samples (5×10^6 cells) were taken for total RNA extraction.

After confirming a robust growth-reduced phenotype in the previous experiment, a final temperature shift experiment consisting of triplicate shake flasks with 30 mL following the same protocol was performed. Samples for total RNA extractions were collected and total RNA extracted and stored at – 80°C for analysis.

2.8. Validation of identified temperature sensitive miRNAs by RT-qPCR

Differential miRNA expression after temperature shift for a subset of targets was analysed using TaqMan microRNA assays as per Methods Section 1.3.3. From the 10 miRNA that were identified as differentially expressed between 37°C and 31°C, 8 were selected after applying a Mean Normalised Count cut-off higher than 100. U6 snRNA was used as endogenous control. cgr-miRNA-204, which was used as model miRNA in previous chapters, and was not found to change upon temperature shift was also measured for reference. Ct values and relative quantification between samples at 37°C and samples at 31°C are shown in Figure 34 and Figure 35. For each miRNA, abundance data, estimated using the mean of read counts (BaseMean) and linear fold-change (lfc) values obtained from the RNASeq data analysis are also shown for comparison.

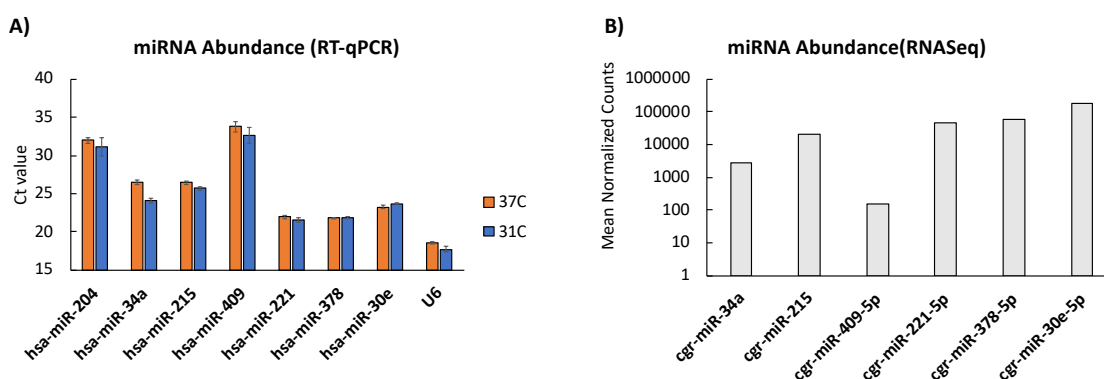


Figure 34 – Abundance (Ct values) of the identified temperature sensitive miRNA. miRNA expression was analysed by qPCR (A). Mean normalized counts from the RNASeq experiment are shown for comparison (B). Bars in plot A) represent the average Ct value for 3 biological replicates (N=3) and error bars represent the standard deviation. In Plot B) each bar represents the BaseMean counts for each miRNA as obtained after sequencing and analysis.

Regarding miRNA abundance, the Ct values obtained by RT-qPCR data matched well the RNASeq counts data. miRNA which showed higher abundance such as cgr-miR-221, cgr-miR-378 and miR-30e on the RNASeq showed the lowest Ct values at around 21-23. miRNAs with moderate abundance, cgr-miR-34a and cgr-miR-215, showed Ct values of

24-26 and low expressed miRNAs, such as cgr-miR-409-5p and cgr-miR-204 showing higher Ct values above 30 (Figure 34).

In terms of relative quantification, only 3 out of the 6 tested candidates, cgr-miR-34a, cgr-miR-378 and cgr-miR30a showed fold changes validating the results obtained in the RNASeq experiment. It was possible to confirm that cgr-miR-34a expression is up-regulated when shifting cultures from 37°C to 31°C whereas cgr-miR-378 and cgr-miR-30a are down-regulated with the decrease in temperature (Figure 35).

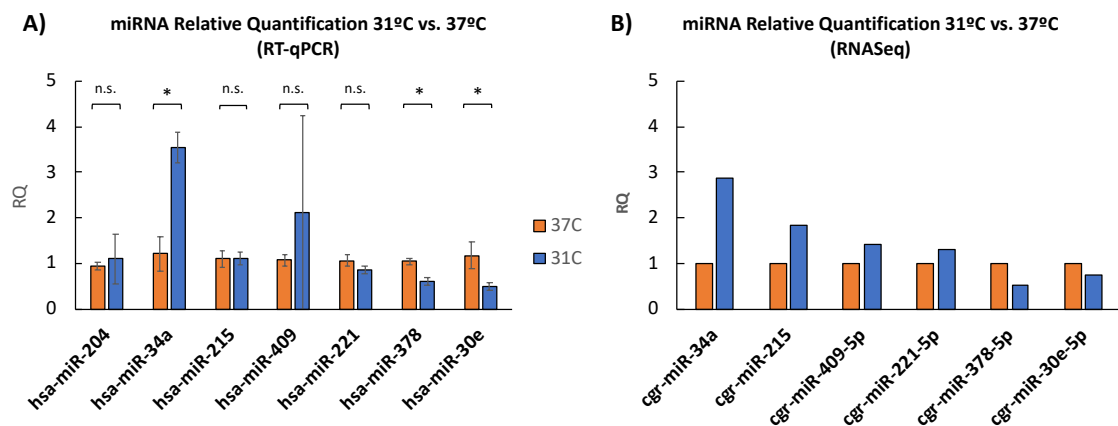


Figure 35 – Relative quantification for the analysed temperature sensitive miRNAs analysed by RT-qPCR using TaqMan assays (A) Relative quantification of the same miRNAs calculated from the RNASeq experiment (B). For the qPCR data, each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. RT-qPCR results were statistically tested using a two-tailed homoscedastic Student t-test - * = p -value < 0.05, n.s. = not significant.

Having verified their differential expression after temperature shift, cgr-miR-34a, cgr-miR-30e-5p, cgr-miR-378-5p were moved forward as candidates to be tested as temperature sensitive triggers to modulate gene expression using specific miRNA sponges. Before doing that another element in the expression vector that could potentially be affected by temperature, the promoter, was investigated.

2.9. Reducing promoter-related variation upon temperature shift to allow evaluation of miRNAs for transgene expression control

Before cloning specific sponge sensors for the validated temperature sensitive miRNA, the effect of shifting temperature from 37°C to 31°C on the expression of GFP from different promoters was assessed. The aim was to select a promoter that would allow minimal changes of GFP expression after temperature shift as a starting point to build

an expression cassette containing miRNA responsive elements and be able to evaluate them under temperature shift conditions with minimum noise.

Three different promoters, CMV, UBC and EF1a, were tested in GFP expressing plasmids. Cells were transfected in triplicate wells in 24-well suspension plates with each specific plasmid and GFP was evaluated 24h after transfection by flow cytometry. In order to minimize differences in GFP expression derived from differences in transfection efficiency, it was important to ensure that the same cell population was used to seed cultures to be used for temperature shift. This was achieved by pooling the different transfection replicates and re-measuring GFP expression in that population. As temperature shift experiments could only be performed in a static incubator, the pooled cells were re-seeded in 6-well cell culture treated plates at 3×10^5 cells/well in 2 mL DMEM:F12 media + 5% FBS. Cells were cultured for 24h at 37°C. At this point, half of the cultures were shifted to 31°C. The other half was left at 37°C as a control. Samples were taken at 24h after the temperature shift and GFP fluorescence, cell density and viability were measured.

As previously observed, temperature shift had an effect on cell growth. Cells treated with the only Mirus® control and untransfected cells grew to higher densities at 37°C than at 31°C (Figure 36 – A). In terms of GFP expression (Figure 36 – B), mean fluorescence intensity, either of the GFP positive fraction (MFI) or the whole population (MFI all events) was higher in cells shifted to 31°C degrees than in cells maintained at 37°C. Differences in the effect of temperature shift on GFP expression were observed depending on the promoter used. Whereas expression from the CMV GFP plasmid was increased by 43% upon temperature shift, expression from the EF1a GFP plasmid was increased by only 10% and expression from the UBC GFP plasmid was increased by 21%. The increase in GFP expression after temperature reduction was also apparent when looking at the percentage of GFP positive cells, with 61% and 40% more GFP positive cells at 31°C for the CMV-driven and the UBC-driven GFP plasmids respectively, while a difference of only 3% was observed for cells expressing GFP from an EF1a promoter. Differences can also be seen when looking at the MFI for all recorded events and in this case, the temperature shift effect on UBC-driven GFP cells was the strongest, with +108% more expression at 31°C, followed by CMV-driven GFP plasmid, with 64% and a minimal increase of 7% was observed for EF1a-driven GFP. To account for changes in

cell numbers, caused by high levels of expression of GFP potentially affecting growth as well as from differences caused by plasmid dilution in faster growing cells, normalised MFI (all events) / VCD values were calculated.

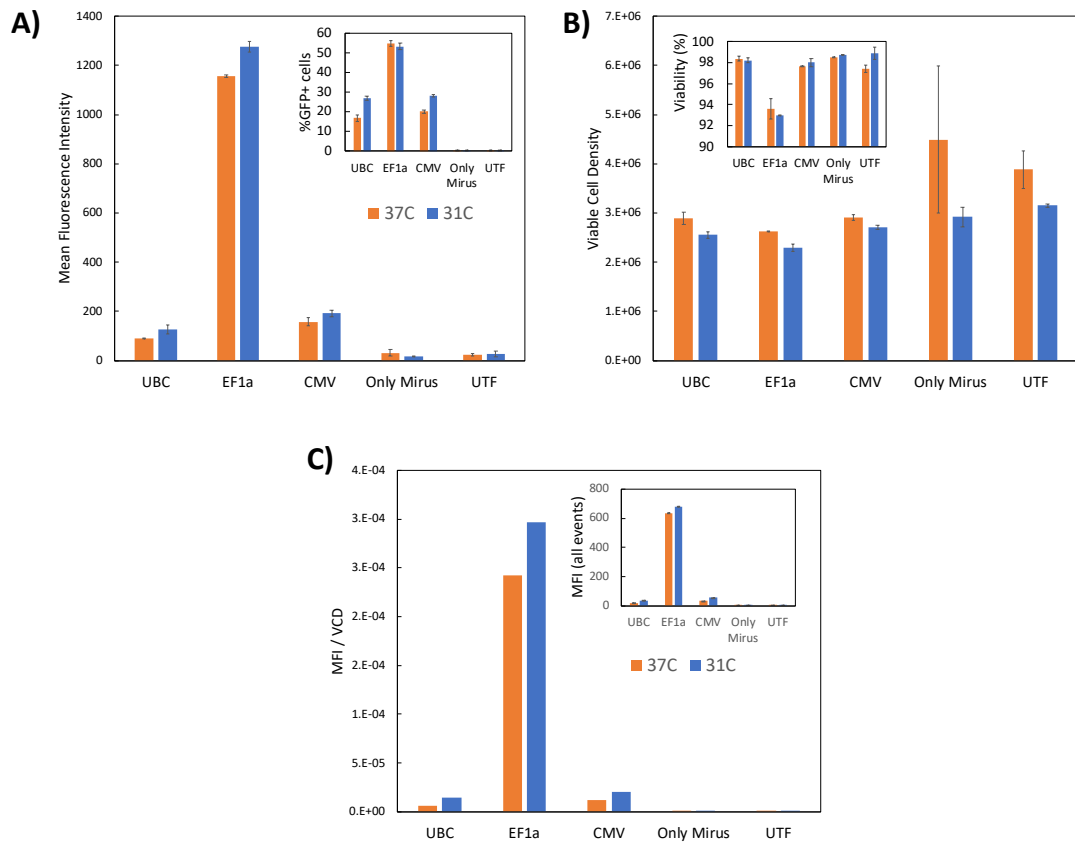


Figure 36 – A) GFP expression (Mean Fluorescence Intensity in large plot and %GFP cells in small plot); B) Cell density (Viable Cell Density in large plot and %Viability in small plot); C) Normalised MFI all events/VCD. CHO-K1 cells were transfected with a plasmid expressing GFP under different promoters: UBC, EF1alpha and CMV and GFP was measured 24h after temperature shift from 37°C to 31°C. Only Mirus correspond to samples treated with only the transfection reagent. UTF = untransfected cells. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

As before, temperature shift to 31°C led to an increase of the MFI/cell of 135% for the UBC-driven GFP plasmid, 22% for the EF1a-driven GFP and 75% for the CMV-driven GFP plasmid (Figure 36 and APPENDIX B – SUPPLEMENTARY MATERIAL FOR CHAPTER 4 Supplementary Table 1).

Based on the above results, EF1a was chosen as the promoter to use for the construction of miRNA sponge expression plasmids due its lower reactivity to temperature in comparison with CMV or UBC.

2.10. CMV and EF1a as drivers for miRNA sponge sensors

In order to further characterize the EF1a promoter and evaluate its less responsive nature upon temperature shift, a backbone GFP vector containing an EF1a promoter driving the expression of the GFP miRNA-sponge cassette was constructed by replacing the original CMV promoter with an EF1a promoter. Both vectors, the original CMV-GFP and the newly constructed EF1a-GFP, were transfected in CHO-K1 cells and GFP expression after temperature shift was evaluated as described in Section 2.9. Data corresponding to viable cell density, viability, Mean Fluorescence Intensity (MFI) and percentage of GFP positive cells (%GFP+) is available in Supplementary Table 2 – Appendix B.

As observed previously, for all conditions, shifting cells to 31°C had an effect on cell growth and lower cell densities were observed after 24h and 48h at 31°C when compared with control cells grown at 37°C. No differences in viability were observed between cells at 37°C and 31°C (Figure 37 – top and middle panels). Regarding GFP expression as measured by the mean fluorescence intensity of the transfected population and by the percentage of GFP positive cells (Figure 37 – middle and bottom panels), no differences in GFP expression between the two tested plasmids were observed. In terms of the effect of temperature shift on GFP expression (Table 11), cells grown at 31°C showed higher MFIs and %GFP positive cells than cultures grown at 37°C, for both the CMV-driven plasmid (CMV-GFP) and the EF1a-driven GFP plasmid (EF1a-GFP), at both time points, 24h and 48h following temperature shift (Table 11). In terms of MFI, once more, CMV-driven GFP expression was increased by 56% between 37°C and 31°C whereas EF1a-driven GFP expression was increased by 34% after temperature reduction.

Based on these results, the EF1a-GFP backbone plasmid was selected as the vector to clone the different miRNA-responsive elements / sponges for the temperature sensitive miRNAs.

This experiment was also used as a reference to design an appropriate protocol to be able to evaluate temperature shift effect on miRNA sponges. It is worth noting that at 48h post temperature shift, both cells growing at 37°C and cells growing at 31°C had reached 100% confluency under the microscope. This was taken into account in

subsequent experiments as growth inhibition by contact or cells floating in the supernatant being lost during trypsinization was to be avoided.

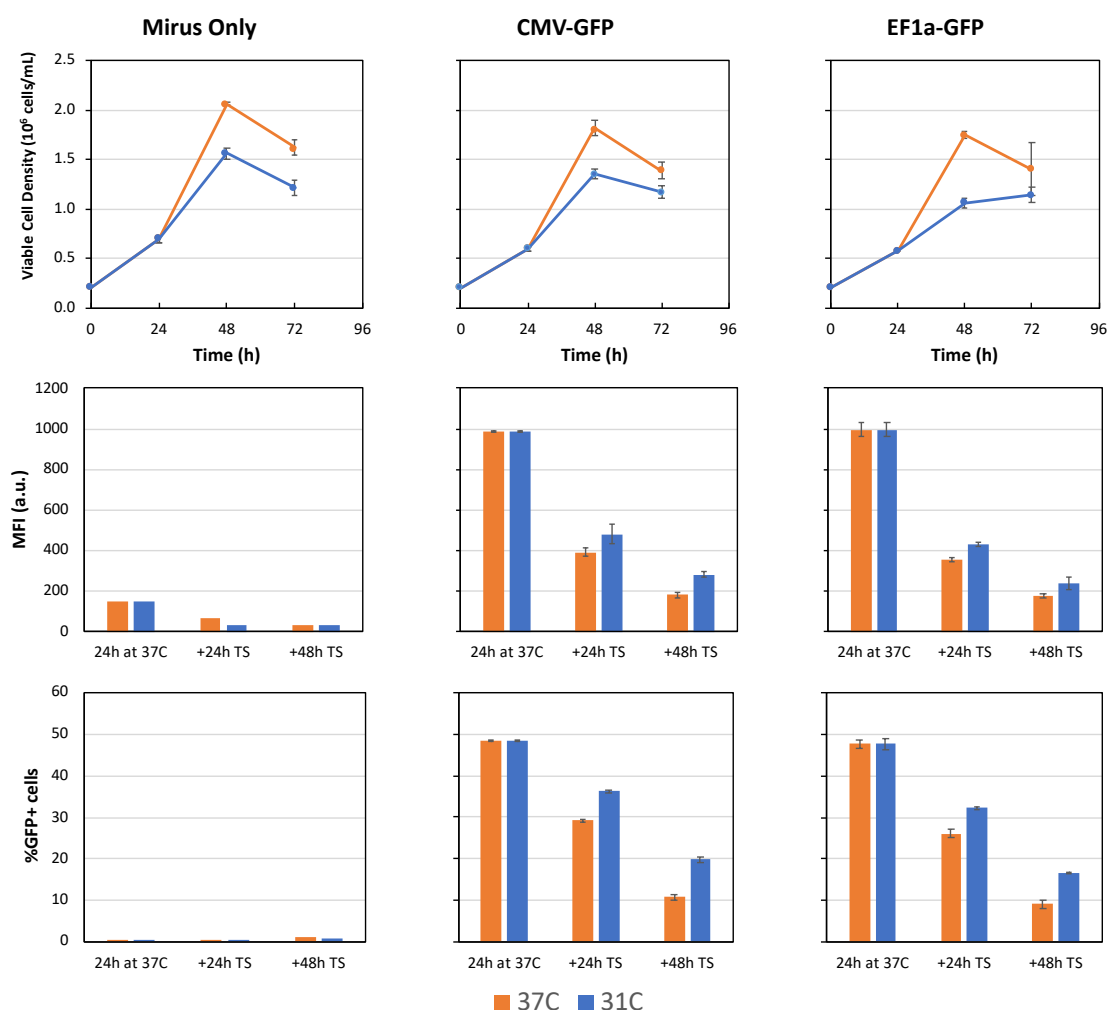


Figure 37 – Comparison of cell growth and GFP expression between cells transfected with a CMV-GFP plasmid and cells transfected with an EF1a-GFP plasmid after temperature shift. Viable cell density (top), Mean Fluorescence Intensity (middle) and percentage of GFP positive cells (bottom) are shown. Transfected cells were re-seeded in 6-well plates cultures at time 0h and cultures were grown at 37°C for 24h before temperature shift (TS) was performed. Control cultures were kept at 37°C (orange) and TS cultures were moved 31°C (blue) for 48h. A Mirus™ only (MO) mock transfection was grown in parallel as negative control for flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

Table 11 – Effect of temperature shift on the different measured parameters (VCD, Viability, MFI and %GFP+) for cells transfected with each different plasmid. Ratios (%) for between 31°C and 37°C calculated based on data provided in Supplementary Table 2

Viable Cell Density (VCD)		Time	
Temperature	Plasmid	+24h TS	+48h TS
31C / 37C	MO	-24%	-25%
31C / 37C	CMV-GFP	-26%	-16%
31C / 37C	EF1a-GFP	-39%	-19%
Viability (%)		Time	
Temperature	Plasmid	+24h TS	+48h TS
31C / 37C	MO	0%	-1%
31C / 37C	CMV-GFP	-1%	1%
31C / 37C	EF1a-GFP	-1%	0%
MFI (Y-Mean)			
Temperature	Plasmid	+24h TS	+48h TS
31C / 37C	MO		
31C / 37C	CMV-GFP	23%	56%
31C / 37C	EF1a-GFP	20%	34%
%GFP+ cells		Time	
Temperature	Plasmid	+24h TS	+48h TS
31C / 37C	MO		
31C / 37C	CMV-GFP	24%	84%
31C / 37C	EF1a-GFP	23%	82%

2.11. EF1a-driven miRNA sponges for temperature sensitive mRNAs

Using the EF1a-GFP backbone described above – miRNA sponges specific for a subset from validated temperature sensitive miRNAs were constructed as per Methods Section 1.5.5. Hybrid sponges containing binding sites for cgr-miR-30e-5p and cgr-miR-378-3p were also constructed.

As previously described, negative control (NC) sponges were cloned using a non-miR binding sequence. Sequences for these different binding sites are shown in Table 8. Based on the work presented in Chapter 4, fully complementary binding sites were used. Whenever possible, sponges containing 2 and 8 binding sites for each targeted sequence were selected. For hsa-miR-34a-3p, no sponge plasmids with 8 binding sites were obtaining during colony screening and therefore one with 6 binding sites was selected.

For the mixed sponges, different combinations of the two different binding sites were selected with sponges ranging from 6 to 14 total binding sites.

Table 12 - Oligo sequences used to construct miRNA sponges for the selected temperature sensitive miRNAs . A negative control sponge was constructing using A-rich sequence (NC-sponge) as in Chapter 4. The overhanging ends for cloning using KflI (SanDI) and spacers are marked in bold. miRNA binding regions are underlined.

	Sequence
cgr-miR-34a-5p Forward Reverse	GTCCC <u>ACAACCAGCTAAGACACTGCCA</u> AATT <u>ACAACCAGCTAAGACACTGCCAGG</u> GACCC TGGCAGTGTCTTAGCTGGTTGT AATT TGGCAGTGTCTTAGCTGGTTGT GG
cgr-miR-221-3p Forward Reverse	GTCCC <u>GAAACCCAGCAGACAATGTAGCT</u> AATT <u>GAAACCCAGCAGACAATGTAGCT</u> GG GACCC CAGCTACATTGTCTGCTGGGTTT CAATT AGCTACATTGTCTGCTGGGTTT C GG
hsa-miR-30e-5p Forward Reverse	GTCCC <u>GCTTCCAGTCGAGGATGTTTACA</u> AATT <u>GCTTCCAGTCGAGGATGTTTACA</u> GG GACCC TGTAAACATCCTCGACTGGAAG CAATT TGTAAACATCCTCGACTGGAAG C GG
cgr-miR-378-3p Forward Reverse	GTCCC <u>GCCTTCTGACTCCAAGTCCAGT</u> AATT <u>GCCTTCTGACTCCAAGTCCAGT</u> GG GACCC ACTGGACTTGGAGTCAGAAG GCAATT ACTGGACTTGGAGTCAGAAG G C GG

Table 13 – Number of inserts and binding sites for the selected miRNA-sponges for each temperature sensitive miRNA.

Sponge	Binding sites	Details	
NC-spg-2	2		
NC-spg-8	8		
34a-spg-2	2		
34a-spg-8	6		
30e-spg-2	2		
30e-spg-8	8		
221-spg-2	2		
221-spg-8	8		
378-spg-2	2		
378-spg-8	8		
378/30e-spg-1	6	1x 378	2x 30a
378/30e-spg-2	10	1x 378	4x 30a
378/30e-spg-3	14	3x 378	4x 30a
378/30e-spg-4	8	3x 378	1x 30a

The next step was to optimise the experimental setup to be able to perform transient transfections and temperature shift experiments using the newly constructed plasmids. Building on the work detailed in section 2.10, plasmids NC-spg-8 and 378-spg-8 were used as test sponges in the following optimisation experiments.

2.12. Optimisation of EF1a-driven miRNA sponge expression under temperature shift to assess miRNA-driven changes in GFP expression

In order to test whether temperature driven miRNA expression changes could be detected by measuring changes in the expression of the GFP reporter linked to the miRNA sponge, the NC-sponge and a miR-378-sponge with 8 binding sites (NC-spg-8 and 378-spg-8) were used in a series of optimisation experiments with the aim to optimize the transfection method, and the temperature shift strategy to identify miRNA-driven changes in transgene expression (GFP). In short, CHO-K1 cells were transfected in 24 well suspension plates. Four wells were transfected per plasmid including an only Mirus® negative control and the “no- sponge” backbone vector EF1a-GFP as positive control. 24h after transfection, individual wells were measured to confirm transfection efficiency and GFP expression. As previously described, cells were pooled to obtain homogenous populations to seed the temperature shift experiment. Pools were also measured for GFP expression to ensure homogeneity. A small decrease in GFP expression from EF1a-NC-sponge was observed when compared with the EF1a-GFP control vector, both in the percentage of GFP positive cells and the MFI for all recorded events.

Six-well cell culture plates were seeded with 1×10^6 cells/well in SFM II media + 5% FBS from the transfected pools – this high density was chosen to saturate the well and start from time 0 with high cell numbers, in an attempt to minimize differences in cell growth between the two culture temperatures, based on the observation regarding cell density made in 2.6. Half of these plates were incubated at 37°C and the other half at 31°C from time 0 (seeding time). Cells were monitored every 24h under a bright field microscope to assess the degree of confluency of the wells and the eventual need to re-feed with media to ensure cell viability.

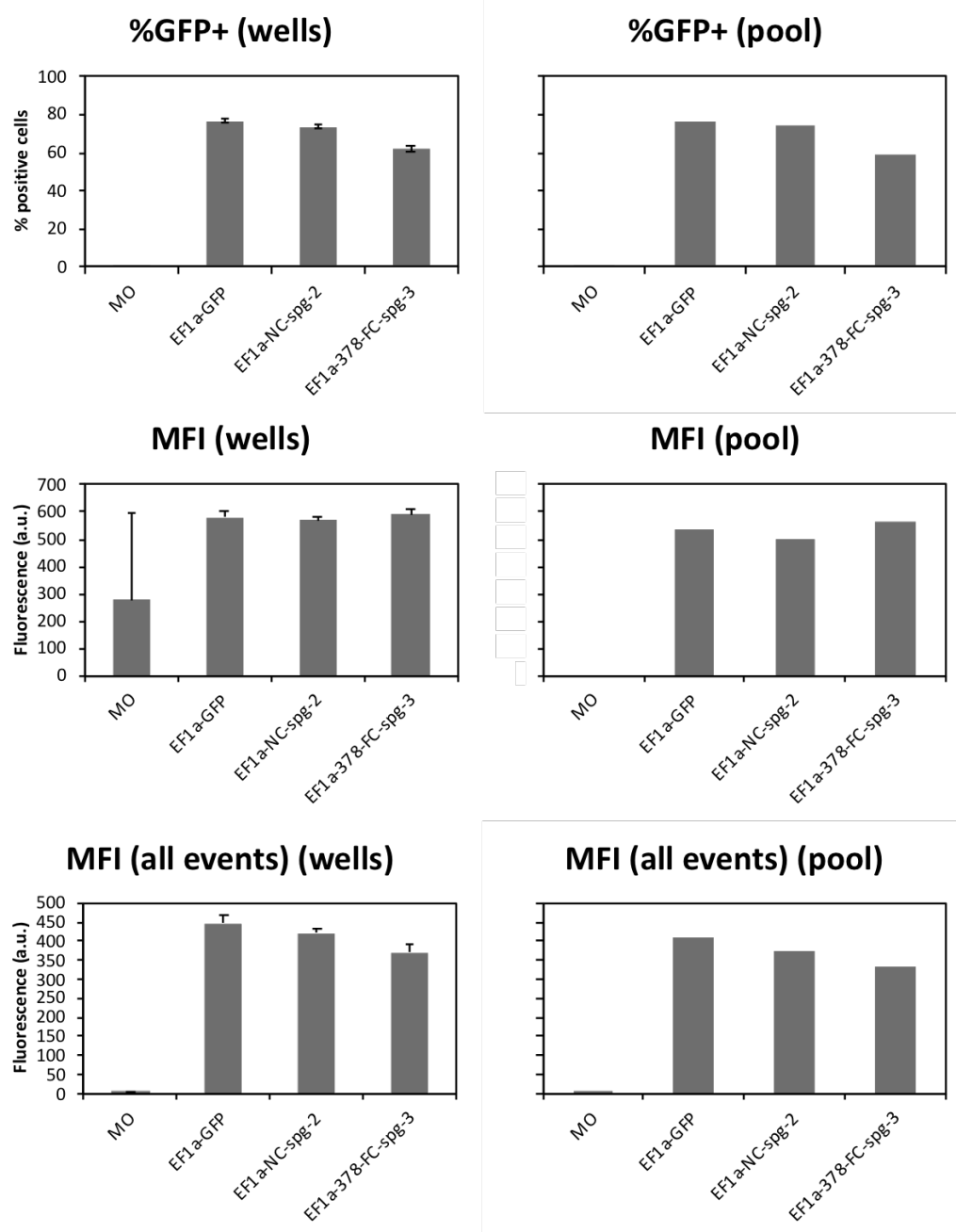


Figure 38 – GFP expression in transient populations of cells transfected with EF1a-GFP vector, EF1a-NC-sponge and EF1a-378-sponge, measured 24h after transfection before seeding the temperature shift experiment. %GFP positive cells (top), MFI (Y-Mean) for the GFP+ sub-population (mid), and MFI (Y-Mean) for all measured event (bottom) are shown. Measurement of the 4 replicate wells are shown in the left panel and measurements of the pooled populations are shown in the right panel. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation (only one measurement per pool).

A larger decrease (17%) was observed for the EF1a-378-sponge vector when compared with the control, presumably due to the effect of miR-378 binding the transcript (Figure 38).

At 24h post seeding, a lot of floating cells were observed in the supernatant, with adherent cells covering 100% of the surface at both temperatures. Floating cells were sampled for flow cytometry measurements (GFP and Viability) and media was replaced. At 48h after seeding, many floating cells were observed at both temperatures. Supernatant was sampled again for flow cytometry measurements of GFP fluorescence and viability. Attached cells were then trypsinized and measured for GFP fluorescence and cell viability using flow cytometry. Data for cell density, viability and GFP expression is shown in Figure 39 and Figure 40.

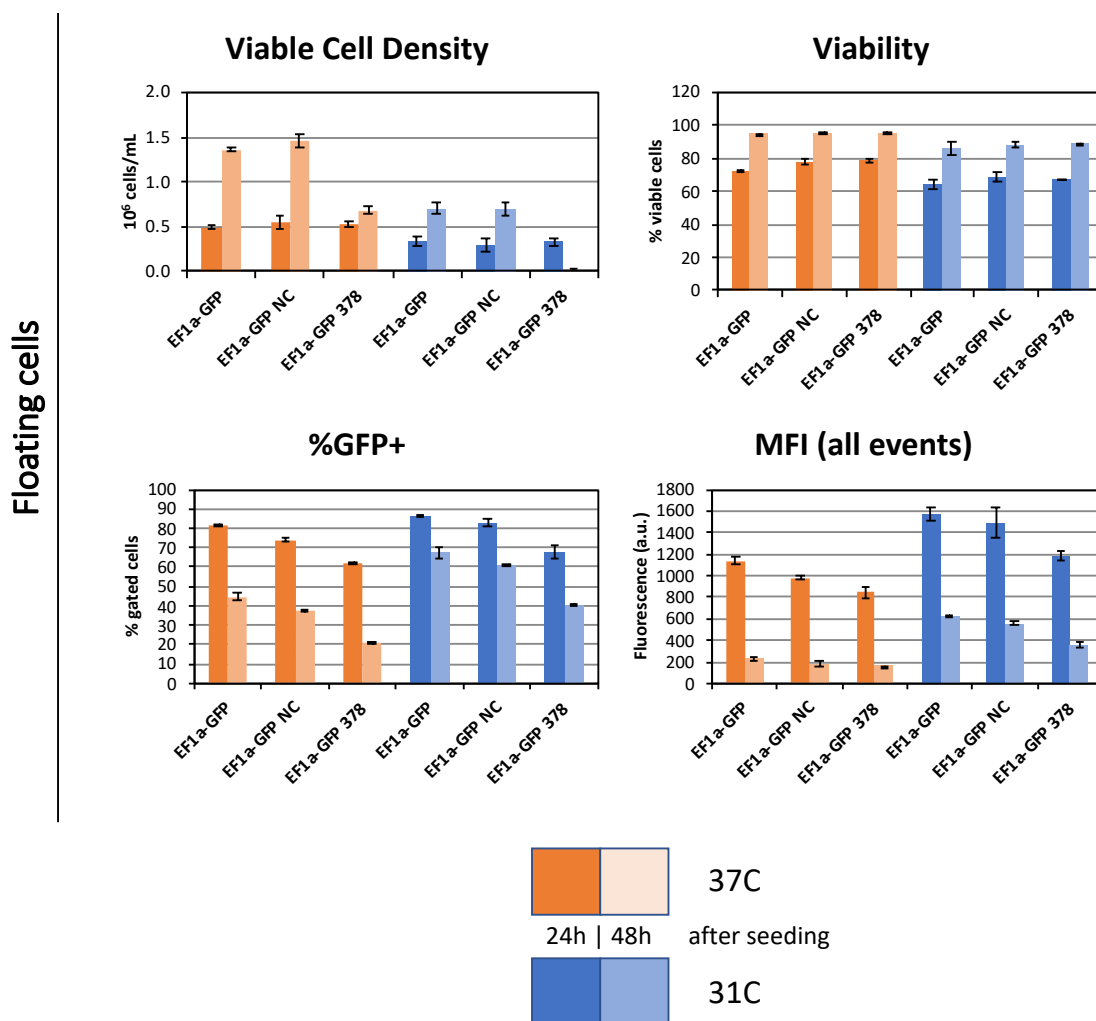


Figure 39 – Viable cell density, viability, and GFP expression (%GFP positive cells and MFI (all events) for the different sampled floating cells after 24h (dark colours) and 48h (light colours). Cells were seeded in 6-well cell culture plates after 24h from being transfected with EF1a-GFP vector, EF1a-NC-sponge and EF1a-378-sponge. Floating cells were sampled from the supernatant at different time points. Measurements correspond to the average of 2 technical replicate wells and error bars represent \pm standard deviation.

Large differences in viable cell density were observed between the 37°C and the 31°C conditions when looking at viable floating cells (Figure 39 – VCD), both at 24h and at 48h after temperature shift. Cells at 31°C grew up to 50% less than cells growing at 37°C and remained floating in the supernatant not having any surface available. Looking at adherent cells, smaller differences were observed in cell density between 37°C and 31°C, but small differences between cells transfected with different plasmids and grown at the same temperature were observed (Figure 40 – VCD).

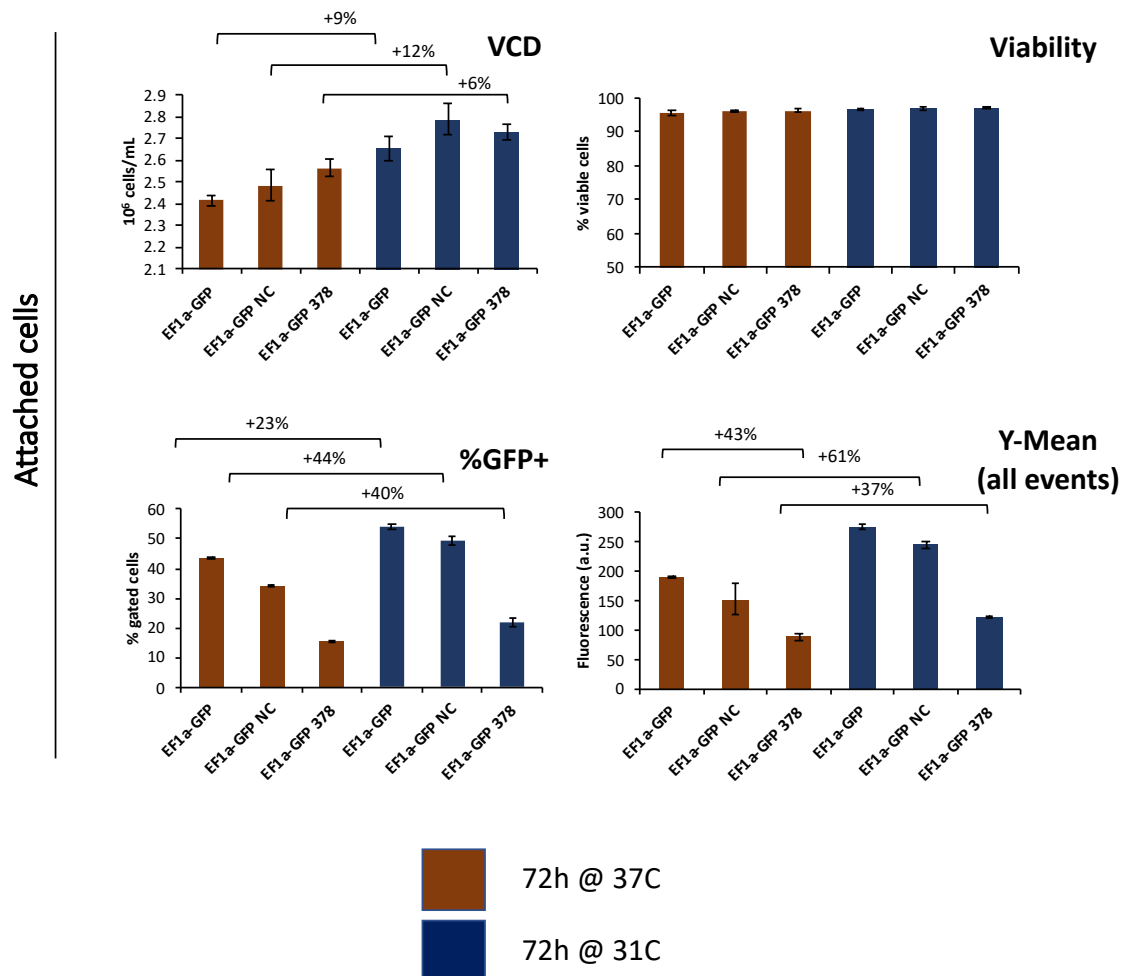


Figure 40 – Viable cell density, viability, GFP expression (%GFP positive cells and MFI all events) for transient pools 72h after temperature shift. Cells were seeded in 6-well cell culture plates 24h after being transfected with EF1a-GFP vector, EF1a-NC-sponge and EF1a-378-sponge. Plates were incubated at 37°C or 31°C for additional 72h with media being replaced every 24h. Measurements correspond to the average of 2 technical replicate wells and error bars represent \pm standard deviation.

Overall, viability for floating cells was low, between 60%-80% for the first 24h at both temperatures, but recovered for both temperatures by 48h (Figure 39 – Viability). GFP expression was higher for the three plasmids at 31°C, as observed on the % GFP cells

and MFI (Y-Mean – all events) both in floating and attached cells (Figure 39 and Figure 40 -%GFP and Y-Mean – all events).

In order to be able to compare differences in GFP expression between NC-sponges transfected cells and 378-FC-sponge transfected cells, it was considered that the transient expression model used required several normalisation steps in order to compensate for differences in cell growth derived from both different growth rates at each temperature as well as differences in GFP expression levels that can affect growth.

Different growth rates would lead to different plasmid dilution rates in the different cell populations, as faster growing cells would dilute their plasmid content quicker, thus having less template from which to express protein. To factor this in, a normalisation step using the GFP signal from the control vector EF1a-GFP at each temperature condition could be used with data being normalised according to the formula below:

$$\text{GFP Expression (Norm.EF1a)} = \text{NC or 378-FC GFP Expression} / \text{EF1a-GFP Expression}$$

However, this normalisation using the control vector signal, would not take into account the different levels of GFP expression and the potential effect these may have on growth. As previously observed, different levels of GFP levels can affect cell growth, with cells expressing more GFP potentially growing slower than cells transfected with more repressed transcripts (e.g. NC-sponges) This might not be very apparent in this optimisation experiment but later when using sponges for high expressed miRNAs such as 30e-spgs or 378/30e-spgs. To account for that, a normalisation using the Viable Cell Density (VCD) measurement for each sample at each temperature condition using the below formula could be more appropriate.

$$\text{GFP Expression (Norm. VCD)} = \text{NC or 378-FC GFP Expression} * \text{VCD}$$

Normalized GFP expression data using the two different approaches is shown in Table 14 and Figure 41.

Table 14 - Normalized values for GFP expression using the EF1a-GFP fluorescent signal at each temperature or the VCD measured for each sample. The resulting data is corrected for differences resulting from different plasmid dilution rates between 37°C and 31°C and difference in cell growth between the two conditions.

GFP Expression (Norm.EF1a)

		24h Floating		48h Floating		48h Attached	
		Average	SD	Average	SD	Average	SD
37C	EF1a-GFP	1.00	0.02	1.00	0.08	1.00	0.01
	EF1a-GFP NC	0.86	0.01	0.78	0.11	0.79	0.14
	EF1a-GFP 378	0.74	0.05	0.68	0.01	0.47	0.03
31C	EF1a-GFP	1.00	0.04	1.00	0.01	1.00	0.02
	EF1a-GFP NC	0.95	0.09	0.89	0.02	0.89	0.02
	EF1a-GFP 378	0.76	0.03	0.57	0.03	0.44	0.01

GFP Expression (Norm. VCD)

		24h Floating		48h Floating		48h Attached	
		Average	SD	Average	SD	Average	SD
37C	EF1a-GFP	1.00	0.02	1.00	0.08	1.00	0.01
	EF1a-GFP NC	0.95	0.01	0.78	0.11	0.82	0.14
	EF1a-GFP 378	0.80	0.05	0.73	0.01	0.50	0.03
31C	EF1a-GFP	1.00	0.04	1.00	0.01	1.00	0.02
	EF1a-GFP NC	0.85	0.08	0.92	0.02	0.93	0.02
	EF1a-GFP 378	0.75	0.03	0.58	0.03	0.46	0.01

GFP expression from NC-sponges and 378-sponge between 37°C and 31°C was compared using the above-mentioned normalised data. It must be noted that no major differences were observed between the two proposed normalisation methods (Figure 41). In both cases, differences in GFP expression between NC-sponge constructs and the 378 constructs at each temperature were different depending on the analysed population in the 3 analysed populations (floating cells at 24h and 48h and attached cells at 48h). While in floating cells, a larger 378-sponge-dependent reduction in GFP expression was observed at 31°C compared to 37°C, smaller differences were observed in the attached cell populations between 37°C and 31°C.

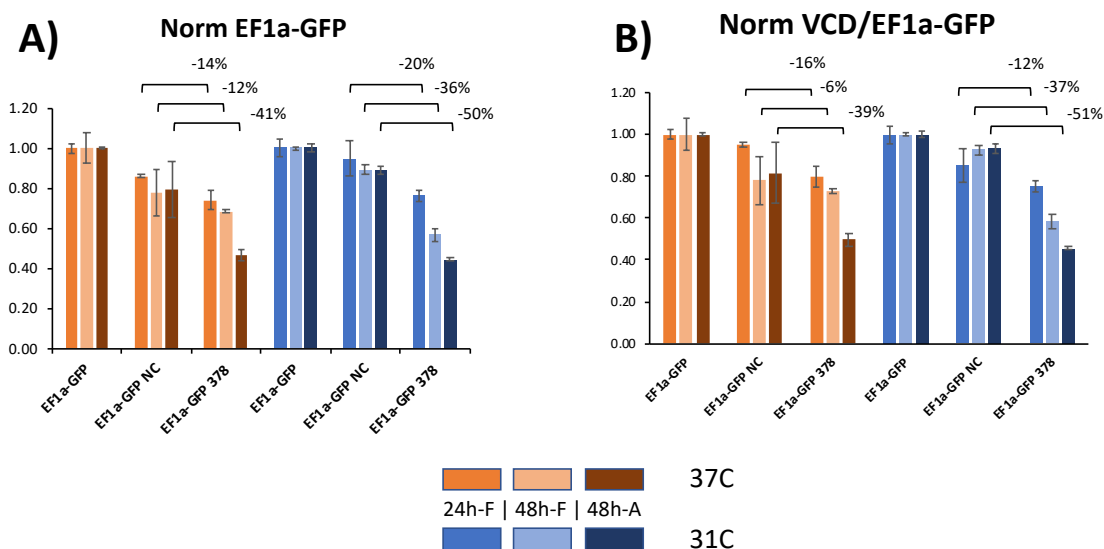


Figure 41 - Normalized values for GFP expression using EF1a-GFPs signal (A) or VCD (B). Data is always shown in relation with the EF1a-GFP vector for each temperature. Shown values correspond to 2 technical replicates transfections and error bars represent \pm normalized standard deviations.

In order to solve the problem resulting from having floating cell in the cultures, a different set of 6 well culture plates was seeded at a lower cell density of 1×10^5 cells/well – this low density was chosen with the aim to not reach 100% confluence over the 72h of the experiment (24h at 37°C then 48h at 37°C/31°C).

In terms of temperature shift conditions, two different combinations of 37°C and 31°C incubation times were investigated:

- 72h at 31°C (directly after seeding the well, without any attachment period at 37°C)
- 24h at 37°C then 48h at 31°C
- A set of plates at 37°C for the 72h was included as control

No floating cells were observed over the course of the 72h in this experiment, so cells were only sampled by trypsinization as described above. In this case, cells only reached 90-100% confluency in the 37°C plates after 72h.

As shown in Figure 42 - VCD, for both temperature shift setups, cells at 31°C reached lower viable cell densities than cells at 37°C. A reduction of 17-20% in final cell density was observed for cells grown for 24h at 37°C and then shifted to 31°C for 48h, while a reduction of 43-50% in final cell density was observed for cells grown at 31°C for 72h. Viability was over 95% for all conditions. In terms of GFP expression, cells shifted to 31°C

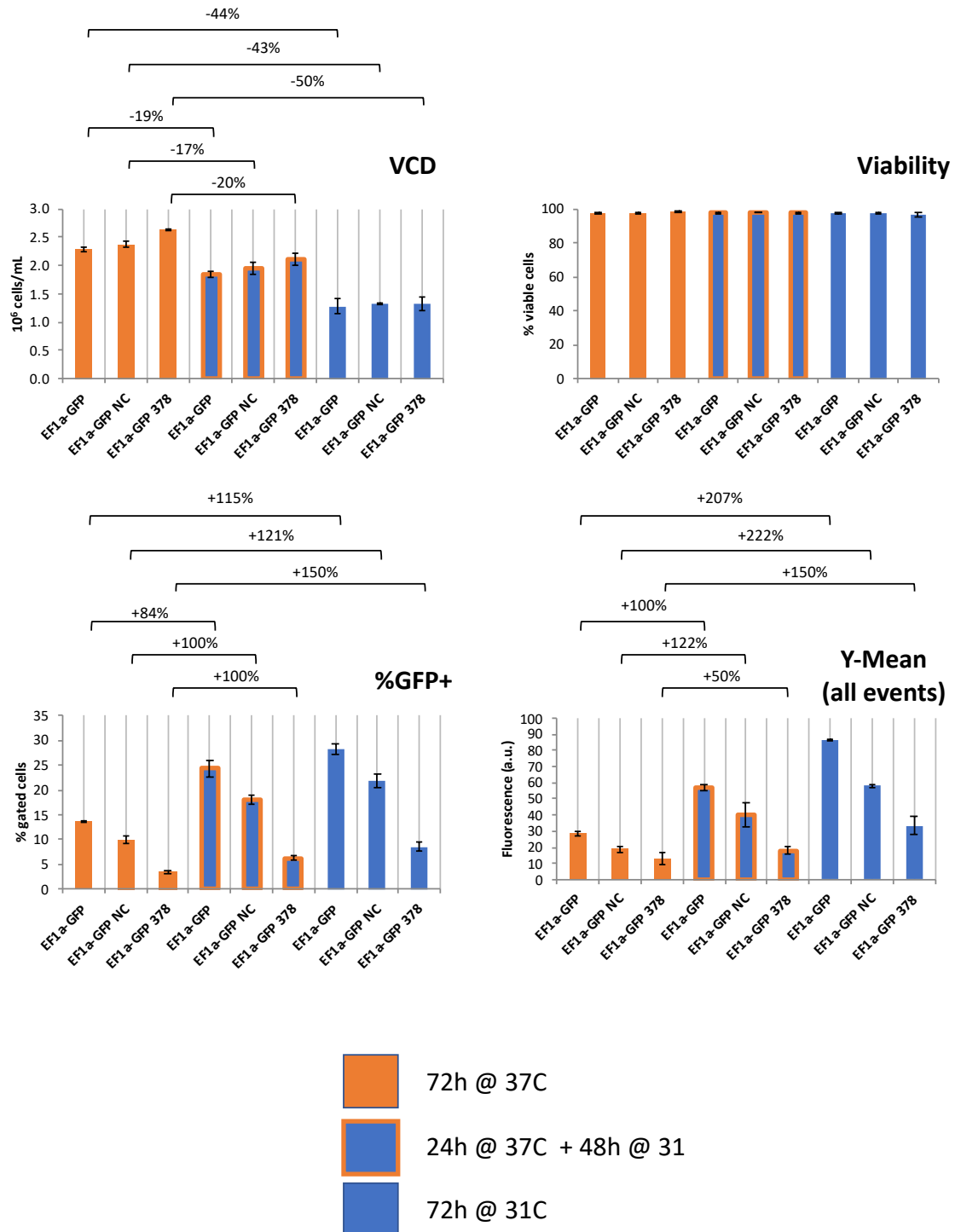


Figure 42 – Viable cell density, viability, %GFP positive cells, mean fluorescence intensity for all measured events (Y-Mean – all events) for transient pools 72h after temperature shift. Cells were seeded in 6-well cell culture plates at 10^5 cells/well 24h after being transfected with EF1a-GFP vector, EF1a-NC-sponge and EF1a-378-sponge. Plates were incubated at their respective temperature (see legend). Measurements correspond to the average of 2 technical replicate wells and error bars represent \pm standard deviation.

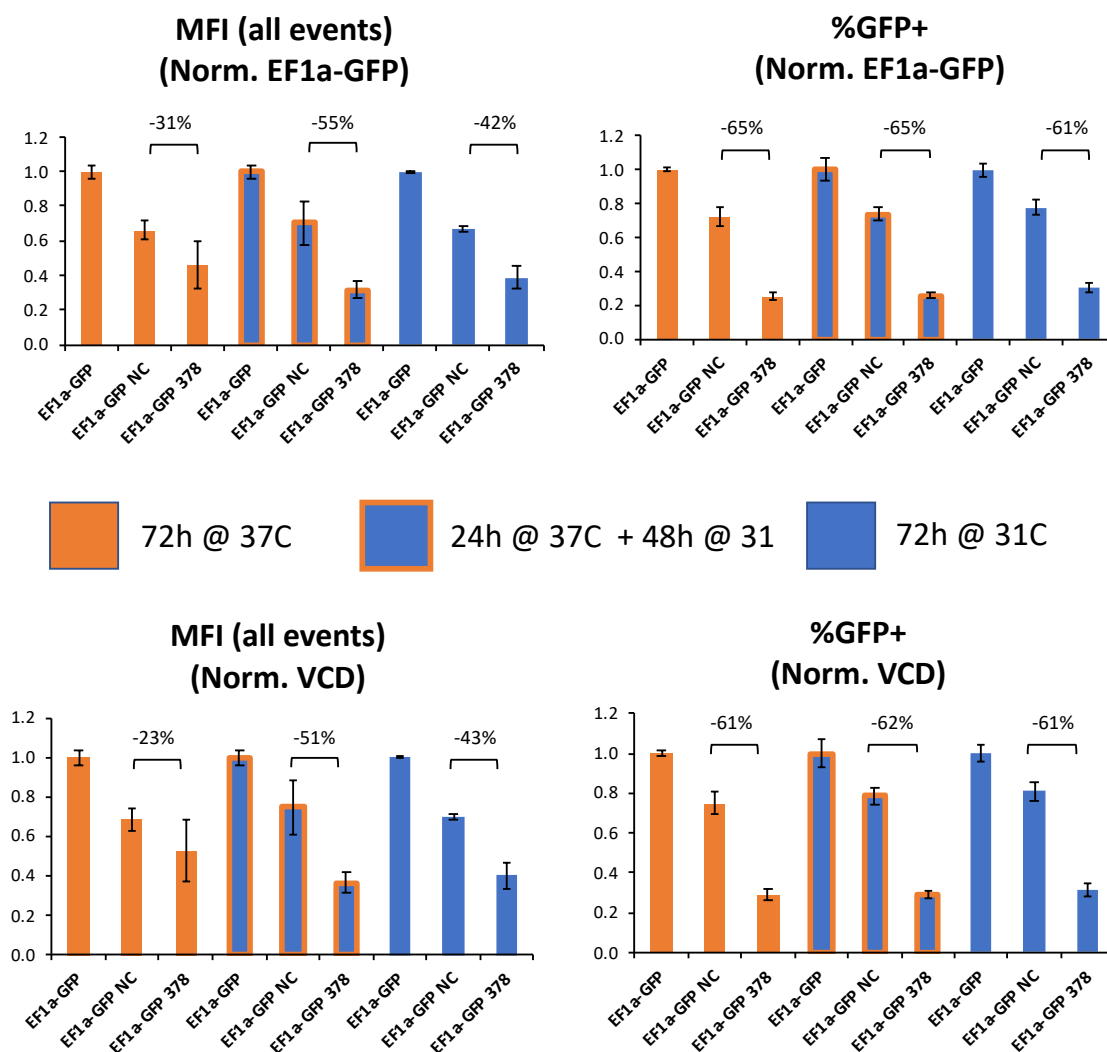


Figure 43 - Normalized data using EF1a-GFP from attached cells 72h after seeding. Cells were seeded in 6-well cell culture plates at 10^5 cells/well after 24h from being transfected with EF1a-GFP vector, EF1a-NC-sponge and EF1a-378-sponge. Measurements correspond to the average of 2 technical replicate wells and error bars represent \pm standard deviation.

showed higher levels of GFP both in terms of percentage GFP positive cells and mean fluorescence intensity than cells grown at 37°C in any of the two tested temperature shift setups. In order to be able to compare differences in GFP expression between NC-sponges transfected cells and 378-sponge transfected cells across the two temperatures, expression data was normalized using the control vector EF1a-GFP at each temperature or the VCD data for each sample as described above. Differences in expression between NC-sponge and 378-sponge expression for each temperature were calculated (Figure 43).

For these, normalised percentages of GFP positive cells were slightly increased at 31°C versus 37°C, normalised MFIs were found to be consistently lower in cells grown at 31°C

than in cells grown at 37°C. No differences were observed in terms of GFP positive cells between cells grown at 37°C and cells grown at 37°C for 24h and shifted to 31°C for 48h (2-3%) whereas there were 7-21% more GFP positive cells in the 31°C wells after 72h incubation. Mean fluorescence intensity (all events) was lower in both cases at 31°C than at 37°C, with a 31% reduction for cells grown at 37°C for 24h and shifted at 31°C for 48h and a 15-24% reduction for cells grown at 31°C for 72h. Looking at the differences observed between NC-sponge and 378-sponge at different temperatures, the presence of the 378-sponge caused a 60-65% reduction in %GFP positive cells when compared with the NC-sponge (Figure 43 – %GFP+). In terms of MFIs, the presence of a 378-FC-sponge led to a reduction of the MFI (all events) of 23-31% when compared to the NC-sponge control at 37°C, with a larger reduction observed at 31°C, 51-55% in the 37°C>31°C condition and 42-43% in the 31°C condition (Figure 43 – MFI all events).

Differences in GFP expression between temperatures for each individual plasmid were also calculated (Table 15). When looking at the effect of temperature on GFP expression, small differences were observed between NC-sponge plasmids at different temperatures in the range of 2-9% in terms of percentages of GFP+ cells and for GFP fluorescence intensity (MFI all events) whereas larger differences were observed for the 378-FC-spg transfected cells at the different temperatures (Table 15).

Table 15 – Differences in GFP expression between cells grown at 37°C and cells shifted to 31°C. 37°C > 31°C = 24h at 37°C followed by 48h at 31°C and 31°C = 72h at 31°C. Data is shown as percentage change from the reference 37°C condition.

Normalised using EF1a-GFP signal

MFI(all events)

	37C > 31C	31C
NC-spg	6%	1%
378-FC-spg	-31%	-15%

%GFP+

	37C > 31C	31C
NC-spg	2%	7%
378-FC-spg	3%	21%

Normalised using VCD data

MFI (all events)

	37C > 31C	31C
NC-spg	9%	2%
378-FC-spg	-31%	-24%

%GFP+

	37C > 31C	31C
NC-spg	5%	8%
378-FC-spg	2%	9%

With this experiment, the experimental setting to evaluate miRNA sponges in transient expression and temperature shift conditions was established. The two proposed normalisation methods led to similar results. Using VCD as normalizer was deemed more

appropriate considering EF1a-GFP signal might not be a good control for sponges that show high levels of repression and would therefore express less GFP.

2.13. Evaluation of transiently transfected EF1a-driven miRNA sponges: effects on GFP expression

The previously established protocol was used in order to evaluate GFP expression from the different miRNA-sponges and to assess miRNA-dependent changes in GFP expression after temperature shift.

In short, equal amounts of the previously described miRNA-sponge plasmids (Table 13) as well as a positive EF1a-GFP control vector were transiently transfected in CHO-K1 cells in 24-well suspension format. GFP expression was measured at 24h and 48h after transfection. GFP expression was verified for all the tested miRNA-sponge plasmids and again, viability was lower for cells expressing higher levels of GFP (Figure 44).

As previously observed, GFP expression obtained from the short NC-sponge plasmid was similar to GFP expression obtained from the EF1a-GFP control, but a reduction in %GFP positive cells and MFI was observed for the longer NC-sponge plasmid.

As expected, GFP expression was reduced by the presence of miRNA-specific sponges when compared with their respective NC-sponge control. Interestingly, for most of the constructs the observed effect of miRNA-sponges was increased over time, with larger reductions of GFP expression observed after 48h and 72h (Figure 45). By contrast, 34a-spg-2 showed the same effect independently of time, while the effects observed from 378-spg-8 and 378-30e-spg-4 only increased after 72h.

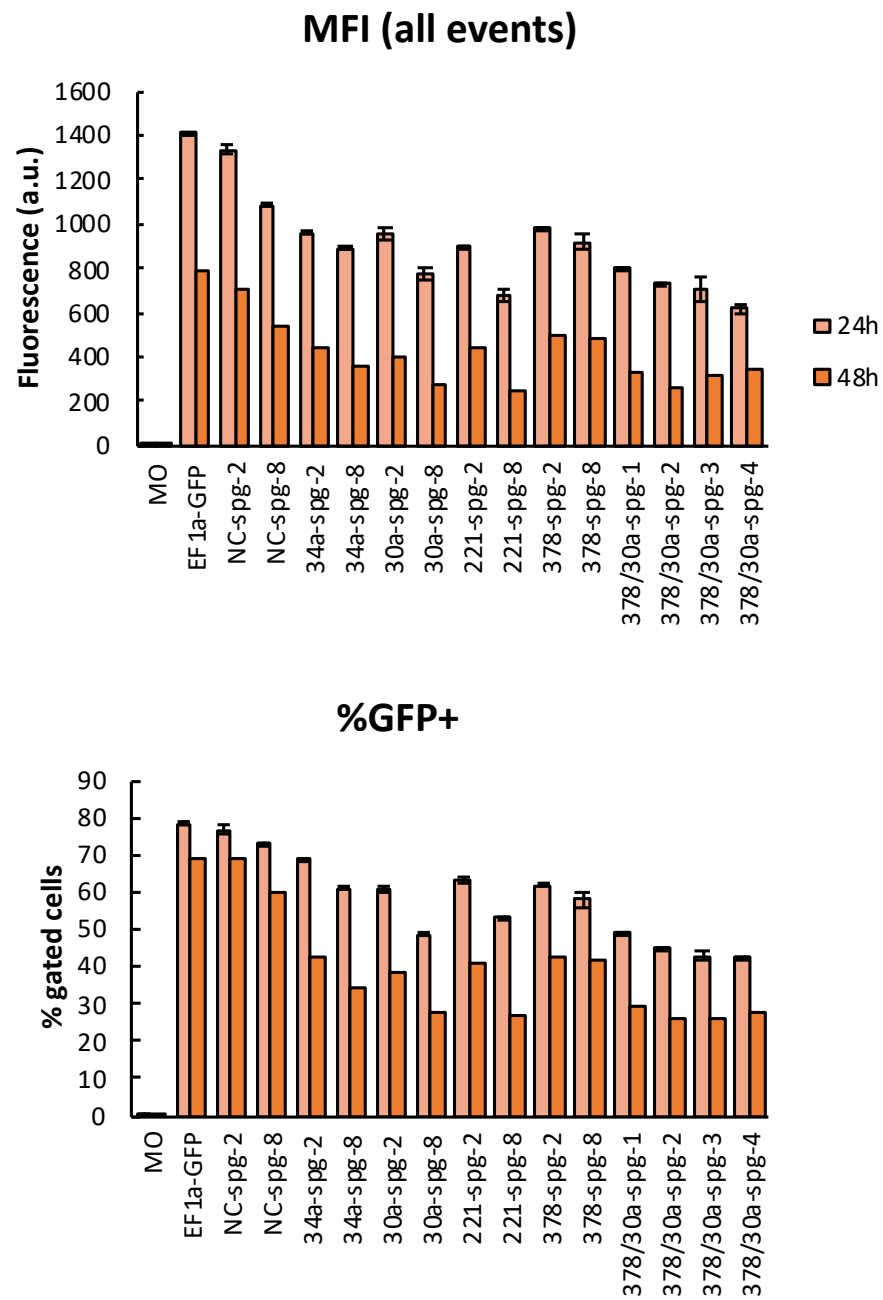


Figure 44 – GFP expression (MFI – all events and %GFP+ cells) from the different temperature sensitive miRNA sponges grown at 37°C. GFP expression were measured at 24h and 48h after transfection in 24-well suspension plates. Each bar represents the average of 3 technical replicates (N=1) and error bars represent \pm standard deviation.

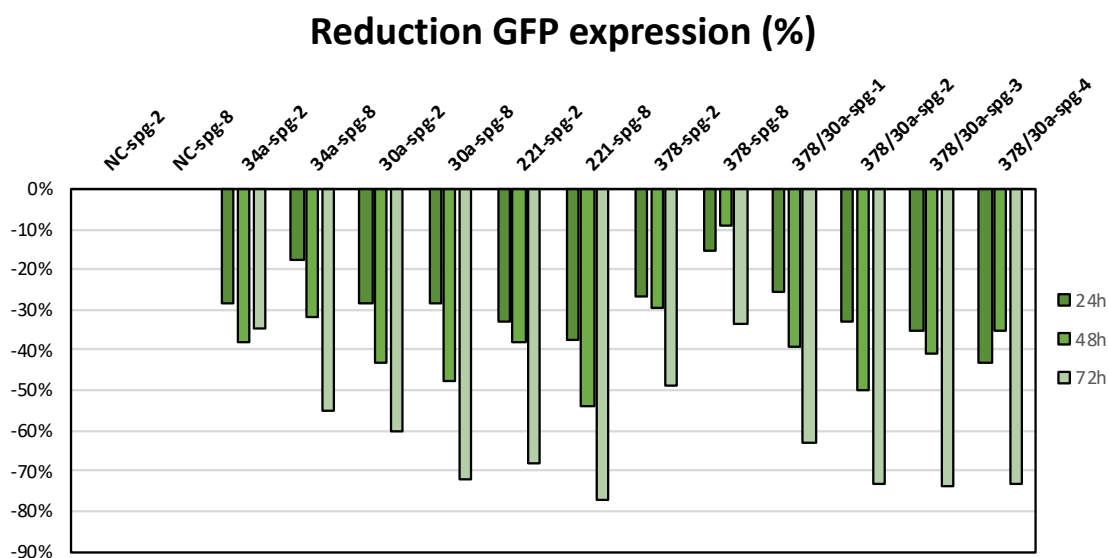


Figure 45 - Reduction (%) of GFP expression by specific miRNA-sponge in relation to their corresponding NC-sponge control. GFP reduction was calculated for each miRNA-sponge versus its respective NC-sponge control (2 or 8 binding sites) for measurements at 24, 48 and 72h after transfection. Each bar represents the average of 3 technical replicates (N=1).

Having verified that the different miRNA-sponges not only express GFP but are also able to cause downregulation of the reporter, transient pools at 24h after transfection were used to investigate the effect of temperature shift on the observed miRNA-dependent GFP inhibition.

2.14. Transient evaluation of miRNA-dependent changes in GFP expression in response to temperature shift

A temperature shift experiment was setup to assess the potential miRNA-driven changes in GFP expression after reducing culture temperature. In short, transiently transfected cells described in Section 2.13 cells were seeded in duplicate 24-well plates in DMEM:F12 +5% FBS at 50,000 cells per well in triplicate wells per construct. Half of the plates were incubated at 37°C and the other half were incubated at 31°C. After 48h, cells were trypsinized and GFP expression and viability measurements were taken.

In terms of cell growth, viable cell density between 37°C and 31°C was reduced at the lower temperature (Figure 46 - middle plot). This was previously observed during optimization experiments and was a good confirmation that the temperature shift phenotype remained. It is interesting to note that these growth differences were larger in cells with repressed GFP expression than in cells transfected with control EF1a-GFP

plasmid and control NC-sponge which expressed higher GFP and for which smaller or no growth differences were observed between 37°C and 31°C.

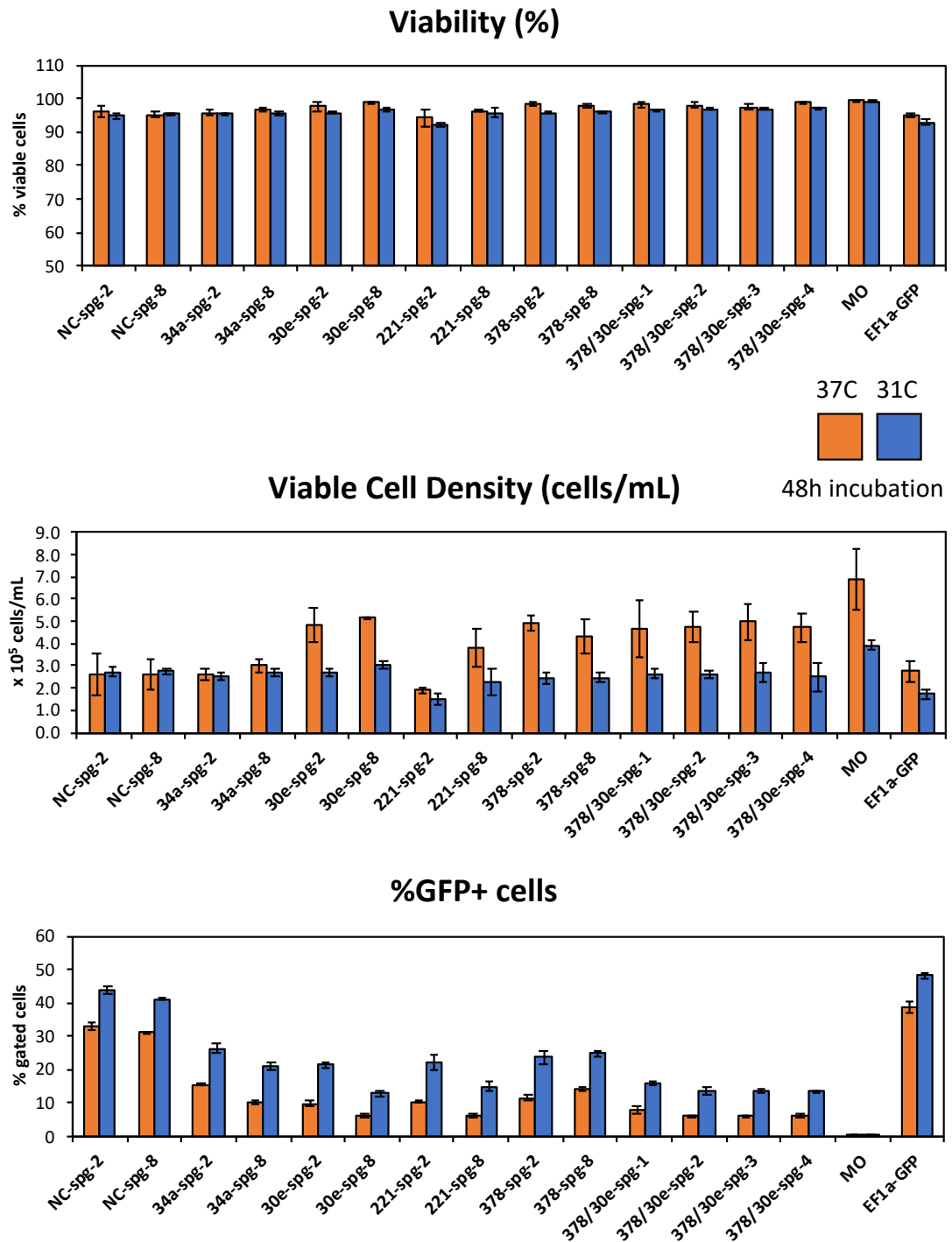


Figure 46 – Viability, viable cell density and %GFP positive cells measured in miRNA-sponge transiently transfected cells after 48h incubated at 37°C or 31°C. Each bar corresponds to the average of 3 technical replicate wells and error bars represent \pm standard deviation.

Regarding GFP expression after temperature shift, transient pools grown at 31°C exhibited larger percentages of GFP+ cells (Figure 46 - bottom plot), as well as larger

mean fluorescence intensity than the respective pools grown at 37°C (Figure 47). Flow cytometry GFP histogram plots for all the analysed samples are provided in Supplementary Figures 3-5 – Appendix B.

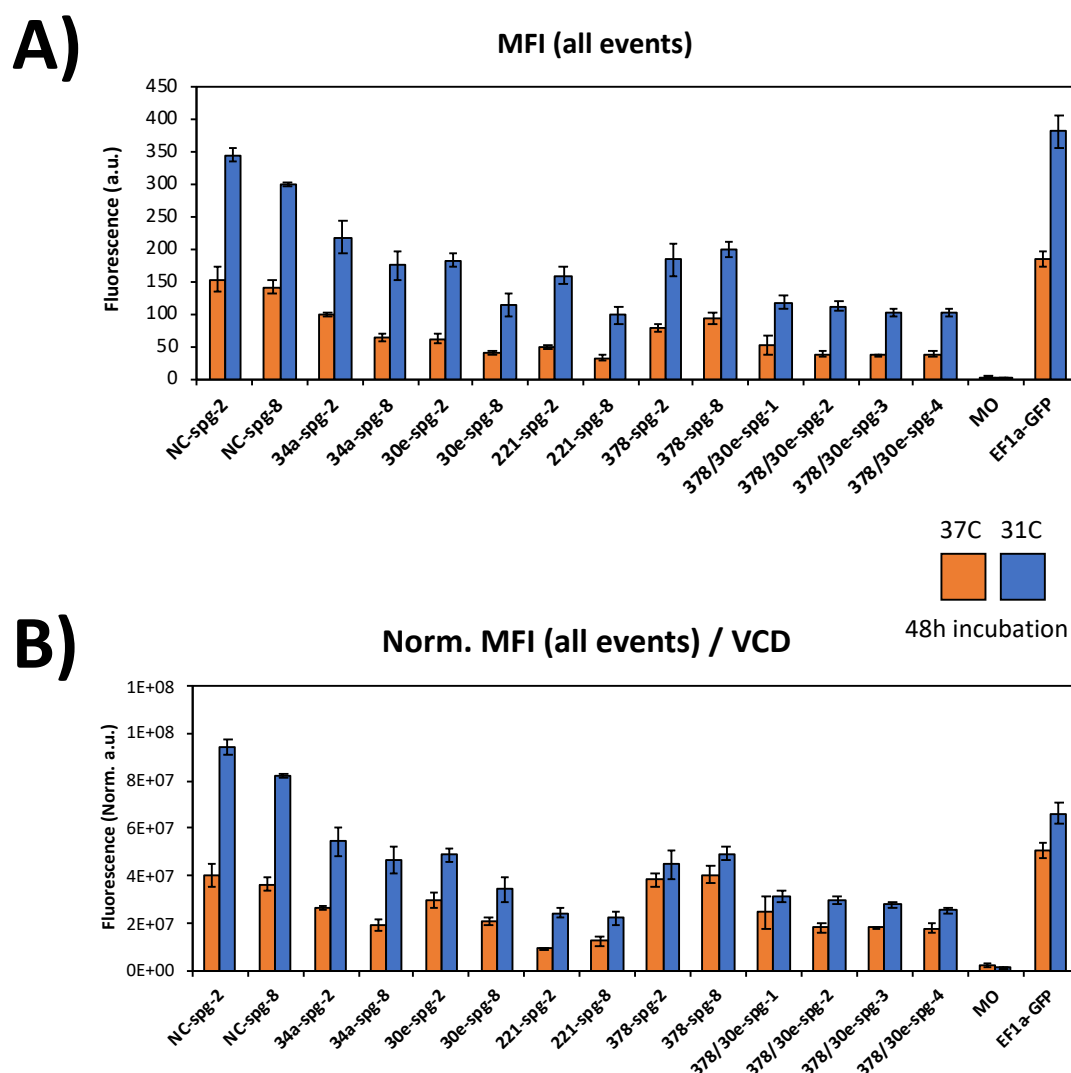


Figure 47 – A) GFP expression (Mean Fluorescence Intensity – all events) and B) Normalized GFP expression measured in the different miRNA-sponge transiently transfected populations after 48h incubated at 37°C or 31°C. Each bar corresponds the average of 3 technical replicate wells and error bars represent \pm standard deviation.

As previously described, in order to compensate for the observed differences in growth derived from the temperature shift itself and the different levels of GFP observed, cell density (VCD) values were used to normalised MFI (all events) for each sample (Figure 47 - bottom plot). It can be observed how taking into account cell number slightly changes the GFP expression pattern observed from the different plasmids. In general, MFI (all events) for cell populations grown at 31°C were increased when compared with

those of cell populations grown at 37°C for all the tested sponges. However, looking at the raw (not normalized) data would lead to the conclusion that a reduction in temperature leads to more than a 2-fold increase in protein expression for most of the plasmids. However, normalised data using VCD depicts a different scenario. While the increase in protein expression is still maintained for all the plasmids, the magnitude of the changes varied between different sponges.

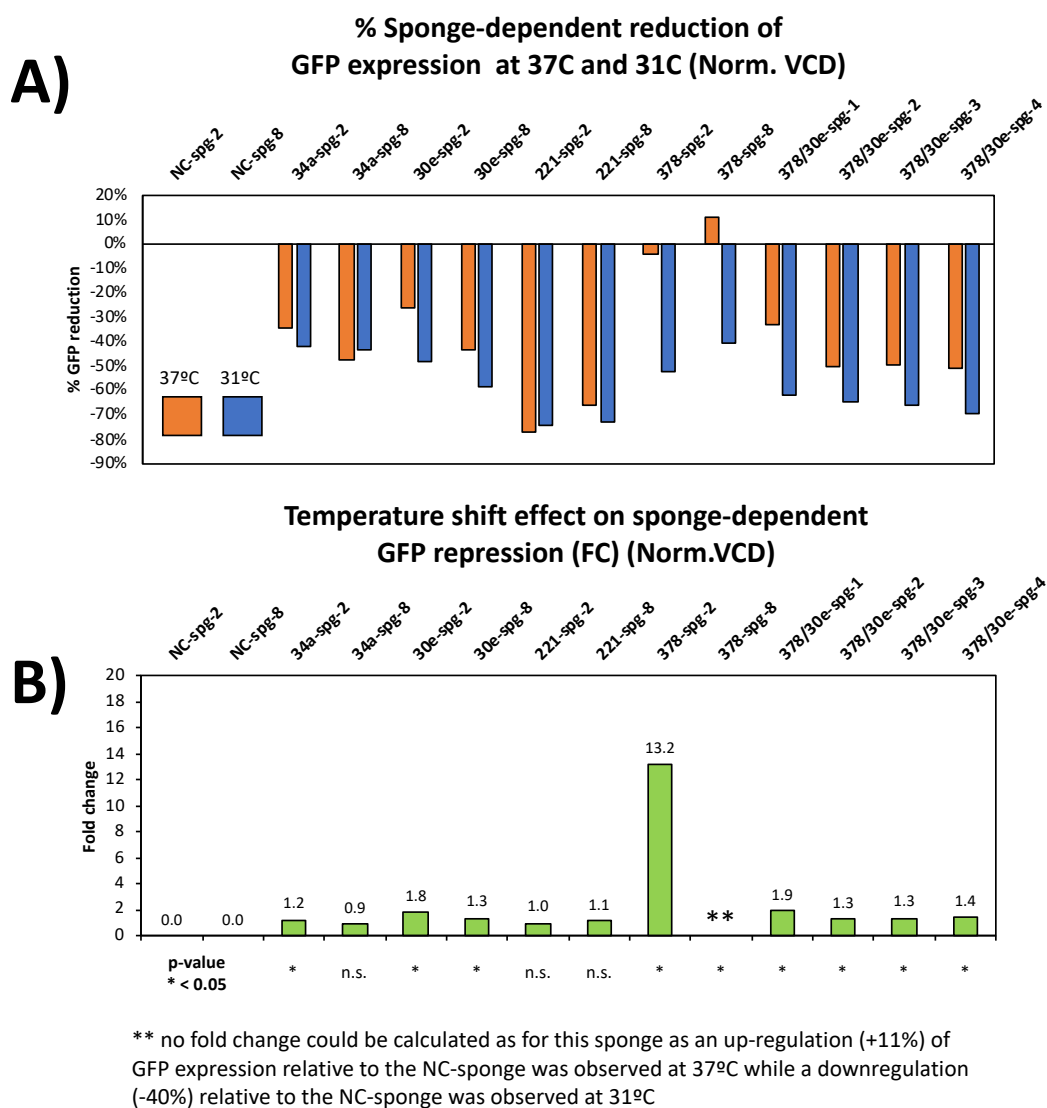


Figure 48 – A) Reduction of GFP expression for each miRNA sponge plasmid in relation to the NC-sponge control at 37°C (orange) and 31°C (blue) and B) Effect of temperature shift on the miRNA-sponge dependent GFP repression. GFP expression data measured in the different miRNA-sponge transiently transfected populations after 48h incubated at 37°C or 31°C. Each bar represents the ratios between the average of 3 replicate wells for each condition.

For example, the increase in GFP expression upon temperature shift for the EF1a-GFP plasmid was much lower than the increase observed for the two NC-sponge control vectors (Figure 47 – B).

In a second step, the percentage of change between the reduction observed at 37°C and the reduction observed at 31°C was calculated to compare the observed sponge-dependent GFP repression at 31°C versus 37°C (Figure 48).

Overall most of the miRNA-sponges tested were slightly more efficient at driving a reduction of GFP expression at 31°C than at 37°C. Strikingly, the largest efficiency changes were observed for 378-sponges. The 378-sponge with 2 binding sites was 13-fold more efficient at repressing GFP at 31°C than at 37°C. Even more surprisingly, GFP expression from the 378-sponge with 8 binding sites was not repressed at 37°C, where normalize expression from this plasmid was found larger than the expression observed for the NC-sponge. However, at 31°C, the 378-sponge-8 plasmid caused a 52% reduction of GFP expression in comparison with the NC-sponge control at that temperature.

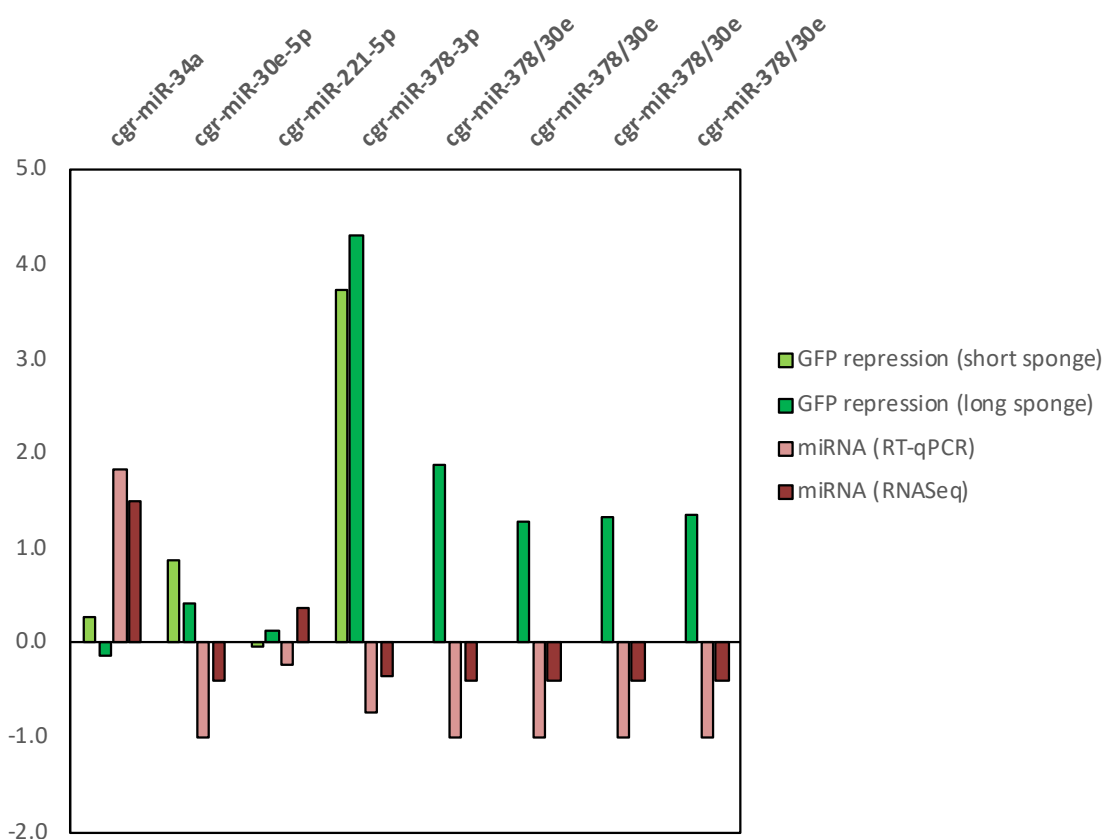


Figure 49 - Side by side comparison between temperature-driven changes in miRNA levels (red bars) and changes sponge-dependent GFP repression upon temperature shift (green bars) observed in transient experiments.

The initial hypothesis in this work was that temperature driven changes in miRNA expression should be able to drive changes in the expression of a transgene placed under the control of a specific miRNA-sponge. However, no correlation was found between the observed changes in sponge dependent GFP repression and the changes in miRNA expression previously established. Looking at each individual miRNA/sponge pair, it was expected that an increase in miRNA expression at 31°C would lead to an increase in miRNA sponge repression efficiency and vice versa. Instead, for the miRNAs with decreased expression upon temperature shift, an increase in their respective sponge-dependent GFP repression was observed (Figure 49).

Taking into account the challenges encountered when performing transient expression experiments and the above-mentioned need to control for several different factors affecting GFP expression other than miRNA expression that might be masking the investigated miRNA-related effects, stable pools were generated to further investigate those in depth.

2.15. Evaluation of EF1a-driven miRNA sponges in stable mixed pools: effects of miRNA sponge on GFP expression

In order to further explore the effect of temperature on miRNA-driven changes in GFP expression, and to simplify the experimental model used, stable mixed pools expressing the different miRNA-sponge constructs were generated.

In short, approximately 3×10^5 cells from the transiently transfected populations described in Section 2.13 were seeded in 6-well plates and sub-cultured for 2+ weeks in Hygromycin selection media. After an initial drop in GFP fluorescence signal during the first two passages, cell pools with up to 50% GFP+ cells were obtained from P3 onwards (Figure 50 - left). As expected, stable pools expressing miRNA-sponges showed different levels of GFP expression. As previously observed, the GFP expression from NC-sponge plasmids was slightly lower than the GFP expression obtained from the EF1a-GFP control, with a small reduction in %GFP positive cells and MFI observed in stable pools expressing the longer NC-sponge plasmid when compared with stable pools expressing the shorter NC-sponge plasmid (Figure 51). GFP expression was reduced by the presence of miRNA-specific sponges when compared with their NC-sponge control (Figure 52).

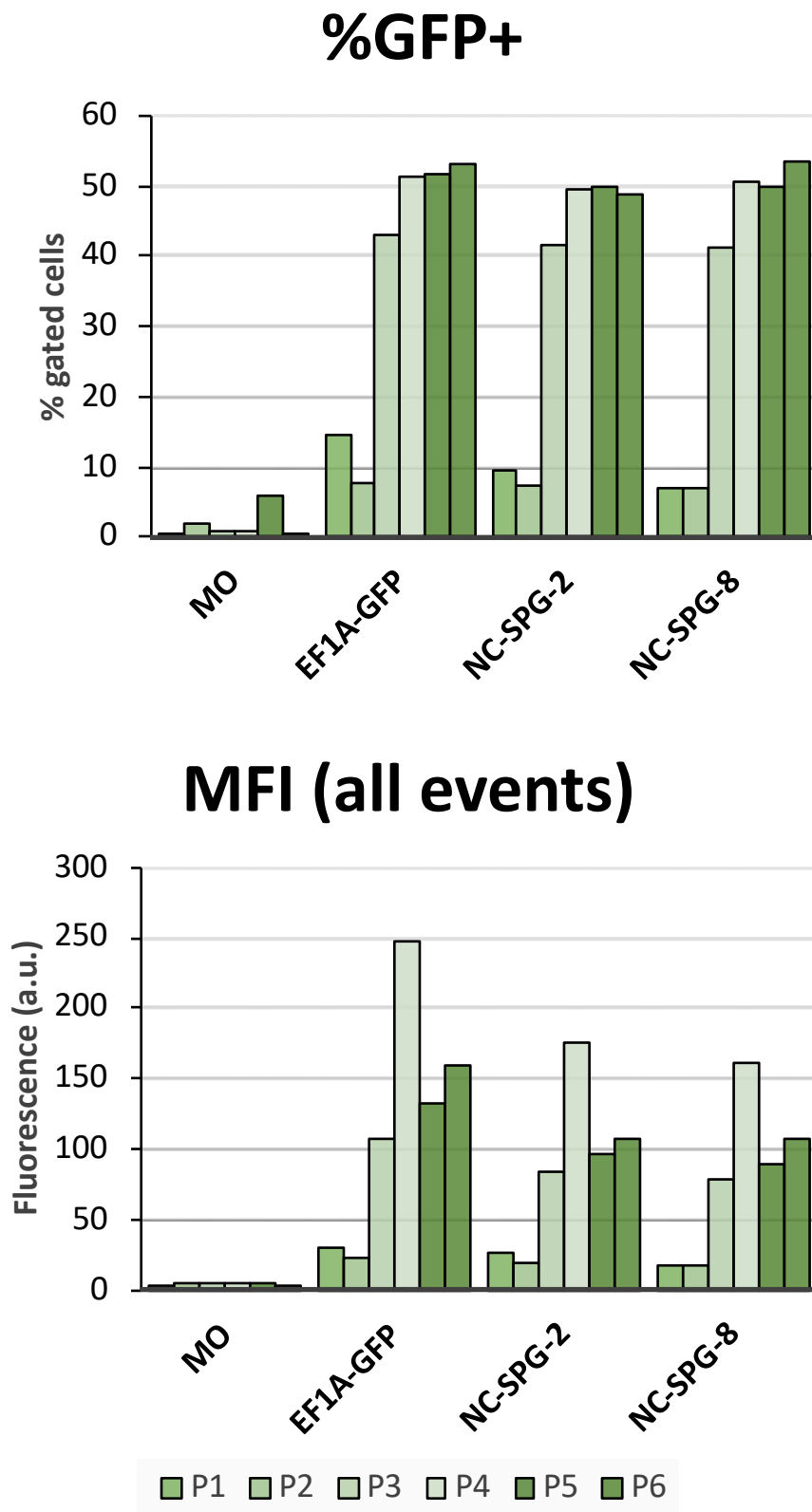


Figure 50 - Generation of GFP-sponge stable cell lines by selection with Hygromycin. Percentage of GFP positive cells (left) MFI for all recorded events and MFI for the GFP positive population for Passage 1 to Passage 6. Note passage 4 was shorter (2 days instead of 4 days as cell were used to set experiments). Top plot: each bar represents the percentage of GFP+ cells in the measured sample. Bottom plot: each bar represents the average mean fluorescence intensity (MFI) for +5000 recorded events.

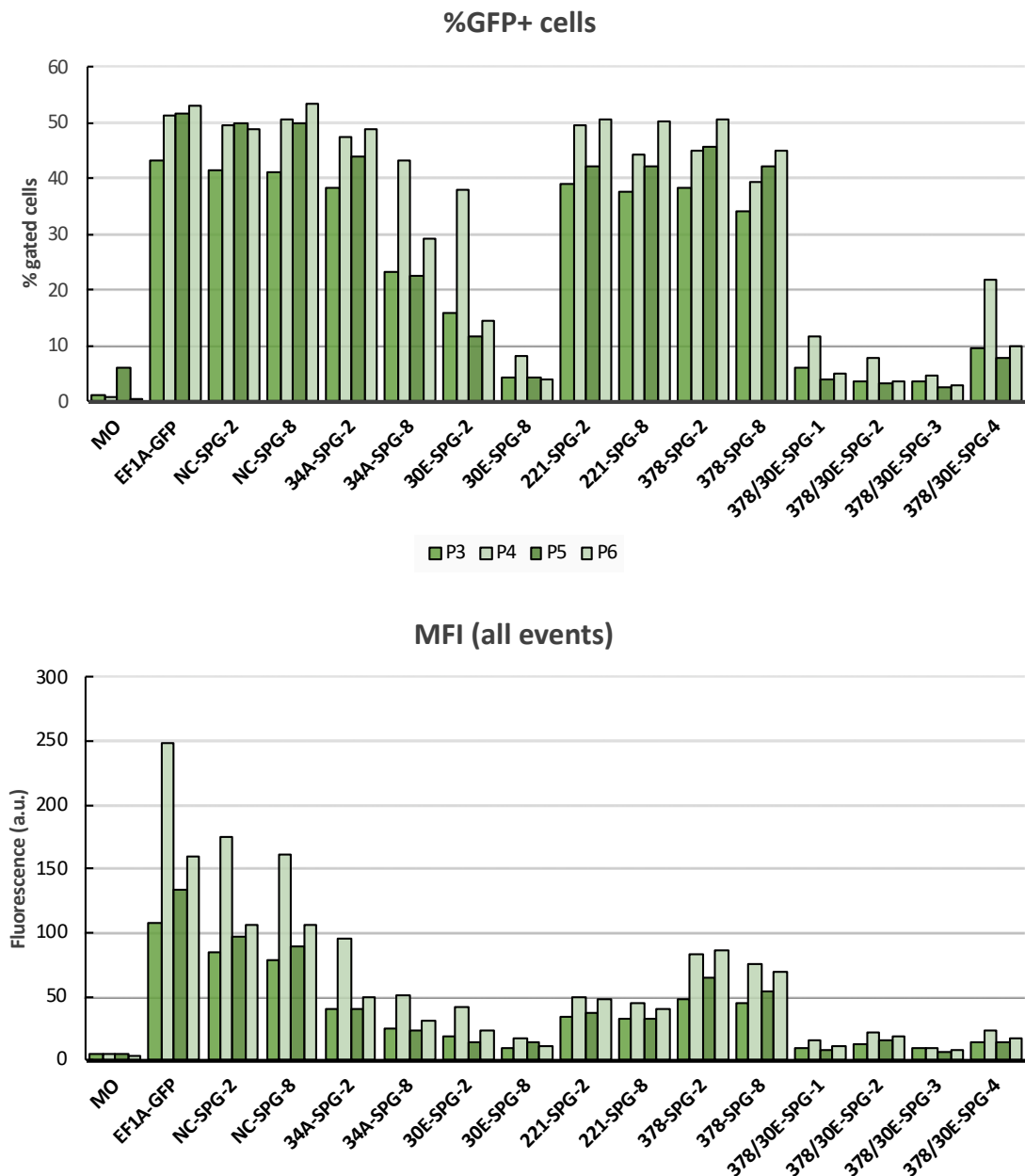


Figure 51 - GFP expression obtained in stable expression pools expressing miRNA-sponges. Shown data corresponds to cells selected under Hygromycin selective conditions (500 µg/mL). Top plot shows percentage of GFP positive cells and bottom plot shows mean fluorescence intensity for all recorded events (5000+). Top plot: each bar represents the percentage of GFP+ cells in the measured sample. Bottom plot: each bar represents the average mean fluorescence intensity (MFI) for +5000 recorded events.

It is interesting to note that the observed GFP expression in the stable pools was similar to the ones observed in the previous transient experiments using both MFI and normalised MFI data (Supplementary Figure 7 – Appendix B). However, the GFP reduction caused by miRNA-sponges observed in the stable pools, which remained very consistent during the various passages, did not correlate well with the GFP reductions

observed in the transient model for the different sponges (Supplementary Figure 8 – Appendix B).

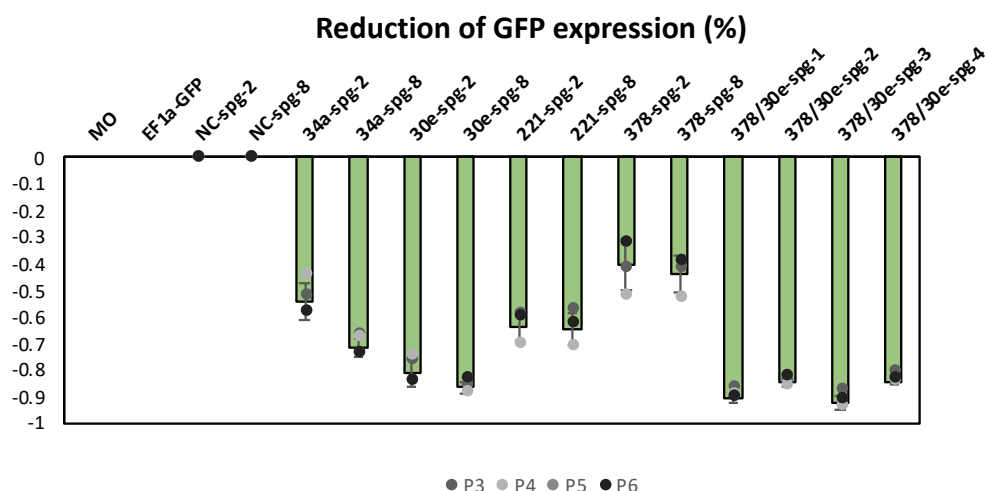


Figure 52 - Reduction (%) of GFP expression by specific miRNA-sponge in relation to their correspondent NC-sponge control. GFP reduction was calculated from average values for all passages (P3-P4 shown with dots) for each miRNA-sponge versus its respective NC-sponge control (2 or 8 binding sites) for cells before seeding new plates (data shown in Figure 51). Each bar represents the average of triplicate wells and error bars represent \pm standard deviation.

As previously done for the transiently transfected pools, stable pools were used to investigate the effect of temperature shift on miRNA-sponge mediated GFP repression as described in the following section.

2.16. Evaluation of temperature shift on miRNA-driven changes in GFP expression using stable mixed pools expressing miRNA-sponges

As done before for the transient pools, a temperature shift experiment was setup to assess miRNA-driven changes in GFP expression from miRNA-sponges after reducing culture temperature from 37°C to 31°C. In short, cells from the stable mixed pools described in Section 2.15 at passage 4 (P4) were seeded in duplicate 24-well plates in DMEM:F12 +5% FBS at 50.000 cells per well in triplicate wells per each construct. The temperature shift setting was adapted to make it closer to the one used to profile miRNA expression in Section 2.8. Namely, plates were incubated at 37°C for 48h then half were shifted to 31°C and incubated for 24h before sampling. cells were trypsinized and GFP expression measurements were taken by flow cytometry (Figure 53).

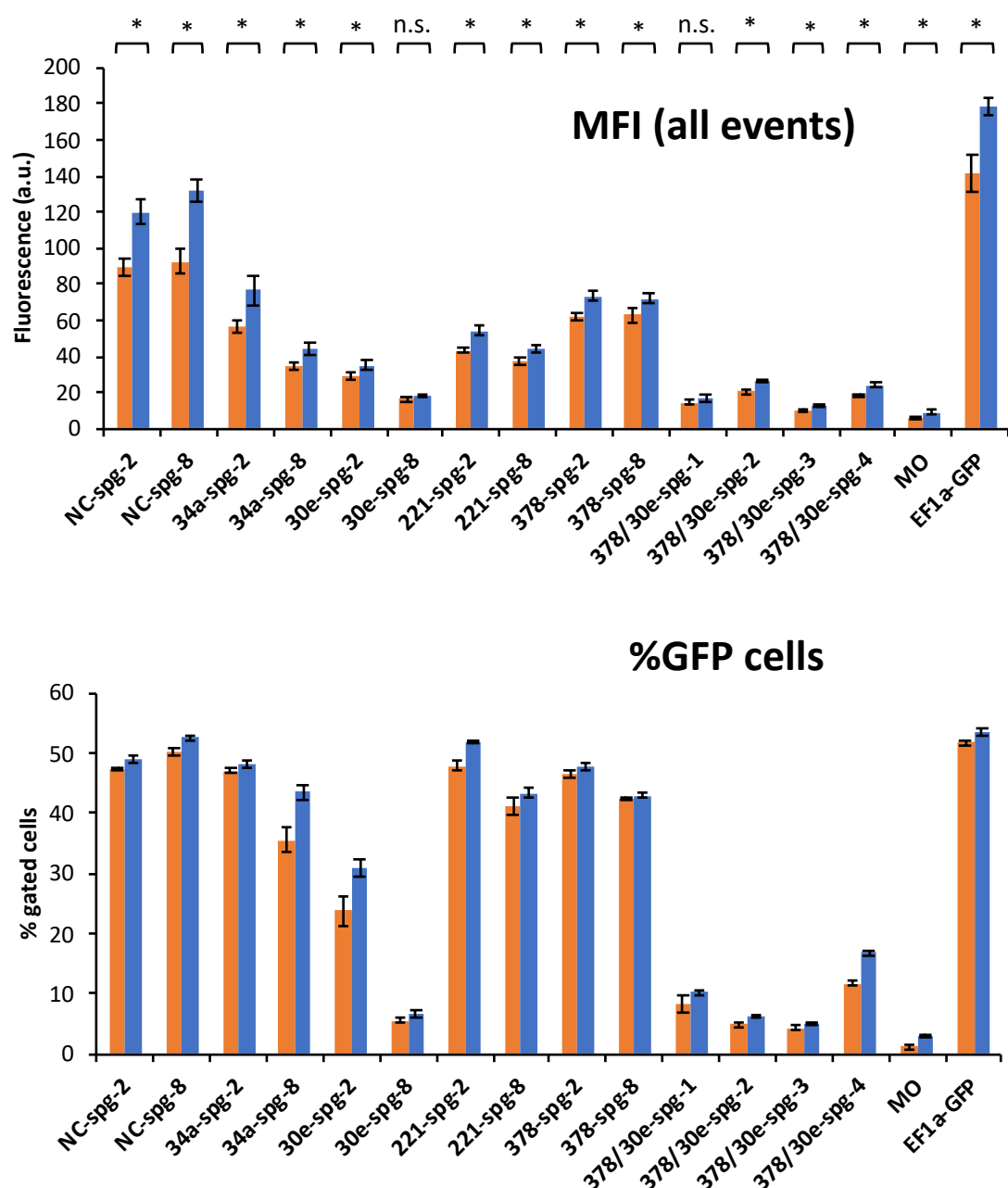


Figure 53 - GFP expression (MFI all events and %GFP) for stable pools after temperature shift: cells were grown at 37°C for 72h (orange) or at 37°C for 48h then shifted to 31°C for 24h (blue). Each bar represents the average of triplicate wells and error bars represent \pm standard deviation. MFI differences between 31°C and 37°C were evaluated using a two-tailed homoscedastic Student T-Test. * = p-value <0.05 and n.s. = not significant.

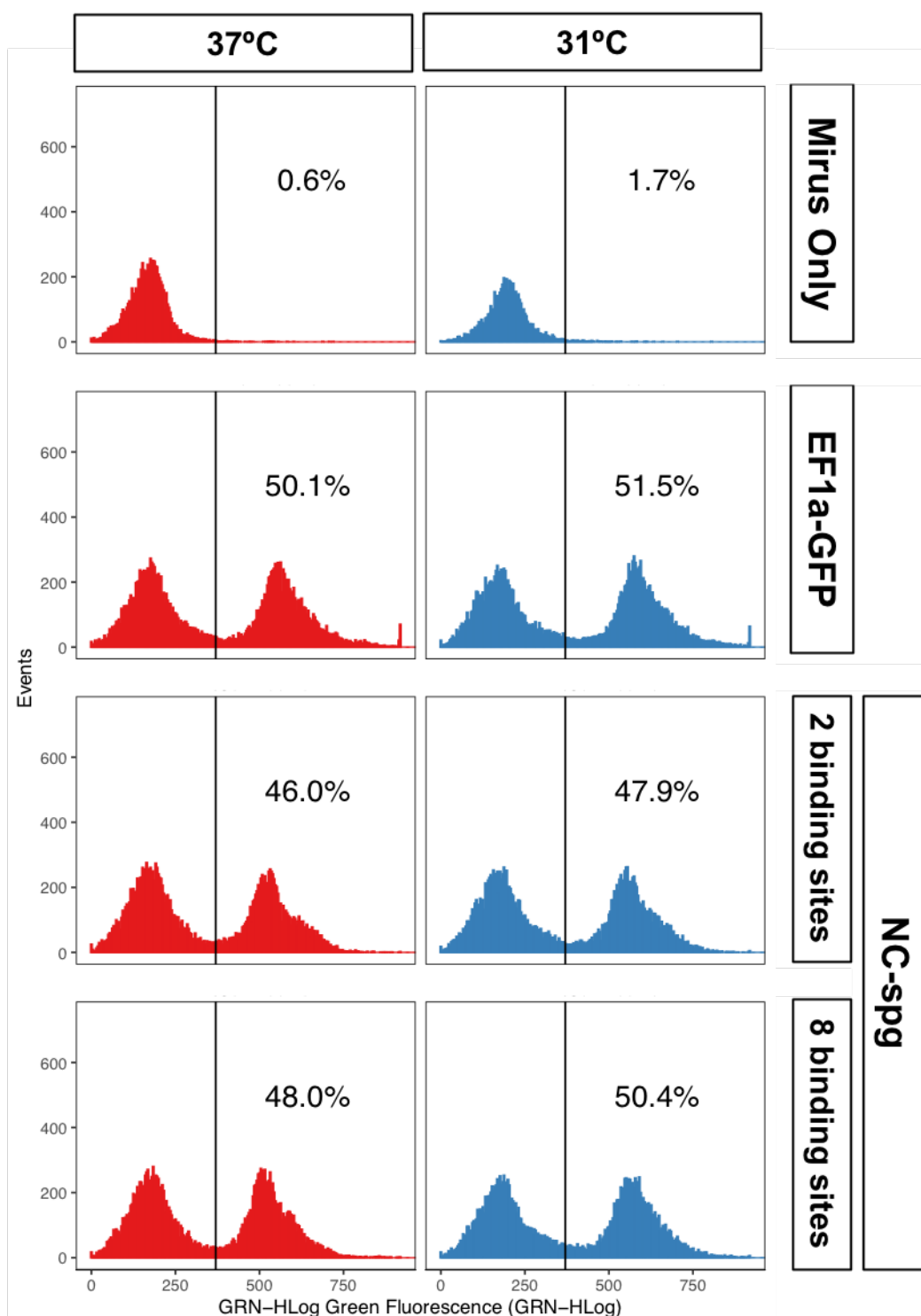


Figure 54 – Histogram plots showing flow cytometry data for GFP fluorescence (GRN – HLog) for stable pools grown at 37°C and 31°C: negative control for selection (only transfection reagent - MO), a EF1a-GFP positive control and the two negative control NC-sponges with 2 and 8 binding sites. A total of 5000+ events were recorded for each sample. Percentages correspond to the proportion of cells gated as GFP+ for each pool.

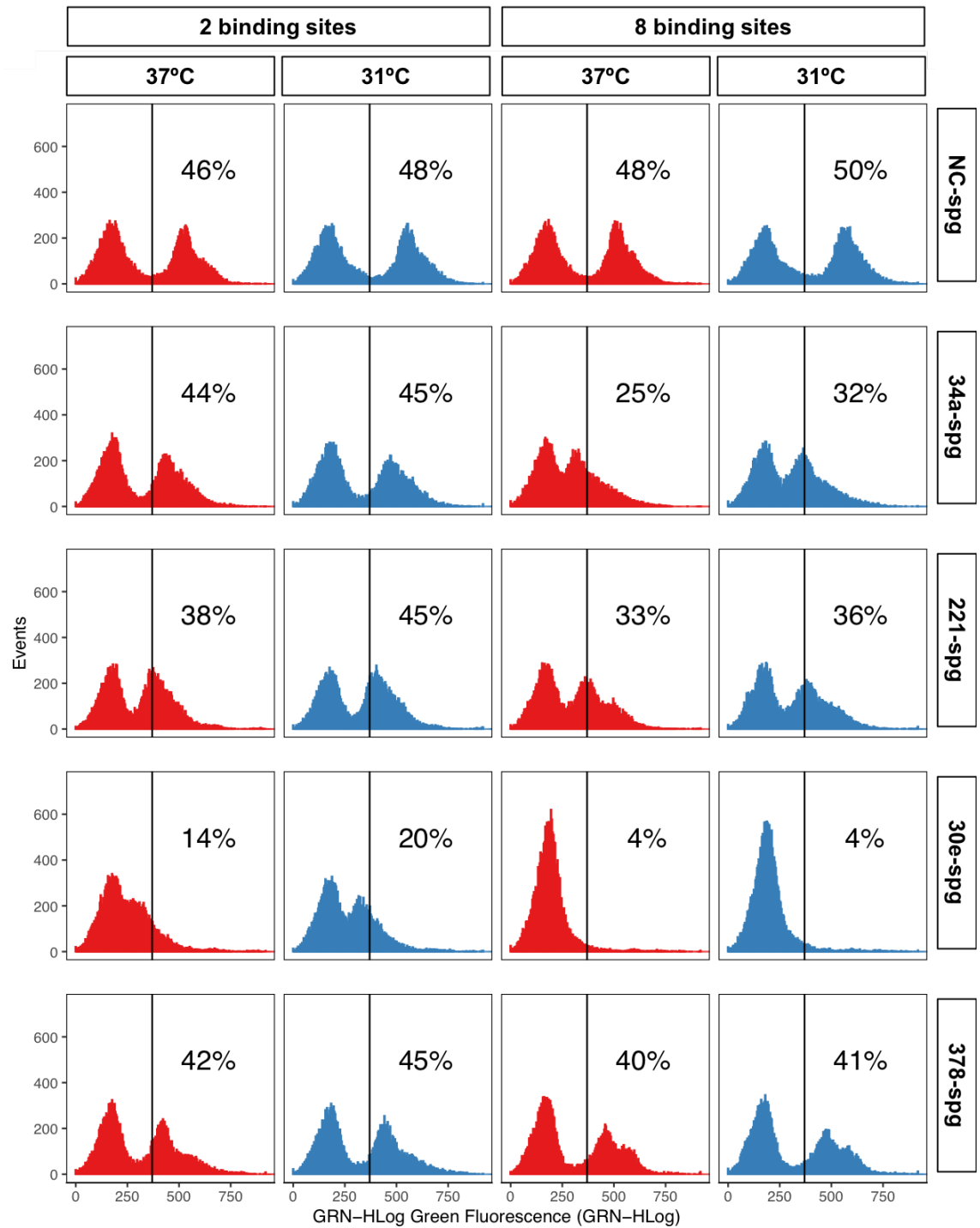


Figure 55 – Histogram plots showing flow cytometry data for GFP fluorescence (GRN – HLog) for stable pools grown at 37°C and 31°C. Pools expressing negative control NC-sponges with 2 and 8 binding sites, as well as pools expression each of the short and long miRNA-specific sponges are shown. A total of 5000+ events were recorded for each sample. Percentages correspond to the proportion of cells gated as GFP+ for each pool.

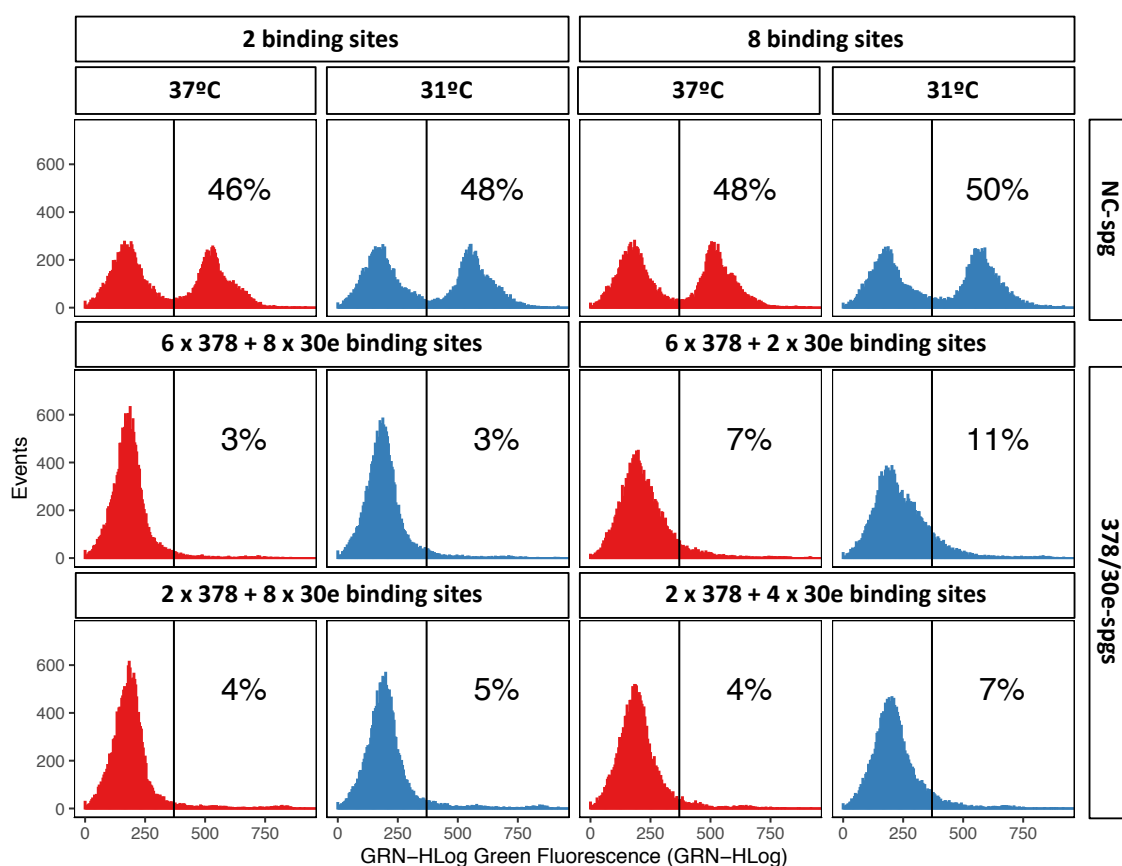


Figure 56 – Histogram plots showing flow cytometry data for GFP fluorescence (GRN – HLog) for stable pools grown at 37°C and 31°C. Pools expressing Negative control NC-sponges with 2 and 8 binding sites, as well as pools expression each of the short and long hybrid miRNA-specific sponges are shown. A total of 5000+ events were recorded for each sample. Percentages correspond to the proportion of cells gated as GFP+ for each pool.

Temperature shift increased GFP expression in all tested mixed pools with MFIs (all events) at 31°C between 9-42% higher at 31°C than at 37°C. No significant changes in expression between 31°C and 37°C were observed in mixed pools expressing 30e-spg (8 binding sites) or 378/30e-spg-1, (Figure 53 – top). Small differences in the percentages of GFP+ cells were observed between pools at 37°C and pools at 31°C. This can be explained by looking at the full flow cytometry data. Each different stable pool displayed a different GFP expression distribution when looking at the fluorescence signal for each individual cell (Figure 54, Figure 55, Figure 56). Most of the pools showed two peaks corresponding to dark and green populations respectively. It is interesting to note that while in some of the pools, these two populations, GFP negative and GFP positive, could be well distinguished (e.g.: 34-spg pools or 221-spg pools) for some others, the miRNA-sponge was so efficient at repressing GFP expression that those populations were

indistinguishable from an untransfected CHO-K1 parental population. In some other pools, an overlap of the two populations could be observed, making it difficult to gate green cells. Once more, the need to look at the mean fluorescence for the whole population is highlighted here.

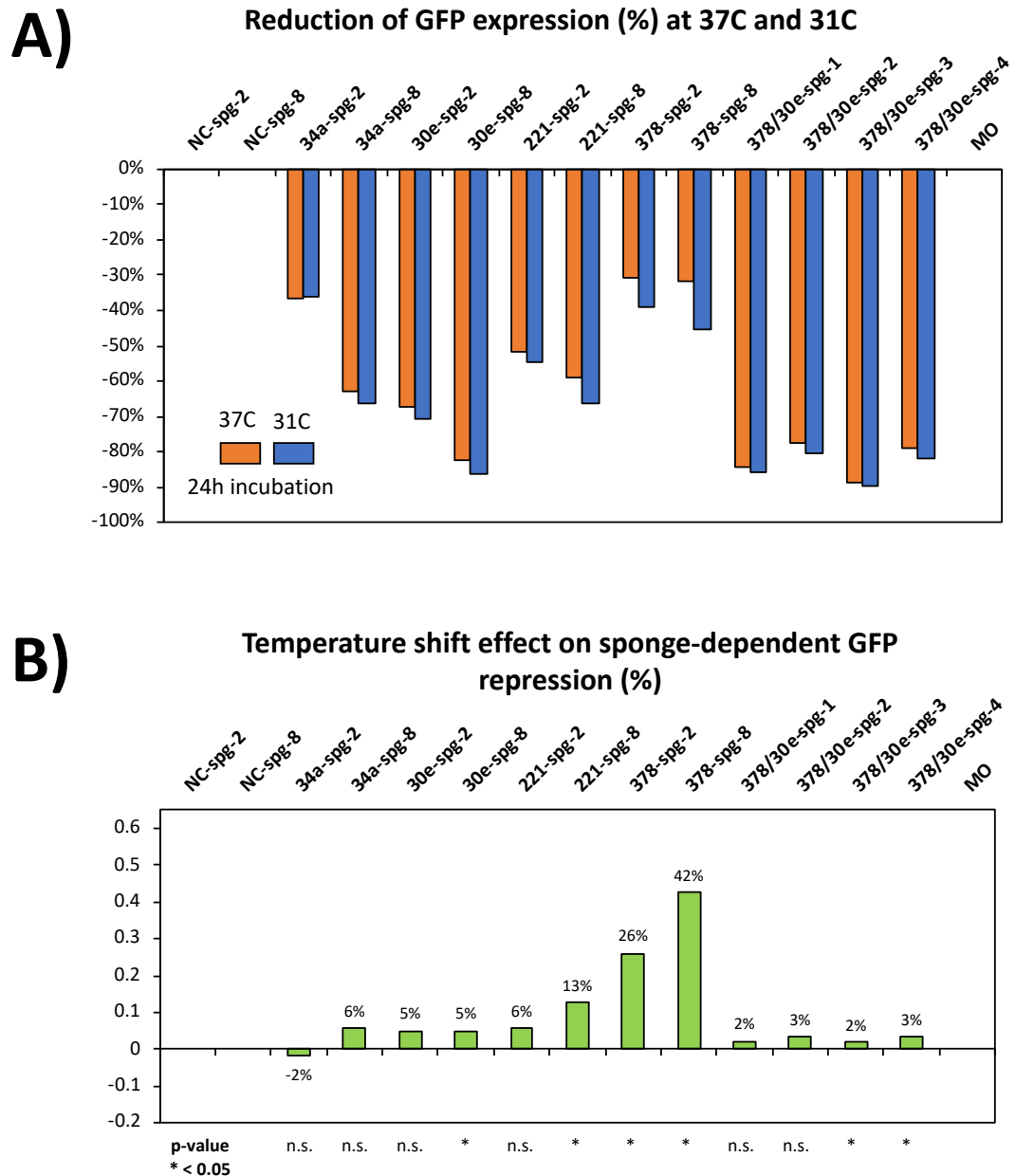


Figure 57 – A) miRNA-sponge-dependent changes in GFP expression in stable mixed pools after temperature shift (dark green for 37°C and light green for 31°C) B) Effect of temperature shift (37°C>31°C) on sponge-dependent GFP repression. Each bar represents the average change calculate from data shown in Figure 53. Differences between 31°C and 37°C were evaluated using a two-tailed homoscedastic Student T-Test. * = p-value <0.05 and n.s. = not significant.

The GFP expression profiles in pools stably expressing miR-378/miR-30e hybrid sponges are also interesting. While short sponges for miR-378-sponge or miR-30e were not able to completely down-regulate GFP expression, the hybrid sponge with 2 miR-378 binding sites and 4 miR-30e displayed very high levels of GFP repression. In a similar way, while the longer miR-378 sponge was not able to fully down-regulate GFP expression, the hybrid sponge containing 8 miR-378 binding sites and 2 miR-30e binding sites again showed strong GFP repression. These results suggest either cooperative effects between the two different miRNAs on the same transcripts or a scenario in which transcripts that are not repressed by one of the miRNAs are then bound by the second miRNA and repressed as well.

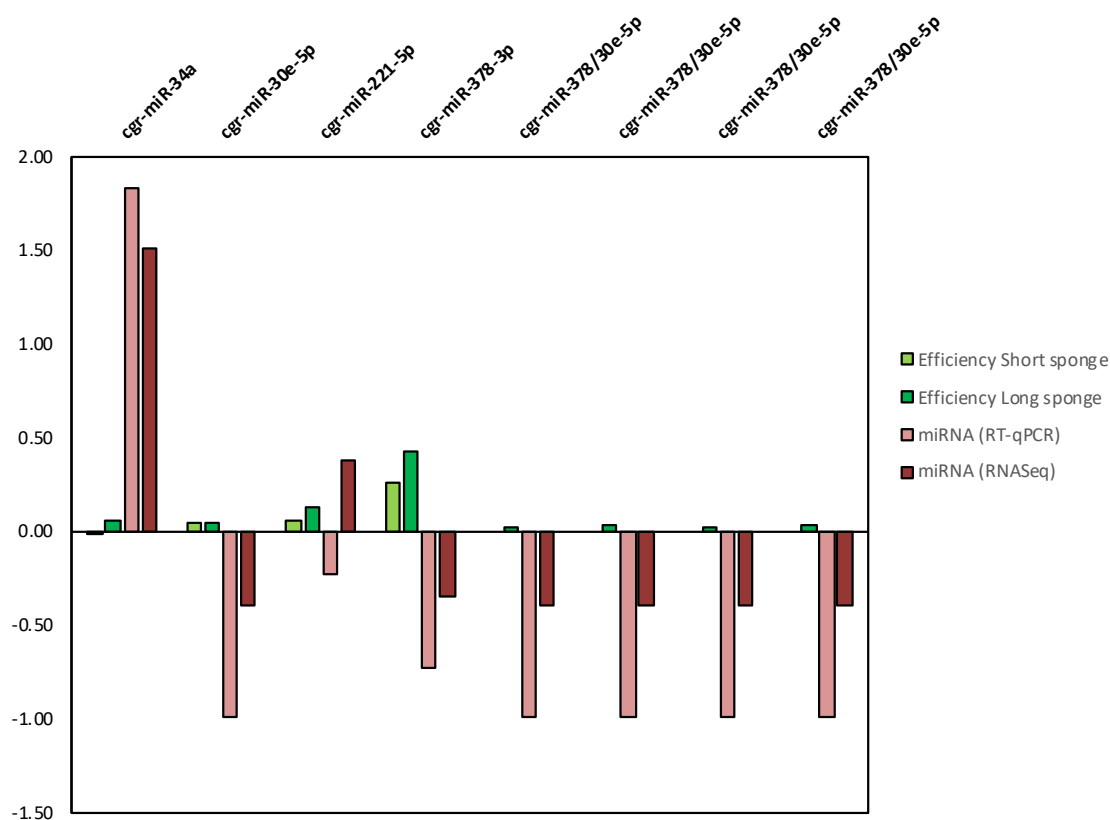


Figure 58 - Side by side comparison between temperature-dependent changes in miRNA levels (red bars) and changes in sponge efficiency upon temperature shift (green bars) observed in stable mixed pools after temperature shift from 37°C to 31°C.

As previously described, the sponge-dependent GFP repression for each individual miRNA-sponge was calculated in relation to their respective NC-sponge control (2 or 8 binding sites) for each temperature (Figure 57 – top). Differences between the observed sponge-dependent GFP repression at 37°C versus 31°C were calculated as a ratio between the two (Figure 57 – bottom). Significant differences in sponge dependent GFP

repression were observed between the two different temperatures for 6 out of the 12 tested sponges. miRNA sponges 30e-spg-8, 378/30e-spg-3 and 378/30e-spg-4 showed a small increase in their ability to reduce GFP expression (2-5%) at 31°C compared with 37°C. Larger differences were observed for miRNA-sponge 221-spg-8 (+13% reduction of GFP expression). Interestingly, even larger differences were observed between 31°C and 37°C for miRNA sponge 378-spg-2 and 378-spg-8 with 26% and 42% more GFP repression respectively.

Consequently, when looking at temperature driven changes for each individual miRNA/sponge in these stable pools, the findings observed in the transient experiments were replicated. Once more no correlations were found between miRNA expression and sponge dependent GFP repression (Figure 58).

3. Discussion

The work performed in the first sections of this chapter aimed to establish whether synthetic 3'-UTRs could be used to modulate GFP expression with a certain degree of control. The results obtained from the stable expression of GFP-sponge constructs showed that while it is possible link miRNA levels and GFP levels by using sponge UTRs, our simple system presented limitations in terms of clearly defining the relationship between sponge mRNA levels, protein output and miRNA levels. These limitations were mainly related to the sponge design and the expression levels of the integrated sponge construct. The sponges used had a rather large number of bulged binding sites (12), therefore it was anticipated to see changes at the protein level driven by translation inhibition rather than by miRNA mediated mRNA degradation, known to occur with full complementary binding sites (Ebert, Neilson and Sharp, 2007). It was also anticipated to observe a so called "sponge" effect due to miRNAs being sequestered upon binding the sponge. In this situation, relating changes at the miRNA to changes in GFP would be harder to do. It must be noted, that these cell lines were originally generated to assess the effect of the knock down of cgr-miR-204-3p. Indeed, due to the strong inhibitory effect observed in the 204-sponge cells, changes at the protein level (GFP fluorescence) could not be detected during earlier time points.

Negative correlations between miRNA expression and GFP protein (fluorescence) were observed in both, NC-sponge cells and 204-sponge cells, which indicates that probably, the decrease in GFP observed towards the end of the culture is due to the cells entering

a death phase (viability drops on day 10 for both cell lines) rather than due to the specific effect of miRNA binding. Regarding GFP mRNA transcript levels, despite being lower during the early stage of the culture for the 204-sponge cell line than for the NC-sponge cell line, these increased over time, reaching similar levels to those observed in the NC-sponge cells, without this being reflected on the measured fluorescence output, which always remained lower in the 204-sponge cell line than in the NC-sponge cell line.

Considering that cgr-miR-204-5p was found to increase over time, it is possible to hypothesize that the low levels of GFP protein detected in the 204-sponge expressing cells are maintained low despite more transcript been observed. It is also interesting to note that both cgr-miR-204-5p levels and GFP mRNA levels correlated strongly with VCD and cell viability in the 204-sponge cell line, whereas in the NC-sponge cell line, only miRNA expression correlated with this two parameters, and no clear pattern was observed for the levels of GFP mRNA transcript. It is counterintuitive to see more mRNA transcripts accumulating when being targeted by increasing amounts of miRNA. However, different studies have shown that miRNA-targeted transcripts accumulate in specialised cytosolic processing bodies (p-bodies). In these p-bodies, miRNA-targeted mRNA transcripts can accumulate without being degraded. It is possible to hypothesize that the overexpression of a heavily repressed mRNA transcript such as a sponge could lead to p-body saturation and explain the observed accumulation (Liu *et al.*, 2005; Parker and Sheth, 2007).

The results from the transient experiments also showed that miRNA sponges not only resulted in different levels of expression of the GFP transgene in CHO cells, but also that these levels of expression can be modulated depending on the number of binding sites, the amount of delivered/expressed transgene and the specific levels of endogenous miRNA present in the cell. The presence of a long synthetic 3'UTR, without any specific miRNA target sequence, had a small effect on transgene expression, but the addition of specific miRNA binding sites clearly reduced GFP levels in a sequence specific manner for the two tested miRNAs, cgr-miR-409-3p and cgr-miR-204-5p. The basal level of expression could also be tuned by varying the number of binding sites present on the 3'-UTR, both, for the synthetic NC-sponge as well as for the miRNA specific sponge plasmids. Increasing the number of binding sites had previously been reported as a strategy to increase the efficiency of miRNA inhibition. Ebert *et al.* already reported that

above 6 binding sites, the improvement on miRNA inhibition to be marginal but highlighted the possibility of an increased number of binding sites being beneficial in the case of stably integrated sponges, where expression levels would be lower than when transiently transfected (Ebert, Neilson and Sharp, 2007). Similar results are reported by Kluiver *et al.*, with 20 binding sites leading to marginal differences when compared with 6 binding sites sponges (Kluiver *et al.*, 2012).

In addition, the degree of expression reduction was further modulated depending on the degree of complementarity of the miRNA binding sites. Fully complementary binding sites were able to drive a stronger reduction of GFP signal than bulged binding sites. The original work on miRNA sponges by Ebert *et al* already mentions that bulged binding sites were introduced in miRNA sponges to strengthen their “sponge” effect (Ebert, Neilson and Sharp, 2007). Blocking the degradation of mRNA transcripts driven by perfect complementary miRNA binding sites by including mismatching base pairs around the 10/11th nucleotide of the binding site made it possible to maintain higher levels of sponge in the cell and obtain a stronger inhibition of miRNA function (Ebert, Neilson and Sharp, 2007; Kluiver *et al.*, 2012). In a scenario where the aim is to obtain control of transgene transcript levels, driving increased turnover of transgene mRNA transcripts by using perfect matching binding sites would be desirable in order to reduce miRNA function inhibition and reduce potential side effects. In fact, most of the work done so far on miRNA-based genetic circuits, reviewed in the introduction, reports perfect complementarity for the sensors. It was interesting to see that further repression could be obtained by changing intracellular concentrations of miRNA using miRNA mimics and that a saturation effect was only seen at very high concentration of mimics. However, it must be considered that transient expression levels are typically much higher than stable expression levels, which, additionally, largely depend on the integration loci of the transgene. As seen in the stable sponge work, an overly strong sensor leading to strong repression of the transgene would lose its ability to further react to increased miRNA concentrations. The same could happen in the reverse scenario in which decreasing levels of miRNA would not be sufficient to allow transgene de-repression leaving the system locked by default in a repressed state.

Besides showing that transgene expression modulation can be obtained in CHO, these results reinforce what other reports have also shown before. Factors such as the initial

gene dose, binding site number and the levels of miRNA in the cells need to be carefully balanced in order to achieve a particular switch behaviour.

As reviewed in the introduction of this thesis, there have been many studies capitalising on synthetic UTRs in order to successfully link miRNA and transgene expression. For example, Mansfield *et al.* used a miRNA sensor to demonstrate expression patterns of *Hox* genes by using miR-10 and miR-196 specific binding sites. These two microRNAs are located in the *Hox* clusters in mouse and their expression replicated the expression of the *Hox* genes (Mansfield *et al.*, 2004). Brown *et al.* also demonstrated how cell specific transgene suppression can be achieved in cells using synthetic UTRs targeted by cell type or lineage specific miRNAs (Brown *et al.*, 2006b, 2007). Based on this idea, several cell classifiers, with different levels of complexity have been developed and have already been reviewed in the introduction. In most cases if not all, miRNA dependent transgene control relies on the presence or absence of a particular miRNA, or in what is referred as miRNA activity in some publications, which often is related with their absolute level of expression in a particular cell type or tissue. To my knowledge, all these studies still rely on static miRNA signatures acting as input signal and driving a particular output from each particular miRNA-based circuit. The results obtained in the first part of this chapter indicate that moving from using static miRNA signatures to a situation where intracellular changes in miRNA expression result in changes in transgene expression will require finer strategies to be developed. Although big changes in miRNA expression could be simulated by using mimics, it is unlikely that such big changes will be found in real biological conditions. Ways around that would involve the use of several miRNAs showing modest changes of expression in particular conditions, but all together amounting to a sufficient change that could drive strong changes at the transcript/protein level. The use of more sophisticated circuits might also be something to test in the future, with designs involving several layers of regulation with different transactivators and repressors placed under the control of miRNAs. In addition, a recent report by Gam *et al.* reported a large-scale screen of miRNA sensors (620) and dissected the potential antagonistic or synergistic effects of adding binding sites for different miRNAs on a UTR. Their findings demonstrate that not only the number of binding sites but also their location are important factors that affect miRNA-directed repression. The authors develop an Antagonistic/Synergistic model to explain the behaviour of different types of multi-input miRNA sensors. In particular, Gam *et al.* report that while synergistic

effects are obtained from concatenating several binding sites for a particular miRNA in the UTR, different miRNAs acting on the same UTR show antagonistic effects. The authors also report synergistic effects occurring when binding sites are placed in different UTRs of the same transgene (3'-UTR and in the 5'-UTR) (Gam, Babb and Weiss, 2018). Capitalising on this improved knowledge of miRNA sensor behaviour and using these improved models, it might be possible to get closer to circuit designs that can efficiently react to lower changes in miRNA levels and thus be applied to effectively respond to endogenous changes. Another observation derived from the presented work is that results observed in transient expression systems should not always be expected to reflect results obtained in stably transfected cell lines. Therefore, being able to adjust transient expression conditions to mimic as much as possible expression conditions that can be obtained in stable will be essential in order to be able to transfer systems tested and integrated in stable cell lines engineered for particular applications. An alternative to transient testing would be to move to a high-throughput stable clone generation platform for testing of miRNA sensors. Some studies have recently advanced in that direction, with the development of rapid prototyping platform to test genetic circuits in mammalian cells at a large scale using Zinc Finger Nucleases targeted integration (Duportet *et al.*, 2014), the discovery of novel landing pads in CHO cells (Gaidukov *et al.*, 2018) and novel CHO endogenous promoters (Nguyen *et al.*, 2019) that when combined, could be used to obtain stable and adjustable expression necessary to be able to control sensor levels depending on the specific levels of particular miRNA sets.

Following the investigation of the use of sponges as transgene modulators, the potential of using temperature sensitive miRNAs to modulate transgene expression by means of specific miRNA-sponge vectors has been investigated. The identification of what we called temperature sensitive miRNA started with robust temperature experiment that allowed us to identify 3 miRNAs that showed differential expression upon shifting the culture temperature from 37°C to 31°C: cgr-miR-34a, that was up-regulated at 31°C compared with 37°C and cgr-miR-30e and cgr-miR-378, whose expression was shown to decrease at 31°C versus 37°C. Temperature sensitive miRNA in CHO have been investigated in the past, not to use them as temperature controllable molecules but based on the rational that their differential expression could be related to the phenotypes observed upon temperature shift such as prolonged cell viability and increased productivity. Gammell *et al.* were the first to report the investigation of

miRNA in the context of temperature shift by using cross-species miRNA arrays. However, in that study, cells in exponential growth phase were compared to cells in low temperature induced stationary growth phase and the differential expression for the identified miRNA, cgr-miR24 and cgr-miR-23, was related to growth phase rather than to temperature (Gammell *et al.*, 2007). A later study by Barron *et al.* identified 10 miRNAs using TLDA analysis (N. Barron *et al.*, 2011). Interestingly, among those, miR-30e was already identified in that study as differentially expressed upon temperature shift and was not pursued for validation. That study described miR-30e as being switch ON by temperature shift, with an increase in expression detected at 31°C compared with its undetectable levels at 37°C. Indeed, this is the opposite of what was found in the work presented here, where the downregulation of cgr-miR-30e upon temperature shift was validated and RNASeq data showed that miR-30e is actually highly expressed in CHO cells. This is another example of how careful one needs to be when choosing targets from previously published datasets, as highlighted in Chapter 1. The study from Barron *et al.* successfully validated the downregulation of miR-7 at 31°C by RT-qPCR, once more, highlighting the difficulty in obtaining consistent results when profiling miRNAs (N. Barron *et al.*, 2011). Another study aiming to characterized temperature sensitive miRNA in HeLa and CHO cells identified cgr-miR-483 as differentially expressed upon temperature shift using a microarray profiling platform (Emmerling *et al.*, 2016). It is interesting to note that in these studies, including this thesis, there were no common hits despite the temperature shift setup being very similar with cells grown for approximately two days at 37°C then shifted to 31°C for 24h. While it is true that there are differences in the profiling technologies, cell lines and culture systems used, it is intriguing to not have more common identifications. While it is entirely possible that the datasets overlap more than it would appear from the data disclosed in the publications, this could also mean that miRNA profiles are specific to cell line or culture conditions. If so, this would be an important drawback to the use of miRNAs as genetic switches, as it would imply specific contexts would require specific profiling and testing. This could be something interesting to further explore by profiling different cell lines in different culture systems aiming to identify or disprove common expression patterns linked to temperature.

Coming back to the work presented in this chapter, the next step was to design and construct plasmids expressing sponges specific to the identified miRNAs. As previously

described, the changes observed in protein expression upon temperature shift are the result of the effect of temperature at multiple stages from the transcription of the mRNA, the translation of the protein. Factors such as the transcriptional activity of the promoter, mRNA stability, or translation efficiency are modified by temperature. While not much could be done regarding transcript stability or translation efficiency, the effect of temperature on the promoter, an essential element of the expression vector, was considered. Different promoters (CMV, EF1a and UBC) were tested upon temperature shift in order to assess the temperature-dependent expression differences that could arise from using a different promoter independently of the miRNA-sponge. EF1a was chosen as it showed the least the promoter less prone to drive expression changes upon temperature. It must be noted that GFP expression from EF1a in the no-sponge plasmid was still increased by over 20-24% upon temperature shift. This up-regulation results from broad changes caused of temperature shift in cell metabolism, that include an overall increase of mRNA transcript stability, and reduced protein turnover.

Sponges of different sizes have been shown to allow different levels of transgene repression, hence long and short sponges specific for each individual temperature sensitive miRNA were constructed and cloned in the 3'-UTR of GFP in the EF1a-GFP plasmid. In addition to the sponges specific for single miRNAs, hybrid sponges containing a combination of binding sites for cgr-miR-30e and cgr-miR-378 were built in order to test potential cooperative effects from the two miRNAs that showed downregulation upon temperature shift. Testing the effect of temperature shift in a transient expression setup involved several challenges. Cell growth differences upon temperature shift had a clear impacting in the transient expression of GFP. As the amount transiently transfected plasmid present per cell is reduced by half upon cell division, cell populations dividing at different rate will show different plasmid dilution rates that will obviously affect the final amount of template present per cell and be reflected in the GFP measurements. In addition, it has been already mentioned that previously published and *in house* studies have shown changes in stability of RNA molecules as well as the total RNA content not higher levels of mRNA transcripts in cell grown at sub-physiological temperatures than in cells grown at 37°C as shows in previously published studies (Fox *et al.*, 2005; Masterton *et al.*, 2010). While levels of mRNA were not measured in the context of temperature, higher levels of GFP were consistently observed at 31°C versus 37°C. In several occasion through this thesis GFP has been

shown to affect cell viability and impair cell growth. The fact that GFP expression levels from NC-sponge vectors and from some of the sponges were quite different added up to the complexity of properly controlling for growth related factors affecting GFP expression in the tested system. Another factor to consider is that temperature shift could be causing expression changes of some of the targets for the selected miRNAs. As discussed later, changes in the expression of mRNA transcripts targeted by miRNA can affect how the miRNA regulates other transcripts by affecting the pool of available miRNA able to bind them. Last but not least, temperature shift most likely affects all the molecular machinery related to miRNA processing and miRNA-related post-transcriptional regulation (Biggar and Storey, 2014; Potla *et al.*, 2015). These temperature-induced changes could be affecting the interaction between miRNA sponges and their target miRNA, thus making the model miRNA up/transgene down a too simplistic model to explain the observed results.

To account for some of the factors mentioned above, two different normalization strategies were used in order to address this. It was anticipated that none of them would be able to correct for all the above-mentioned growth-related factors affecting. However, both EF1a signal and VCD normalizations were able to partially compensate the observed GFP expression up-regulation, presumably related growth-differences, leaving differences between NC and miRNA sponges to allow the side by side comparison of the effects of each sponge at 37°C versus 31°C. In the transient expression model, no large differences were observed between sponge-dependent GFP repression at 37°C or at 31°C except for the miR-378 sponges, that showed a marked increase in GFP repression ability at 31°C when compared with 37°C. Similar results were obtained in the stable populations expressing the different sponge plasmids. In these, the effects of temperature on sponge dependent GFP repression could be investigated in a more robust system, where most of the challenges previously mentioned would not be an issue. Having stably integrated copies of the different GFP expression plasmid removed the problems with plasmid dilution related to differences in cell growth. In addition, the lower levels of expression of GFP found in the stable pools were all in a range where no negative impact on cell viability was observed. Also, the fact that the analysed populations were mixed pools, with random integration of the transgene, meant the system was controlled for things like copy number and clonal variation. In this stable model, marked differences in sponge-dependent GFP repression

between 37°C and 31°C were also found for the 378-sponges. Additionally, under stable conditions, the long sponge for miR-221 also showed higher levels of GFP repression.

Altogether, these results are not in accordance with the original hypothesis that changes in miRNAs upon temperature shift could finely control transgene expression. According to the canonical miRNA-mediated repression model, changes in miRNA expression due to temperature were anticipated to have inverted effects on sponge-dependent GFP expression. In other words, an increase in miRNA expression following temperature shift should lead to a decrease of GFP expression due to more miRNA binding to the sponge. No correlation between miRNA expression changes and sponge dependent GFP repression could be observed that followed that logic. Several hypotheses could explain that as miRNA expression levels are not the only factor affecting miRNA-mediated regulation. Changes of the expression of mRNA targeted by a particular miRNA will also have an influence on the available levels of miRNA, which are more important than the absolute levels of miRNA. The regulation of gene expression by competing endogenous RNAs (ceRNA) has been recently proposed. While there is still controversy and more research needs to be done in this area, the ceRNA hypothesis assumes that specific RNAs can impair microRNA (miRNA) activity through sequestration, thereby up-regulating miRNA target gene expression (Thomson and Dinger, 2016). Following the inverse logic, a decrease in the levels of expression of mRNA targets for a given miRNA would result in higher levels of available miRNA that could then further repress other targets. The validity of this hypothesis could be tested by looking for an increase or a decrease in targets for the miRNAs tested in this work using transcriptomic differential expression data. Ideally, this should be performed with data obtained from the different stable populations to account for the effect that both sponge and selection with antibiotics can have in cells. Alternatively, data from other temperature shift transcriptomic studies could be used, accepting the limitations mentioned above.

Moving beyond the paradigm of temperature dependent control of transgene expression using miRNA-sponges, the results presented in this chapter provide some other interesting insights in the use of miRNA and miRNA-responsive elements as tools to obtain control of transgene expression. While Gam *et al.* have shown that miRNA sensor modulation of transgene expression when miRNA binding sites are placed in the same UTR follows an antagonistic model, where the most repressing sensors

determined the protein output obtained (Gam, Babb and Weiss, 2018). The results obtained here challenge that model by showing that placing binding sites for different miRNAs on the same 3'-UTR, stronger repression can be obtained in comparison with using each individual set of miRNA binding sites on its own. While here miRNA sponges were tested in stable expression conditions, as already discussed above, to date most (if not all) of the research looking at the use miRNA sensors to harness control of transgene expression has relied on transient expression of miRNA sensors and on the use of static miRNA signatures or synthetic miRNA mimics as inputs. While these approaches have proven very valuable to develop cell classifiers or cell computational devices (Xie *et al.*, 2011; Mohammadi, Beerenwinkel and Benenson, 2017), their application in the context of biopharmaceutical production is not clear yet.

4. Conclusion

In summary, by testing miRNA sponge sensors in CHO cells using two different expression platforms – expression from transiently transfected plasmids and stable random integration of sponge containing plasmids – transgene control by endogenous miRNAs using synthetic UTR in CHO cells was investigated and the use of synthetic miRNA mimics to manipulate transgene expression was demonstrated.

To date and to our knowledge, only two reports have been published in CHO reporting on attempts to bring miRNA-based control to practical applications. A 2006 report by Malphetes *et al.* showed that it was possible to achieve modulated levels of siRNA-tagged transgene by using a system where promoters and repressor were combined to be able to adjust the amount of transcript being transcribed and the amount of siRNA being expressed (Malphetes and Fussenegger, 2006). More recently, Jossé *et al.* reported the application of synthetic UTRs targeted by endogenous CHO miRNAs as a tool to repress the expression of a DHFR selection marker thus allowing the generation of antibody expressing CHO pools with lower amounts of MTX (Jossé, Zhang and Smales, 2018).

Regarding the use of temperature as switch to modulate the expression of a specific transgene, an experimental setting was established in which miRNA-sponge dependent transgene repression could be evaluated. miR-378 sponges were shown to change their repressive ability upon temperature shift. However, this did not correlate with the specific changes in miRNA expression observed upon temperature reduction. While

further investigation of the generated cell lines could help to unveil the source of this discrepancy, and it is likely that the systemic changes related to the effect of temperature in the metabolism of CHO cells are playing a role in the observed outcomes.

It is likely that in the near future (probably already happening now), genetic circuits, including miRNA-based genetic switches can be tested in cell lines with known precise integration in high-throughput fashion, allowing the testing of multiple circuit designs and most probably resulting in control systems such as the one proposed in this work. However, in order to be functional and predictable miRNA-based control system require very specific parameters and depend on individual cellular environments cell features, meaning their application in CHO cells and in an industrial biopharmaceutical context might not be practical. Finally, while miRNAs can be seen as the perfect regulatory molecule, easy to harness as a trigger, their endogenous levels and the magnitude of the changes observed will most likely be far lower than they should be to making them more suitable as fine tuning molecules rather to use them on their own as drastic ON/OFF switches.

CHAPTER 5

miRNA-toehold switches as novel tools to
modulate transgene expression

1. Introduction

In the previous chapters, the use of miRNA sponges as miRNA-responsive elements to modulate transgene expression was investigated. However, sponges are not the only miRNA responsive elements available and, as previously described and discussed in the introduction chapter, other systems in combination with miRNAs for controlled transgene regulation have been developed. The work described in this last chapter revolves around the development of a novel miRNA-responsive elements referred to as miRNA-toehold switch.

An RNA toehold switch consists of a sequence of ribonucleotides designed to fold forming a hairpin like secondary structure. In bacteria, when the ribosome binding site (RBS) is placed in this hairpin-like secondary structure, ribosome assembly is blocked, and translation cannot occur. Upon binding of a trigger RNA molecule, complementary to the toehold region, the secondary structure is disrupted, and translation can take place, resulting in an up regulation of the translation of the transgene. Green *et al.* have shown how these type of riboswitches could be used in bacteria to design multi-input circuits allowing controlled expression of different reporter genes by providing specific RNA inputs (Green *et al.*, 2014, 2017). Recently this system has been further developed into a toehold repressor module in which RNA inputs can lead to specific translation repression by also modulating secondary structure around the translation initiation site (Kim *et al.*, 2019).

In the context of miRNAs and their use to modulate transgene expression, the toehold switch concept, with a structured 5'UTR that can be targeted by a miRNA trigger to disrupt the hairpin like structure and promote translation, hence switching on transgene expression, would be very useful to add the ON function to the existing OFF function miRNA naturally already provide (Figure 59).

With that aim, the work presented in this chapter involved the design, construction and testing of a small library of miRNA-toehold switches, all of them having a similar basic toehold design. 5'-UTR synthetic sequences were built containing a miRNA-binding site upstream of a Kozak-start codon sequence, followed by a downstream nucleotide sequence partially complementary to a portion of the miRNA binding site.

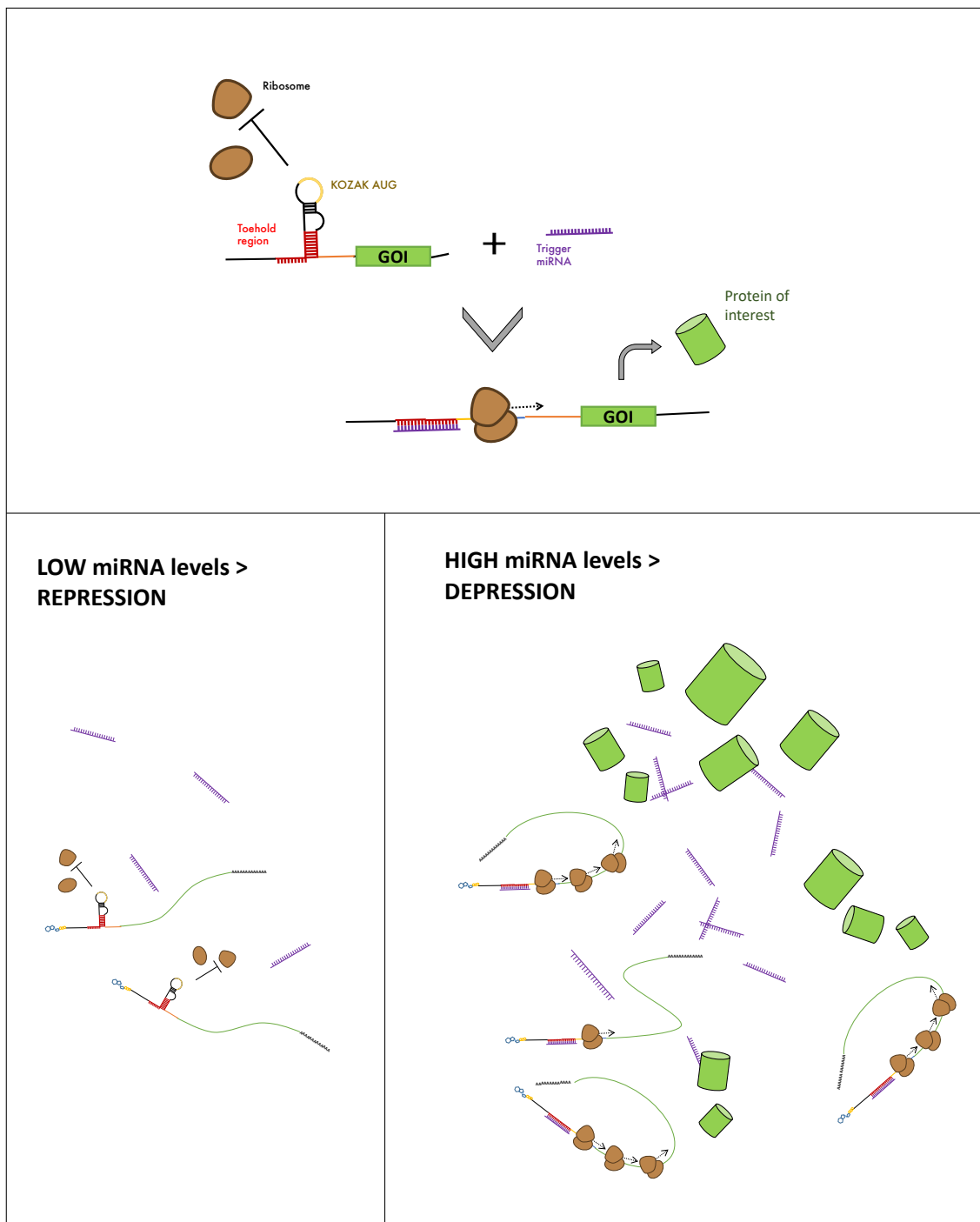


Figure 59 - Diagram of the theoretical mechanism of transgene repression/derepression by a miRNA-toehold switch. The presence of the hairpin-like structure blocks the assembly of the ribosome to around the Kozak-start codon sequence, hence impairing translation. Upon binding of the miRNA-trigger, the secondary structure around the Kozak sequence is disrupted allowing the formation of the full translation capable ribosome and allowing translation to proceed. Under low miRNA conditions, most of the transcripts would remain in a repressed state with low proteins amount being produced, while under upon and increase in miRNA expression, expression of the protein of interest would be switched ON.

This design should allow the 5'UTR to fold forming a hairpin like structure that blocks translation. Following the same principle as for the toehold switches, the binding of a specific miRNA would potentially force the unfolding of the UTR, thus allowing translation initiation and switching ON transgene expression. Combining this system with miRNA changes in response to particular conditions, would allow to have a system in which increase intracellular miRNA concentration would be grant the ability to miRNAs to up-regulate protein expression, as opposed to their canonical repressive natural function.

2. Results

2.1. Design and evaluation of new miRNA-based switches: miRNA-toehold switches

2.1.1. Switch-Backbone (SW-BB) Vector Cloning

All vectors were built based on the in-house expression vector N44, which contains a CMV promoter driving the expression of a d2eGFP reporter gene and a non-palindromic *KflI* (*SanDI*) restriction site downstream of the GFP coding sequence. This vector has been routinely used in the lab for the cloning of sponge vectors.

In order to be able to incorporate the toehold-switch sequences upstream of the reporter d2eGFP, a new *SgsI* (*AsclI*) restriction site had to be incorporated between that gene and the CMV promoter. In addition, as the toehold-switch sequence must contain a Kozak initiation sequence/ start codon (ATG) cassette as part of the loop, the original start codon of the d2eGFP had to be removed.

The d2eGFP gene from N44 was amplified by conventional PCR using a set of primers in which the forward primer included a 5' non complementary tail containing the restriction sites *KpnI* and *SgsI* in that order. The reverse primer was complementary to the 3' end of the d2eGFP and was overlapping with the *KflI* site. This way, *KpnI* and *KflI* could be used to re-insert the amplified modified d2eGFP fragment into N44, in order to obtain the backbone vector SW-BB, which now includes a unique pair of restrictions sites *KpnI-SgsI*, as an insertion point for the toehold switch sequences, and no Kozak-start codon on the d2eGFP gene.

2.1.1.2. Design and cloning of miRNA toehold switch vectors

In order to be able to use miRNAs as RNA trigger molecules and express the constructs in mammalian cells, various changes had to be made to the original *E. coli* toehold switch design.

First, the ribosome-binding site and start codon region had to be replaced by its mammalian equivalent, namely, a *Kozak*-start codon sequence. Second, the toehold and stem region had to be designed to be complementary to the desired miRNA.

The three pairs of different miRNA-toehold switch sequences were designed *in silico* as explained above.

The best way to obtain such oligo-duplexes was to order the corresponding forward and reverse DNA oligonucleotides for each specific miRNA-toehold switch construct and proceed to use the oligo annealing protocol to reconstitute the duplex DNA fragments before cloning. To avoid having to digest the annealed oligos with KpnI/SgsI restriction enzymes prior to cloning them into the backbone vector, both forward and reverse DNA oligos for each miRNA-switch were designed in a way that the corresponding 3'- KpnI overhanging-end and 5'- SgsI overhanging-end would be formed after the annealing step.

For the proof of concept, three different miRNA sequences were used: a synthetic AT-rich Negative Control (NC) miRNA sequence, routinely used in house as negative control for miRNA sponge experiments, cgr-miR-204 and cgr-miR-409-3p (Table 16).

Table 16 – Mature miRNA sequences for the microRNAs chosen for the design of proof of concept miRNA-toehold switches. *miR-NC is a negative control sequence predicted to show low binding to endogenous miRNAs in CHO.*

miRNA	Sequence
miR-NC	UGUCAGCUUUCUGAAAACUU
cgr-miR-204	UUCCCUUUGUCAUCCUAUGCCU
cgr-miR-409-3p	GAAUGUUGCUCGGUGAACCCCU

NC-L-1G
GGTACCAGCTTTTCAGAAA**GCTA**ACA----GCCGCCTTTT**TTCGCCACCATG**ACCACCGTC--ACACACACACACACCGGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

NC-H-1G
GGTACCAGCTTTTCAGAAA**GCTA**ACA----GCCGCCTTTT**TTCGCCACCATG**GCGGCTGTGGC-TGT**TAGCT**TTTCTGGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

204-L-1G
GGTACCAGGCATAGGATGACAAAGGGAA---AGGAGGAAAAAA**GCCACCATG**GCGGCTGTC--GCACAAAGGGAAGGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

204-H-1G
GGTACCAGGCATAGGATGACAAAGGGAA---GCCGCCTTTT**TTCGCCACCATG**GCGGCTGTC--TTCCCTTTGTCAGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

409-L-1G
GGTACCAGGGGTTCAACGAGCAACATTC---AGGAGT**TTCGCCACCATG**GCGGCTGTC--CGAGCAACATTCGGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

409-H-1G
GGTACCAGGGGTTCAACGAGCAACATTC---GCCGCCTTTT**TTCGCCACCATG**GCGGCTGTC--GAATGTTGTCGGGGGGGGGGCGCGCC
ကလေးစုစုကလေးစုစု

Figure 60 – Sequences of the first generation of miRNA-toehold 5'-UTR. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlight in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

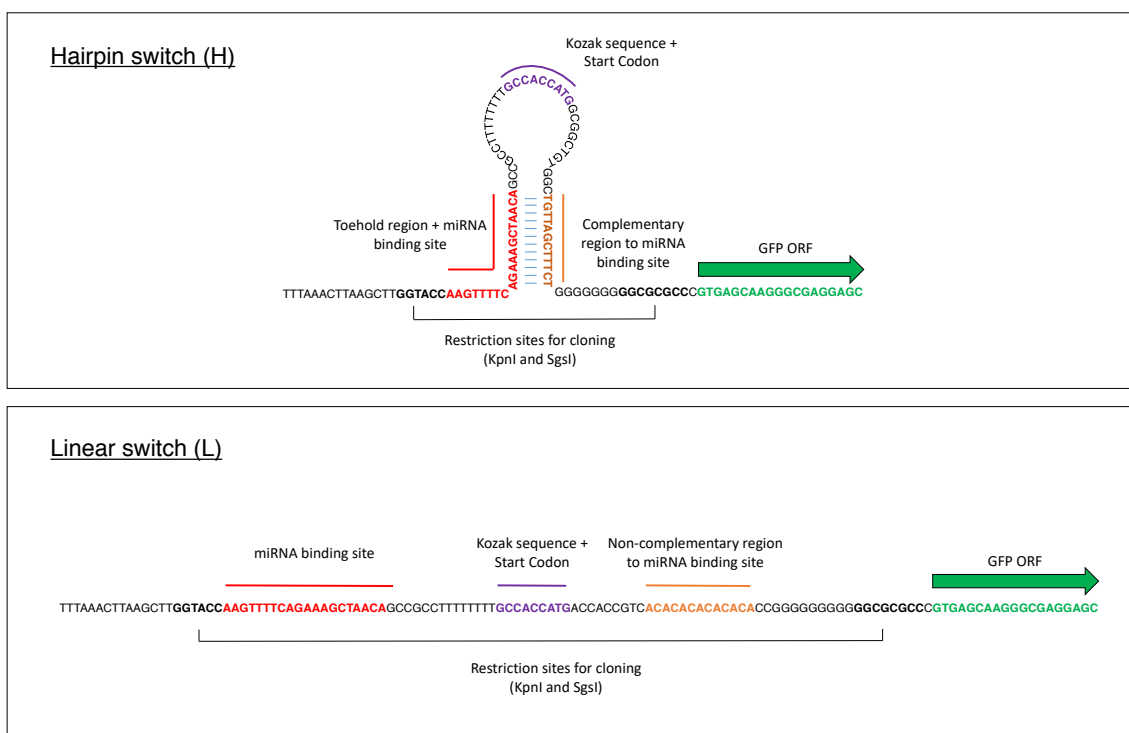


Figure 61 – Sequence design of the hairpin (H) and linear (L) miRNA-toehold switch constructs .The design of a hairpin toehold switch (H), included a sequence downstream of the Kozak-AUG codon is made complementary to the miRNA-binding site placed upstream of it, allowing the formation of a hairpin like structure (top panel). As a control sequence without hairpin, a linear (L) toehold switch sequence as designed by swapping the complementary region by a non-complementary region, with the aim to impair the formation of the hairpin (bottom panel).

For each miRNA sequence, two different switches were designed (Figure 60). First, what from here on will be referred as hairpin switches (H). These hairpin modules were designed based on the toehold switch design original published by Green *et al.* (2014). In short, a Kozak-Start codon sequence (GCCACCAUG) was placed downstream of a miRNA-binding site. Downstream of the Kozak-Start codon sequence a complementary sequence to the last 10 nucleotides of the miRNA binding site was added. This resulted in a sequence able to form a hairpin-like secondary structure (Figure 61 – top).

Considering the miRNA-toehold switches were to be cloned in the above described backbone, it was anticipated that the original GFP expression vector (N44) with its short 34 nucleotides 5'-UTR would not be the appropriate control. Therefore, for each miRNA-sequence a linear (L) switch was also designed by replacing the 3' end sequence complementary to the miRNA binding site with a non-complementary sequence (Figure 61– bottom). This so-called linear construct should allow the transcription of an mRNA without the 5' hairpin structure but similar in length and much of the sequence to the hairpin construct for each sequence.

2.2. miRNA-Toehold-Switch *in silico* characterisation

Before proceeding to clone each individual miRNA-toehold plasmid in the lab, each individual miRNA-toehold switch sequence was analysed *in silico* to assess GC content and thermodynamics for hairpin formation. The short UTR of the backbone plasmid N44 was also analysed for comparison. GC content is associated with a higher degree of secondary structure in transcripts and having toehold sequences with large differences in GC content could lead the formation of secondary structures other than the desired hairpin.

No major differences in GC content were found between each sequence for the 6 constructs (Figure 62 – A) but all had slightly lower GC content than the short UTR found in the N44 plasmid. GC content was between 52 and 61%, with the NC-H/NC-L pair showing the lowest GC content and the 409-H/409-L pair showing the highest GC content.

Larger differences were observed when looking at the GC content of the miRNA-binding site sequence (Figure 62 – B). Values between 30 and 54% were observed. Again, the NC-H/NC-L pair showed the lowest GC content and the 409-H/409-L pair showed the

highest GC content. However, as each pair of hairpin and linear construct for each sequence would have the same miRNA binding site, these differences would only be considered when comparing different toeholds with each other and would not affect comparisons made between hairpin and their linear control constructs for each sequence.

Considering the only different regions for each switch pair (hairpin/linear) are the complementary/non-complementary regions downstream of the Kozak-ATG, their GC content for was also calculated to account for potential differences that would need to be considered (Figure 62 - C). Differences in GC content were observed for the switch pairs NC and 204, with higher GC content in the linear switches than in the hairpin switches.

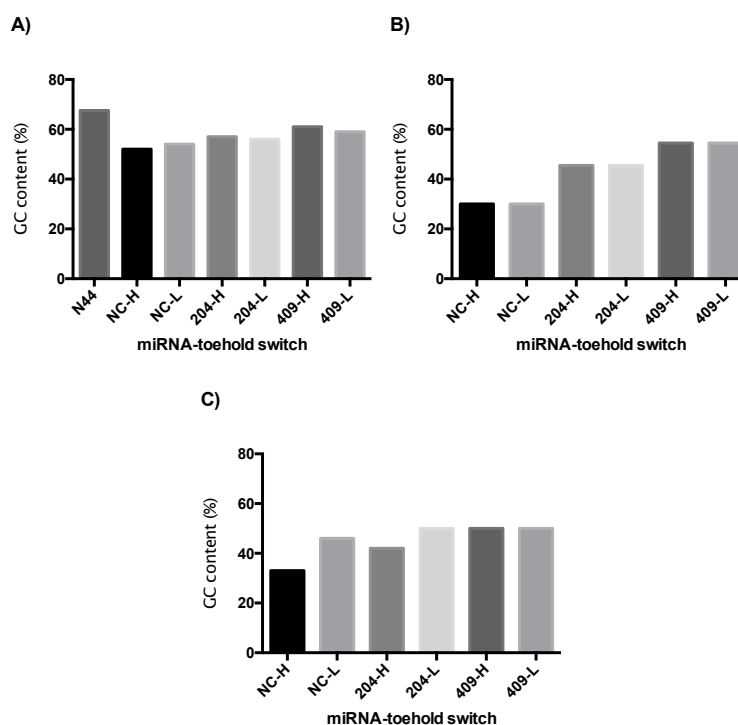


Figure 62- In silico analysis of GC content for each of the miRNA-toehold switch sequences. A) GC content of the full-length miRNA-toehold switch module. B) GC content of the miRNA-binding site upstream of the Kozak-Start Codon sequence. C) GC content of the complementary/non-complementary sequence downstream of the Kozak-Start Codon sequence.

In terms of the thermodynamic calculations for secondary structure formation, the three hairpin constructs showed lower ΔG values than the linear constructs, when analysed with two different prediction algorithms, indicating these regions are thermodynamically more stable, which is a good indication of their ability to form a

hairpin-like structure. Higher $-\Delta G$ values were obtained for the linear constructs with slightly different results depending on the software used for the analysis when comparing with the N44 plasmid short UTR (Figure 63). These 6 constructs will be referred to as 1st generation toeholds (1G).

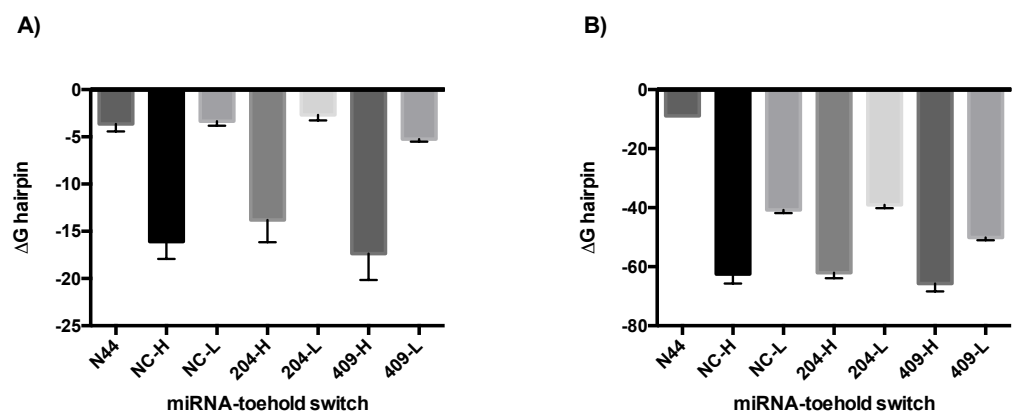


Figure 63 - In silico analysis of the thermodynamics for secondary structure formation for each of the miRNA-toehold switch sequences. A) Thermodynamics for hairpin formation (IDT Oligo Analyser - [Link](#)). B) Thermodynamics for hairpin formation (Predict a Secondary Structure Web Server – University of Rochester -[Link](#))

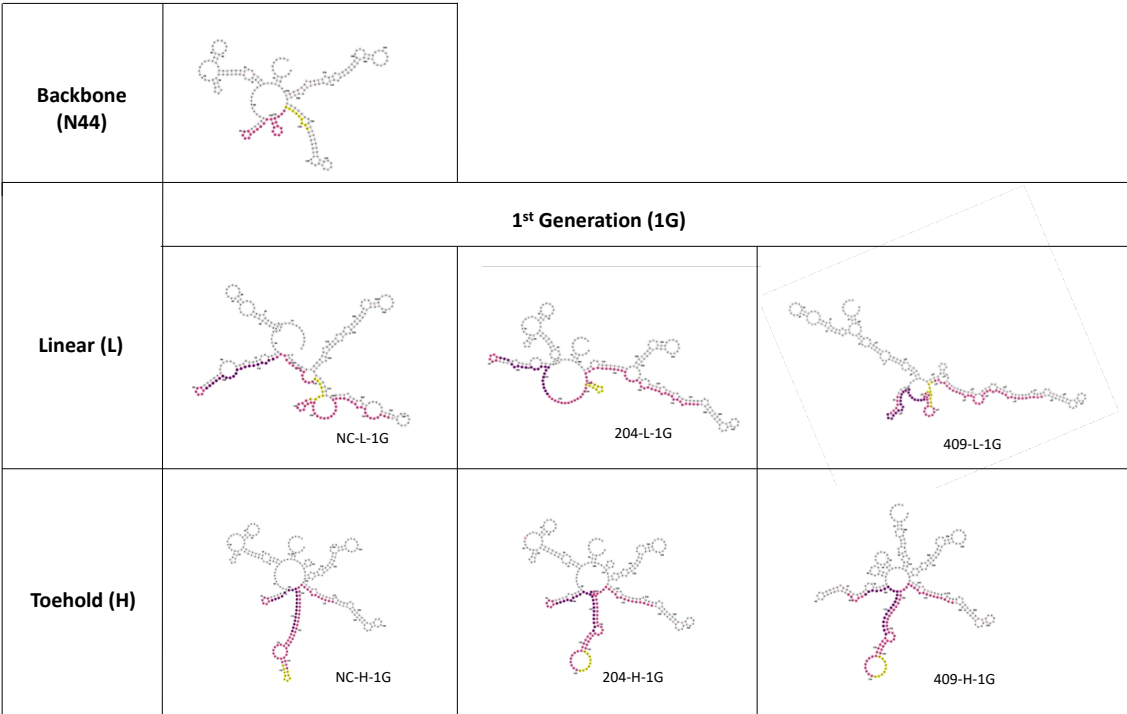


Figure 64 - Predicted secondary structures using RNA fold for the different miRNA-toehold switches (linear and toehold versions) as well as for the GFP backbone plasmid (N44). Predictions include the 5'UTR-region and ca. first 100 nucleotides of the gene (GFP) CDS.

In summary, the structure analysis performed for the different sequence designs indicated that the desired hairpin-like secondary structures around the Kozak-ATG should be obtained for the toehold vectors but not for the linear constructs. Oligos for those sequences were ordered, annealed and cloned in the SW-backbone as previously described. The obtained plasmids were then transfected in CHO cells to assess the effect of the different miRNA-toehold switches on GFP expression.

2.3. Characterisation of miRNA-toehold switch vectors

In order to assess the expression of the GFP reporter gene from the different miR-toehold switches, transient transfection experiments were performed. In addition to the miRNA-toehold hairpin and linear vectors, different controls were used: the empty backbone vector (SW-BB), the original N44 vector as a positive control, and H₂O as no DNA negative control. Four different CHO cell lines, parental CHO-K1, CHO-S, CHO-DG44 and an IgG producer CHO-DP12 were transfected with 500 µg DNA/10⁶ cells in 24-well plates. GFP expression was measured after 24- and 48-hours using flow cytometry.

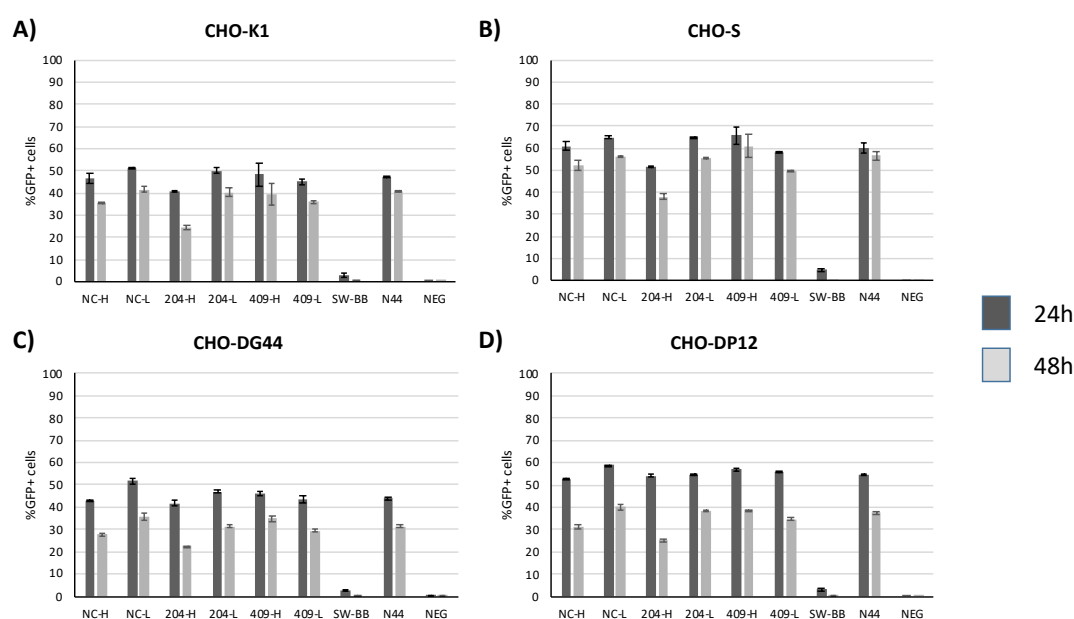


Figure 65 – Percentage of GFP positive (%GFP+) after transfection of the different miRNA-toehold switch plasmids in different cell lines. The data represents the average of the data obtained in measurements performed at 24 and 48 hours after transfection of the different cell lines A) CHO-K1 B) CHO-S C) CHO-DG44 D) CHO-DP12. Data for each cell line was normalized using the N44 values as a reference. Each bar represents the average of 2 biological replicate transfections and error bars represent their standard deviation.

Transfection efficiency for each cell line was estimated from the percentage of GFP positive cells obtained from transfection with the N44 vector (Figure 65). No significant differences were observed between vectors in each cell line. However, the CHO-S cell line exhibited slightly better transfection efficiencies than the other 3 cell lines independent of the vector. The no-ATG SW-backbone vector did not express any GFP and only residual background fluorescence was detected.

As in previous chapters, to account for differences in %GFP+ cells between populations GFP expression will be discussed based on the mean fluorescence intensity for all recorded events (Figure 66). Overall, different levels of GFP expression were observed depending on the host cell line. CHO-DP12 showed the highest MFI values, followed by CHO-S and CHO-DG44. CHO-K1 cells showed the lowest MFIs values.

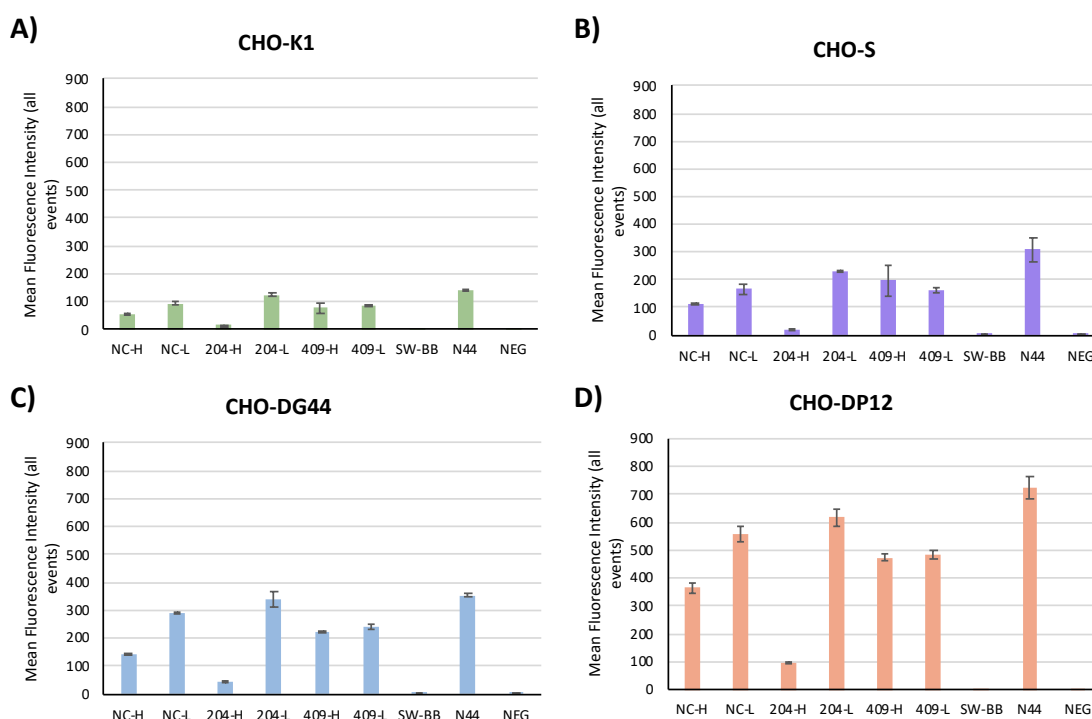


Figure 66 – GFP Mean Fluorescence Intensity (all events) measured 48h after transfection for the different transient miRNA-toehold switch pools. A) CHO-K1 B) CHO-S C) CHO- DG44 D) CHO-DP12. Each bar represents the average of 2 independent transfections and error bars represent the standard deviation.

Expression patterns obtained from the different plasmids were very similar across cell lines. For the NC and 204 sequences, H constructs showed lower GFP expression than L constructs. That was not observed for the 409 sequence. When looking at the differences in GFP expression between each individual plasmid and the N44 positive

control, it can be seen that except for 204-L, the presence of the 5'-UTR switch region caused a reduction in GFP expression when compared with N44 (Table 17 vs N44).

Table 17 - Summary of the data from transient transfection of the different miRNA-toehold switch construct on different cell lines. Values for the changes in GFP expression were calculated for each vector versus N44. Changes between H and L vectors are also shown. P-values are provided for the significance of the differences observed for each vector versus the N44 control (P-value vs N44) and for each pair of hairpin and linear vectors, for each cell line (P-value - H vs L). Differences were evaluated using a two-tailed homoscedastic Student T-test. Differences with p-values < 0.05 were deemed significant and are highlighted in red.

Cell line	Vector	GFP change (%)			
		vs N44	P-Value (vs N44)	H vs L	P-Value (H vs L)
CHO-K1	NC	-53%	0.03	-38%	0.06
	NC-NF	-33%	0.01	-	-
	204	-84%	0.02	-89%	0.00
	204-NF	-8%	0.20	-	-
	409	-38%	0.06	-9%	0.36
	409-NF	-30%	0.01	-	-
	N44	0%	1.00	-	-
CHO-S	NC	-59%	0.08	-31%	0.13
	NC-NF	-39%	0.07	-	-
	204	-90%	0.05	-91%	0.01
	204-NF	-16%	0.28	-	-
	409	-32%	0.14	23%	0.65
	409-NF	-39%	0.10	-	-
	N44	0%	1.00	-	-
CHO-DG44	NC	-52%	0.01	-51%	0.04
	NC-NF	-18%	0.04	-	-
	204	-83%	0.00	-87%	0.04
	204-NF	-2%	0.77	-	-
	409	-35%	0.01	-7%	0.07
	409-NF	-23%	0.01	-	-
	N44	0%	1.00	-	-
CHO-DP12	NC	-46%	0.05	-34%	0.02
	NC-NF	-23%	0.08	-	-
	204	-83%	0.03	-84%	0.03
	204-NF	-14%	0.14	-	-
	409	-35%	0.07	-2%	0.04
	409-NF	-26%	0.09	-	-
	N44	0%	1.00	-	-

This was observed in all cell lines, although surprisingly, differences in the CHO-S were statistically not significant. 204-H had the largest reduction with 83-90% lower MFIs than N44. The NC-H switch resulted in a 46-59% reduction in GFP expression when compared with N44. Comparing GFP expression between NC-H and NC-L, the presence of a hairpin-like structure caused a 2-fold reduction in GFP expression assuming not having to consider for miRNA binding. The 204 H/L pair was the one that exhibited the largest differences between H and L constructs and the one that showed a significant difference between its hairpin and linear versions across the four different cell lines. For the 409 H/L, lesser differences were observed, ranging from a 7% to a 23% reduction in GFP signal in the H construct versus the L construct, only significant for the CHO.DG44 cell line.

In summary, these results showed that including a miRNA toehold switch comprising a structured miRNA binding site region around the Kozak-ATG sequence upstream of the CDS of GFP affected the expression of GFP at the protein level (fluorescence output) and to varying extents depending on the tested sequences. These results also indicated that the presence of the hairpin-like 5'-UTR leads to a stronger decrease of GFP expression in comparison with a linear 5'-UTR (which in NC and 409 lead to a slight decrease when compared with the short 5'-UTR in the N44 control) for 2 of the 3 tested pairs of plasmids (NC and 204 sequences). To further investigate this 1st generation of miRNA-toehold switches, stable pools were generated from the transiently transfected cells.

2.4. Characterisation of miR-toehold switch driven GFP expression in stable pools.

In order to further assess the change in GFP expression observed from the different switch constructs, this time under stable expression conditions, stable cell lines were generated. To do so, transfected cells were transferred to tissue culture plates and grown in adherent mode under selective pressure (Hygromycin). Different cell lines exhibited different sensitivities to the selection antibiotic Hygromycin B, resulting in differences in the percentage of GFP+ cells in the obtained stable mixed populations. However, homogeneity was observed among cell lines derived from a single “parental” cell line (Figure 67). Interestingly, after selection, the different cell pools continued to show drastically different levels of GFP expression as observed in the transient pools (Figure 68 and Table 18). Across the four cell lines tested, constructs predicted to form hairpins exhibited reduced GFP expression in comparison with their linear counterparts

for the NC-H/NC-L and 204-H/204-L pairs. Comparing the different linear versions of the switch constructs for these two sequences to the original N44 vector, the observed effect was cell line dependent. While for CHO-K1, a substantial reduction of GFP was observed in the L constructs compared with the N44 control, in CHO-S and CHO-DG44 cells this difference was smaller, and no difference was observed between L constructs and N44 in the CHO-DP12 cells. The same expression pattern was not observed for 409-H/409-L vectors. 409-H and L constructs did not show differences in GFP expression (CHO-K1, CHO-DG44) or resulted in an inverted situation with the L construct expressing less GFP than the H construct (CHO-S and CHO-DP12).

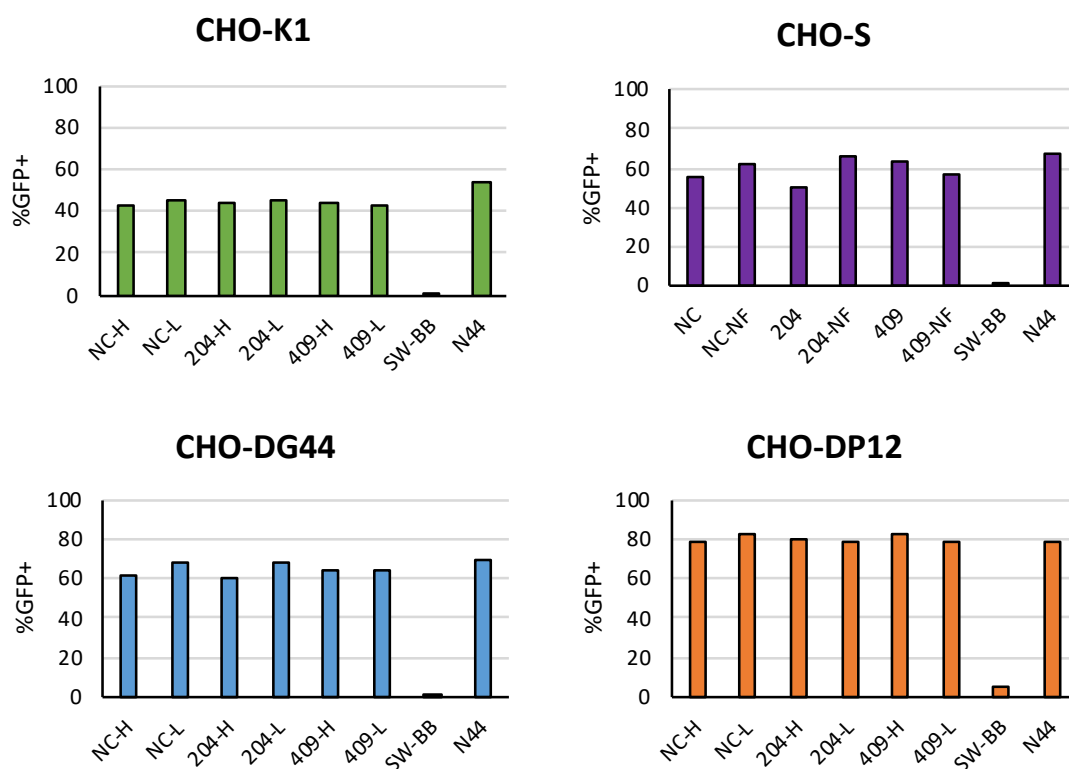


Figure 67 - Percentage of GFP positive (%GFP+) after selection of the different miRNA-toehold switch containing cell lines: A) CHO-K1 B) CHO-S C) CHO- DG44 D) CHO-DP12. Individual gates were set for each cell line taking into account maximum fluorescence measured on the positive control N44 cell line, to ensure all measurement were performed within the detection limits of the instrument. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation.

Table 18 - Summary of the %GFP in the obtained stable mixed populations after selection for each parental cell line.

%GFP+ cells	CHO-K1	CHO-S	CHO-DG44	CHO-DP12
AVERAGE	45.3	60.2	65.2	79.8
ST.DEV.	3.7	6.1	3.3	2.0

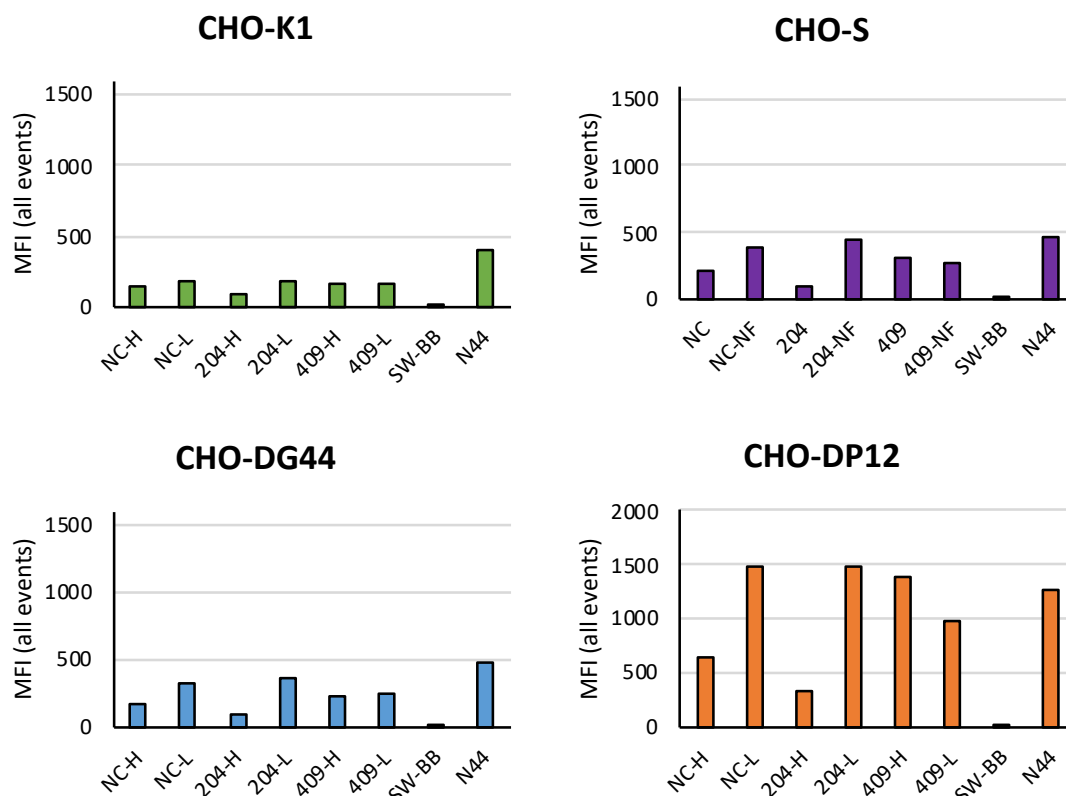


Figure 68 - GFP expression (MFI – all events) after selection of the different miRNA-toehold switch containing cell lines: A) CHO-K1 B) CHO-S C) CHO- DG44 D) CHO-DP12. Each bar represents the average of 3 technical replicate GFP measurements.

In order to assess how well the transient expression results predict the outcome of the generation of stable cell lines, we analysed both data sets using linear regression, for the four cell lines together and for each individual cell line in Figure 69. For each three of the for tested cell lines, CHO-S, CHO-DG44 and CHO-DP12, the transient expression experiments predicted the behaviour of the different constructs tested in stable cell pools.

In summary, similar expression patterns could be obtained for the tested miRNA toehold switches in transient transfections and stable mixed pools. Once more, in the tested

stable mixed pools, results from the NC H/L pair of switches, which are almost identical in terms of sequence but differ substantially in predicted secondary structure, indicate that GFP expression might be affected by these differences in structure, with hairpin-like switches reducing GFP expression.

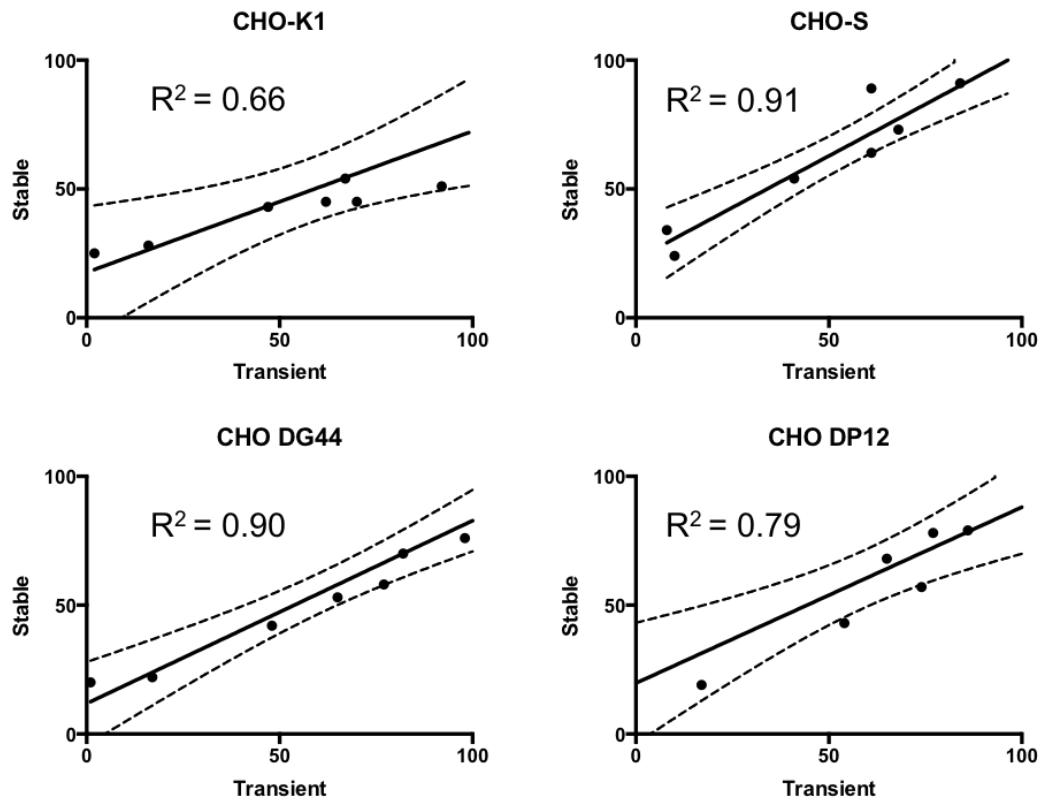


Figure 69 –Linear regression analysis assessing the ability of transient expression experiments to predict the behaviour of the different constructs under stable expression conditions in adherent culture, for the different tested cell lines: A) CHO-K1 B) CHO-S C) CHO-DG44 D) CHO-DP12.

So far in both transient and stable systems, lower GFP expression was observed for the toehold constructs predicted to form hairpin-like secondary structure in comparison with constructs designed to maintain a “linear” secondary structure. In order to be able to further investigate GFP expression from these different constructs in suspension serum-free conditions, the generated stable mixed pools, originally generated in adherent cultures, were transferred to 50 mL Tubespin® bioreactors and serum-free media (SFM) and GFP was monitored over time to track potential changes in expression during the suspension adaptation process.

2.5. Adaptation to miRNA-toehold stable pools to suspension growth

In a step further to assess the functionality of the switch constructs in suspension culture conditions, cells were further adapted to grow in 5-mL cultures in Tubespin® bioreactors in the shaken incubator. DP12 derived cell lines were chosen due to their high levels of GFP+ cells (>80%) after selection in adherent culture (Figure 70). Data for the 409-cell lines is not shown as these cells were not used due to the absence of differences between H and L behaviour in both transient and stable experiments.

During this adaptation no replicates were performed and therefore, only observational trends are described. Transferring cells to suspension caused a drop in the percentage of GFP+ cells to values >60% (From the >80% in adherent culture). Over the >2-weeks adaptation period to suspension growth, the different populations maintained a stable percentage of GFP+ around 60-70%. However, in terms of fluorescence intensity, no differences between the levels of GFP expression in the DP12-NC-H / DP12-NC-L cell lines were observed. The DP12-204-H and DP12-204-L cell lines maintained their different levels of expression, with DP12-204-H, which contains the hairpin switch, exhibiting a reduced level of expression up to 64% in comparison with the cell line containing the linear version of the switch.

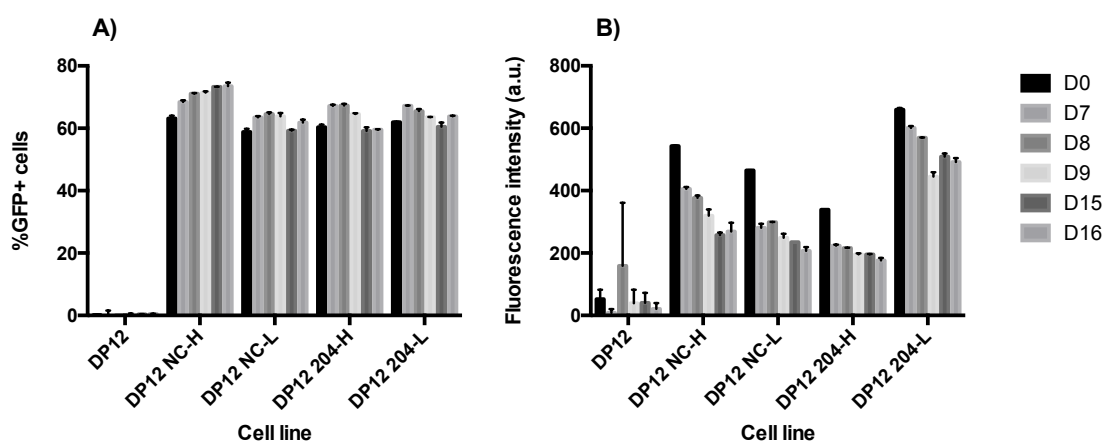


Figure 70 –A) Percentage of GFP+ cells and B) GFP Fluorescence intensity for the different DP12-derived stable pools over the >2 weeks adaptation process to suspension culture. Each bar represents the average of 3 technical replicate fluorescence measurements and error bars represent their standard deviation.

After 3 weeks, stable pools were banked and used in further experiments.

2.6. Evaluation of miR-toehold constructs in stable mixed pools

In order to evaluate the 'responsiveness' of the miR-toehold constructs to the presence of their complementary miRNAs, synthetic miRNA mimics were used to transiently transfect the stable DP12 stable pools expressing the different miRNA-toeholds (NC H/L and 204 H/L).

A proof of concept experiment in order to evaluate the responsiveness of the miR-toehold switches and the specificity of the miRNA to a particular miR-toehold switch was performed. The four stable cell lines containing miR-toehold switches in both hairpin and linear forms, so DP12-NC-H, DP12-NC-L, DP12-204-H and DP12-204-L were transfected with: miR-NC-mimic, and miR-204-mimic.

In addition to measuring GFP expression (fluorescence), samples were collected for total RNA extraction and mRNA transcript levels of GFP, as well as miR-204 levels were evaluated by RT-qPCR.

All transfections with the eGFP-siRNA positive control showed a reduction in the Mean Fluorescence Intensity (MFI) and percentage of GFP positive cells confirming the siRNA transfection protocol worked well (Figure 72 and Figure 73 - siGFP). In addition, for all cell lines, transfecting cells with miRNA mimics not specific to their miR-toehold switch did not have an effect on GFP expression at the mRNA transcript level (Figure 72 and Figure 73 - B). However, at protein level, an increase of the GFP signal was observed in cells transfected with non-specific mimic (Figure 72 and Figure 73 - A). As expected, transfections using a miRNA-mimic specific for the miR-toehold switch contained in a given cell line caused a reduction of the MFI and of GFP transcript levels, independently of the miR-toehold switch being a hairpin or a linear design.

Comparing untransfected samples, it can be observed that GFP mRNA transcript abundances were different between cell lines containing hairpin-forming miR-toehold (DP12-X-H) and cell lines containing the linear version of the miR-toehold (DP12-X-L), which showed higher levels of GFP mRNA (Figure 71). Interestingly and surprisingly, lower levels of GFP mRNA (higher Ct values) were detected in NC pools when compared with 204 pools. This was unexpected as NC transcripts should not be targeted by any endogenous miRNA whereas 204 constructs should be targeted by endogenous miRNA cgr-miR-204-3p and therefore lower levels of GFP would be expected.

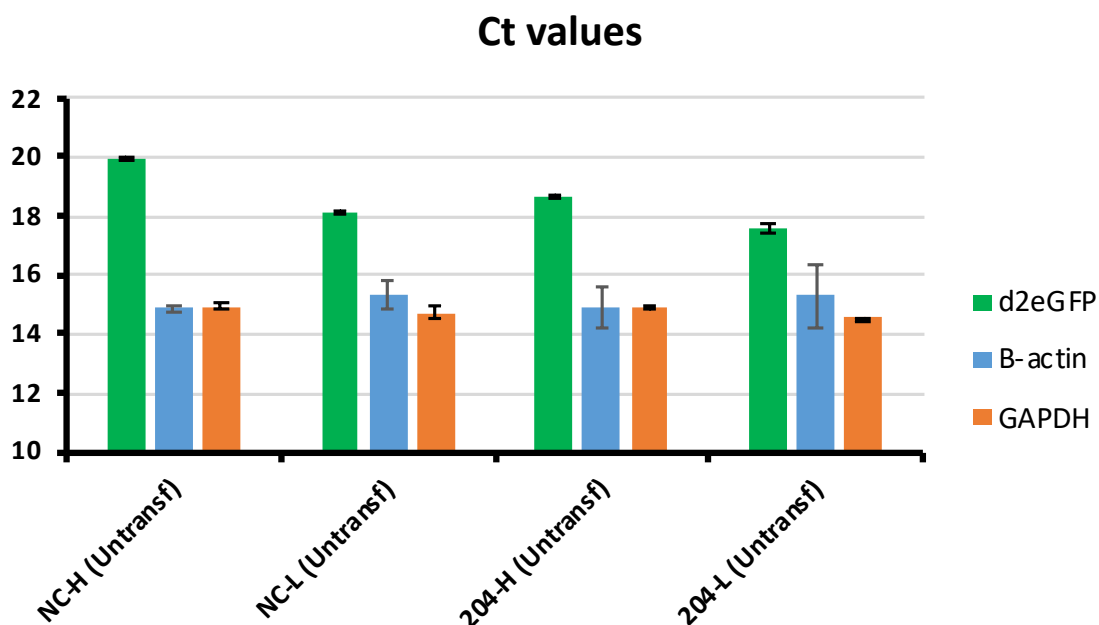


Figure 71 – GFP expression (CT values) for the NC-H, NC-L, 204-H and 204-L pools untransfected with mimics. Endogenous B-actin and GAPDH are shown as controls. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation.

Looking at the effect of miRNA mimics on the different pools, for both sequences, NC and 204, no differences were observed between the pools transfected with the non-specific miRNA mimics when compared with untransfected cells. However, in cell lines transfected with specific miRNA mimics complementary to the miR-toehold switch, a downregulation of the GFP mRNA was observed (Figure 72 and Figure 73 - B). Both NC and 204 hairpin-like toeholds (H) showed lower levels of GFP transcript when compared with the linear toeholds (L) counterpart.

In addition, the magnitude differences observed at the mRNA level did not match the differences observed at the protein level (Figure 72 and Figure 73 – A versus B for untransfected samples). While a 3-fold increase amount of GFP mRNA was observed between NC-H and NC-L, similar levels of GFP fluorescence were detected for the two pools. Regarding the 204 pools, a similar situation was observed. While GFP mRNA levels were approximately 2-fold higher in 204-L versus 204-H, GFP protein was almost 6-fold higher in 204-L when compared with 204-H.

A second experiment to evaluate if the observed knockdown effect would be dependent on the amount of miRNA delivered was performed. Only the DP12-204 pools, were tested taking into account that the pair 204-H/204-L was the one consistently showing

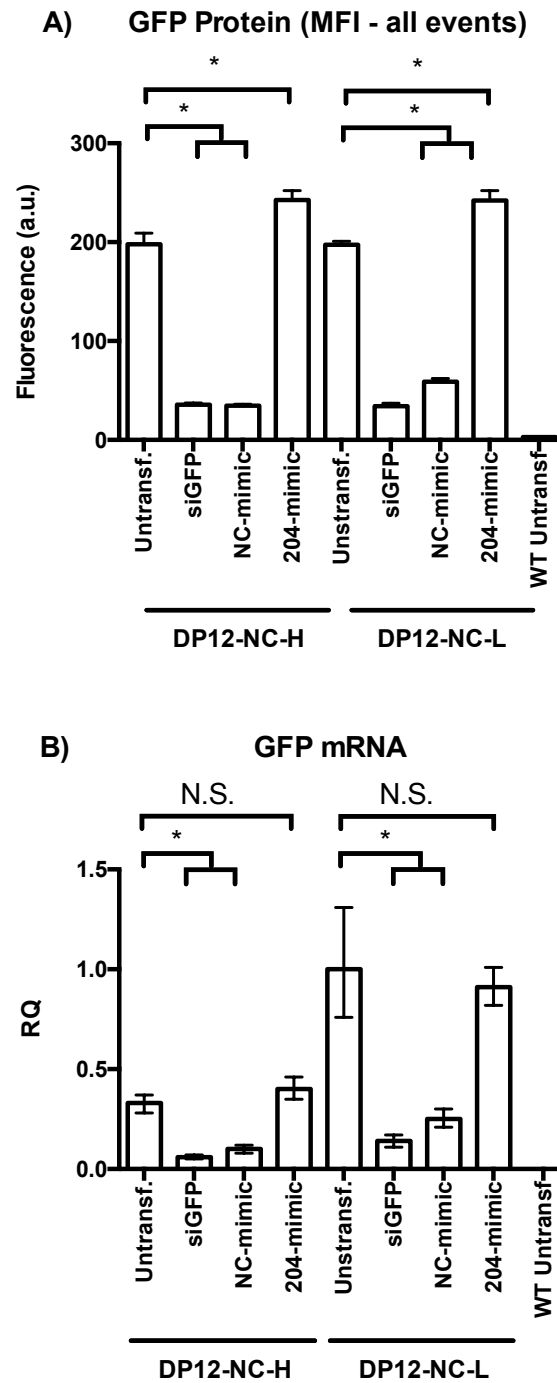


Figure 72 – GFP expression for the miR-toehold transgenic cell lines DP12-NC-H, DP12-NC-L after transfection with miRNA-mimics specific for the NC sequence (NC-mimic) and the miR-204 sequence (204-mimic). A – Mean Fluorescence Intensity of all events cells and B – GFP mRNA transcript relative quantification. A GFP specific siRNA (siGFP) was used as a positive control for transfection. Cells not transfected with mimics for each the cell lines and DP12 WT cells were also analysed as negative controls. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using an two-tailed homoscedastic Student T-test * = p -value < 0.05

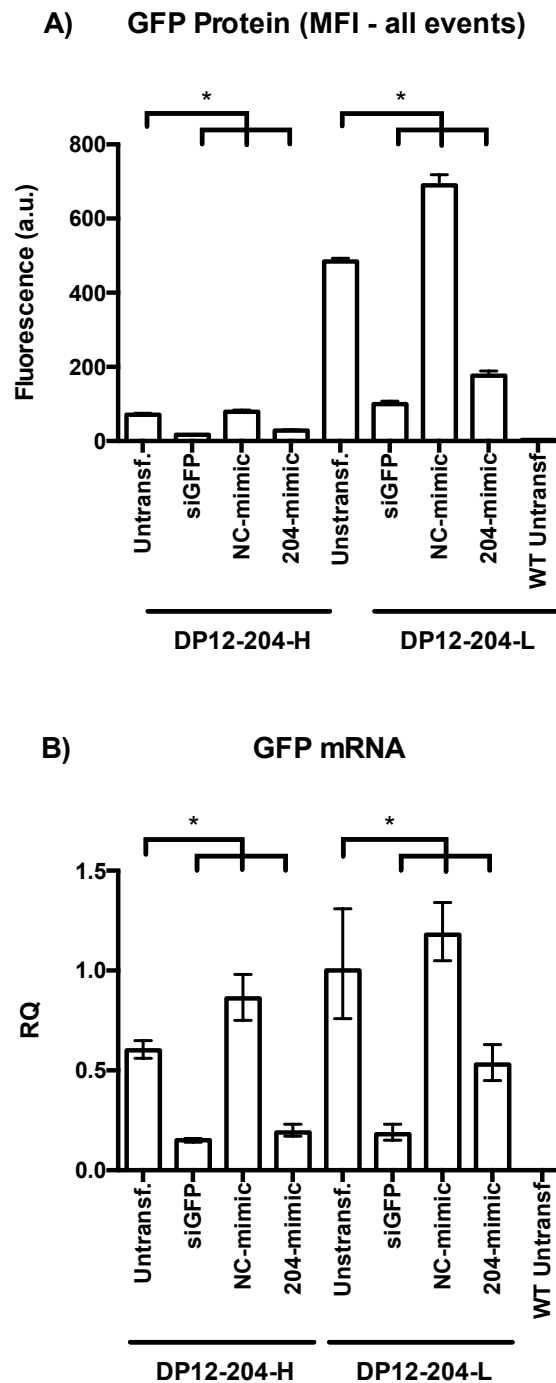


Figure 73 – GFP expression for the miR-toehold transgenic cell lines DP12-204-H and DP12-204-L after transfection with miRNA-mimics specific for the NC sequence (NC-mimic) and the miR-204 sequence (204-mimic). A – Mean Fluorescence Intensity of all events cells and B – GFP mRNA transcript relative quantification. A GFP specific siRNA (siGFP) was used as a positive control for transfection. Cells not transfected with mimics for each the cell lines and DP12 WT cells were also analysed as negative controls. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using a two-tailed homoscedastic Student T-test - * = p -value < 0.05

different levels of expression between the hairpin and linear variants. In this experiment, single stranded DNA oligos with a sequence matching the miRNA mimic sequence were also tested to check if the mRNA downregulation effect was being driven via the miRNA canonical pathway or if simple complementary interference would work.

For that, on the day of transfection, cells were seeded in 24-well plates at a density of 10^6 cells/well in 900 μ L of SFMII media. Cells were transfected with concentrations ranging from 6.25 to 100 nM of 4 different miRNA-mimics/DNA oligos: two mimicking the miR-204 sequence and two mimicking the NC-miRNA sequence. In addition, an siRNA targeting the mRNA transcript of the GFP reporter gene was used as a positive control.

Cells transfected with the miR-NC-mimics did not show any significant reduction of the GFP signal when compared with the untransfected control. Cells transfected with the miR-204-mimics showed a significant reduction of the GFP signal when compared with untransfected cells, to a similar level as the GFP signal measured in cells transfected with the positive control siRNA specific for eGFP (Figure 74). Neither cells transfected with the NC-oligo nor cells transfected with the 204-oligo showed any significant reduction of the GFP signal when compared with the untransfected control (Figure 75).

In summary, miRNA-toehold switches displayed different levels of GFP expression in stable conditions, both at protein and mRNA levels. For both sequences, pools expressing hairpin-like toeholds showed lower levels of mRNA in comparison with the linear counterpart. NC toeholds exhibited lower levels of mRNA transcripts than 204 constructs and these did not correlate with GFP protein. For the 204 toeholds, GFP mRNA levels correlated with GFP protein level. In terms of switchability, all the tested miRNA-toehold switches failed to show increased expression of GFP upon transfection with their complementary miRNA mimics. Instead, the different toeholds switches showed a response to miRNA mimics in a sequence specific manner with further downregulation of GFP expression.

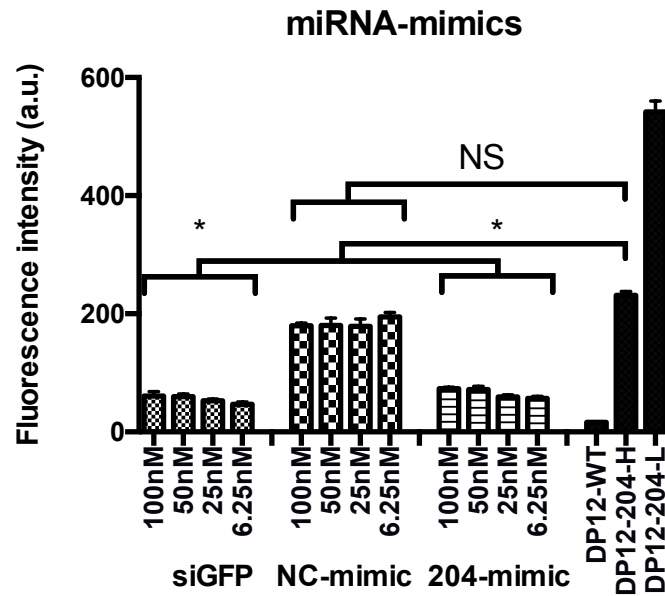


Figure 74 – Fluorescence intensity measured on DP12-204 cells, 24 hours after transfection with the siRNA specific for GFP (siGFP), and two different miRNA-mimics: a NC-mimic and a 204-mimic at different concentrations 100-6.25 nM. DP12 wild type (WT), Untransfected DP12-204-H and DP12-204-L were measured as well as controls. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using a two-tailed homoscedastic Student T-test - * = p -value < 0.05

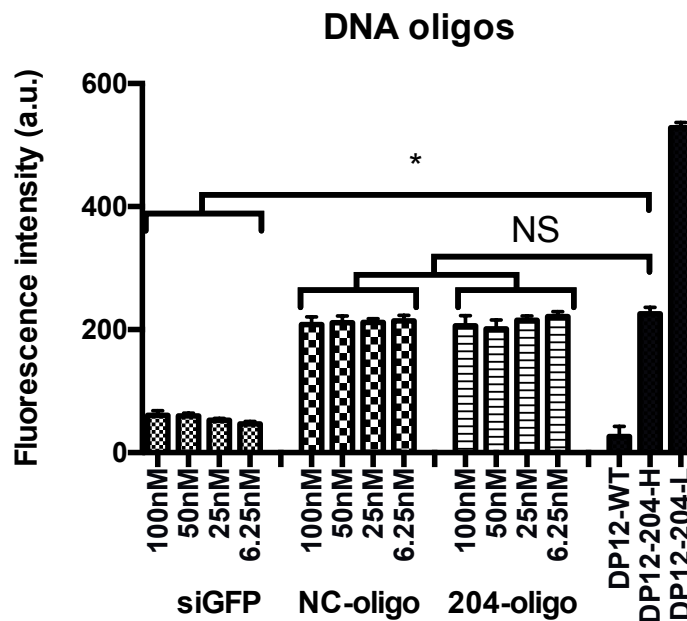


Figure 75 - Fluorescence intensity measured on DP12-204 cells, 24 hours after transfection with the siRNA specific for GFP (siGFP), and the two different DNA single stranded oligos: NC-oligo and 204-oligo at different concentrations 100-6.25 nM. DP12 wild type (WT), Untransfected DP12-204-H and DP12-204-L were measured as well as controls. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using a two-tailed homoscedastic Student T-test - * = p -value < 0.05

To further investigate the mechanism behind the observed downregulation driven by the miRNA, bulged mimics for cgr-miR-204-3p, whose sequence included mismatch regions with the 5'-end, 3'-end and nucleotides 10-11' of the miRNA binding site present in the toehold switch, were transfected in DP12 204-H and 204-L cell lines. There were two rationales for this. On the one hand, this experiment aimed to confirm that the observed downregulation of GFP both at protein and mRNA transcript levels was due to miRNA-mediated transcript degradation (MMTD), which requires fully complementarity of the seed region (5'-end of the mimic) and the 10/11nt for slicing activity. On the other hand, by having bulged mimics unable to drive MMTD, the question was whether the hairpin structure could be disrupted and whether it would cause an up regulation of GFP signal.

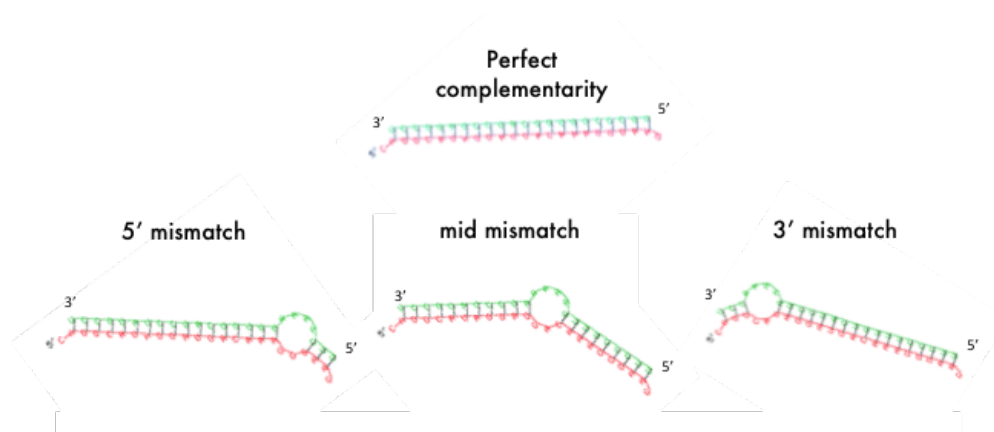


Figure 76 – Bulged mimics used to test miRNA mediated transcript degradation on miRNA-toehold switches.

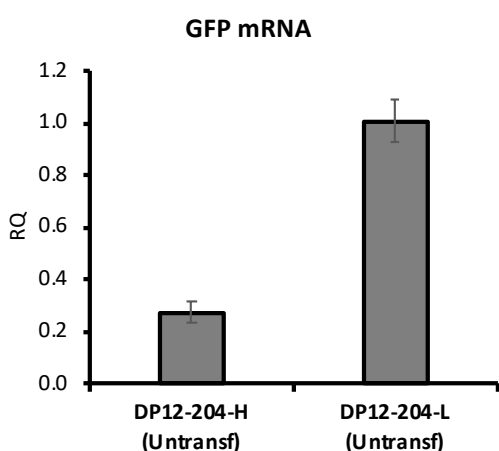


Figure 77 - Relative quantification of GFP mRNA levels between the DP12-204-H pool and the DP12-204-L pool. Untransfected stable pools DP12-204-H and DP12-204-L were analysed by RT-qPCR to perform the relative quantification of the GFP mRNA levels. GAPDH was used as endogenous control. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation.

Looking at GFP mRNA levels in the DP12 204-H and DP12 204-L pools before transfecting them with mimics, it was observed that H-pools had again lower levels of GFP mRNA transcripts than the L-pools (Figure 77).

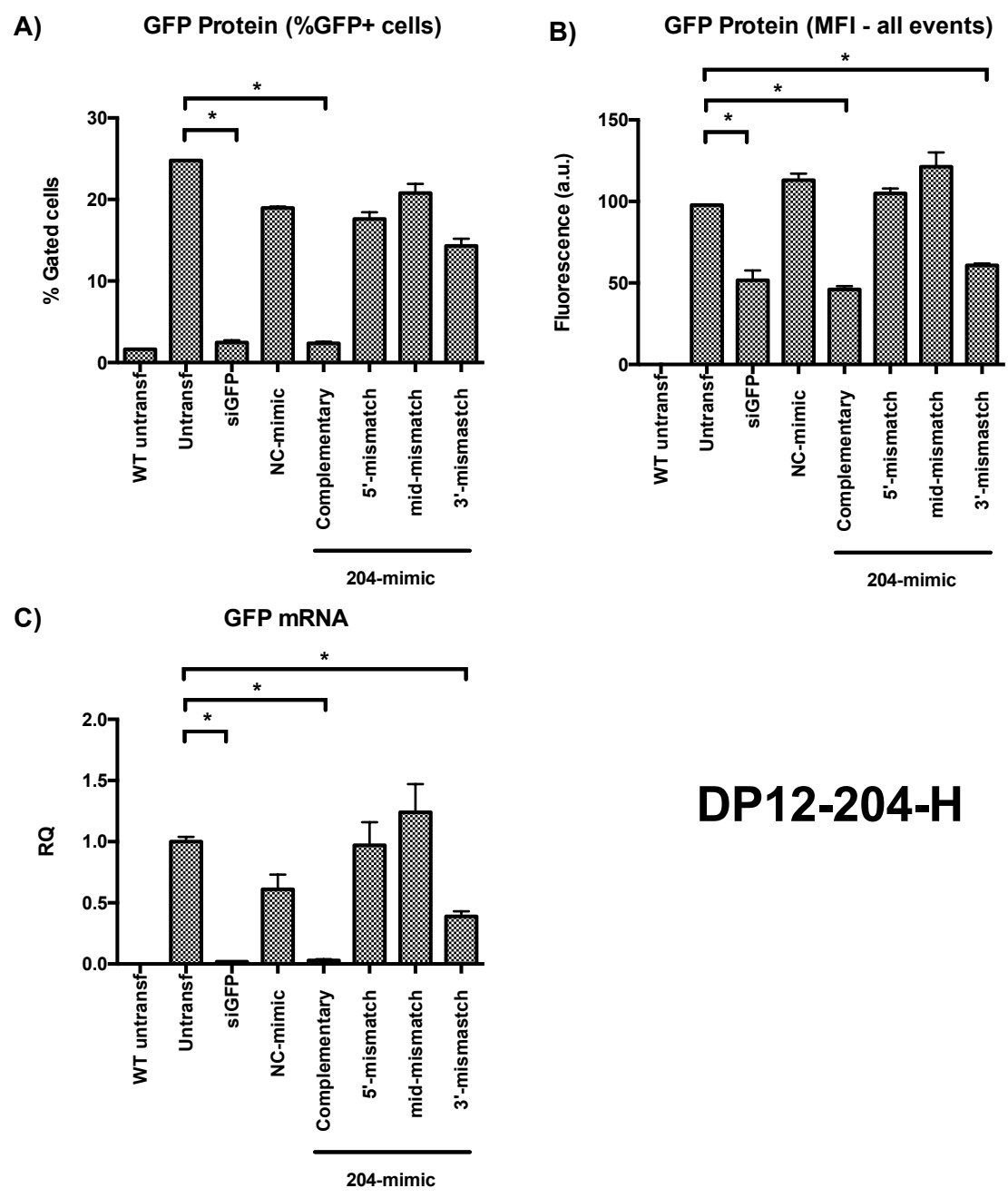


Figure 78 - GFP expression (%GFP cells, MFI and mRNA transcript levels) of the DP12-204-H cell line transfected with bulged miRNA mimics. 5AA – mismatch at the 5'end, midAA – mismatch at the 10/11th nucleotide and 3AA-mismatch at the 3'-end. An eGFP siRNA was used as positive control and a NC-mimic was used as negative control. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using a two-tailed homoscedastic Student T-test - * = p-value < 0.05

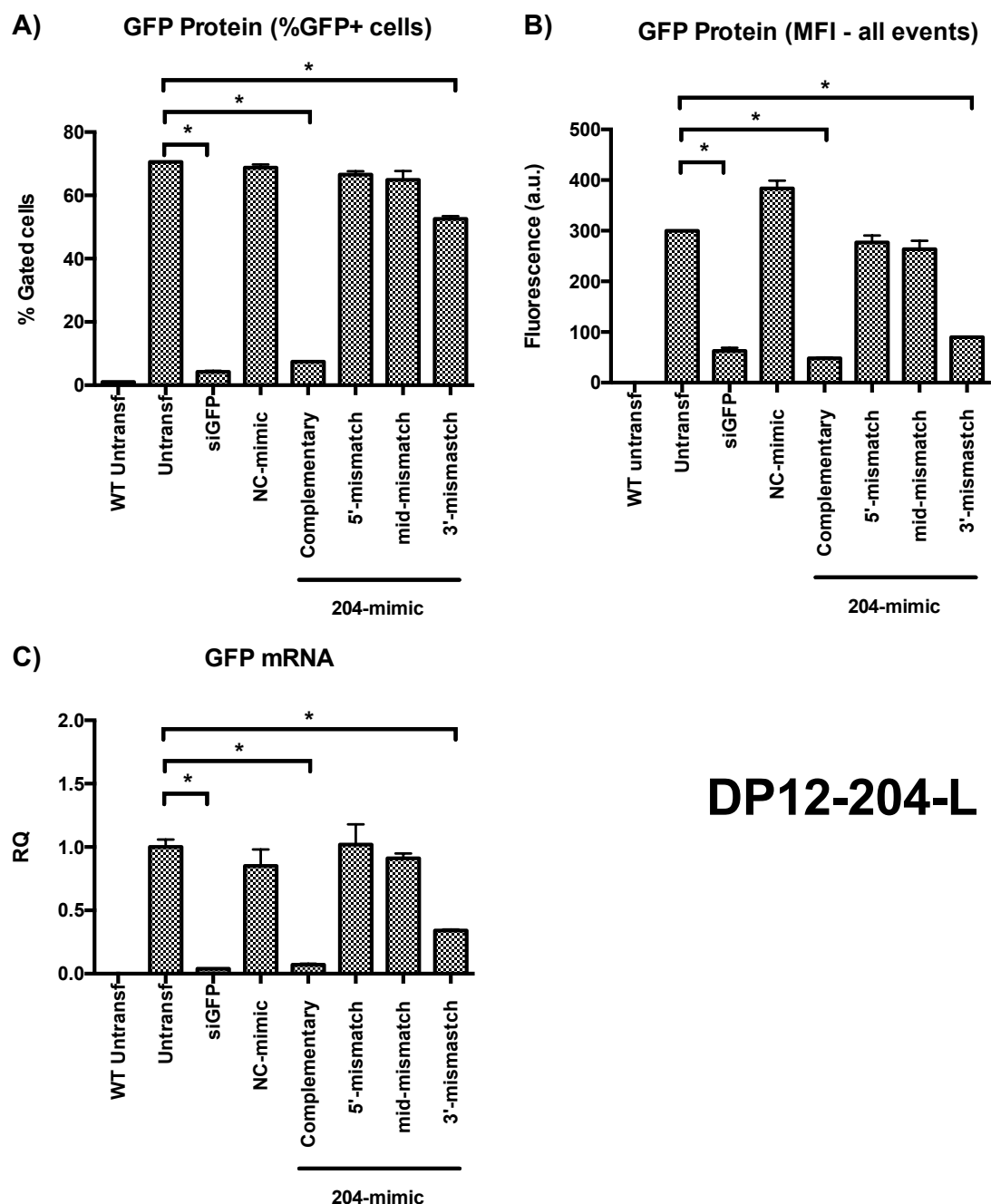


Figure 79 - GFP expression (%GFP cells, MFI and mRNA transcript levels) of the DP12-204-H cell line transfected with bulged miRNA mimics. 5AA – mismatch at the 5' end, midAA – mismatch at the 10/11th nucleotide and 3AA-mismatch at the 3' end. An eGFP siRNA was used as positive control and a NC-mimic was used as negative control. Each bar represents the average of 3 biological replicate transfections and error bars represent their standard deviation. Differences were evaluated using a two-tailed homoscedastic Student T-test - * = p -value < 0.05

Regarding the effect of the different miRNA mimics on GFP expression, transfecting cells with mimics with mismatching sequences at the 5' end and at the 10/11th nucleotide did not cause any downregulation in GFP expression when compared with the NC-mimic control (Figure 78 and Figure 79 – 5' and mid-mismatch). However, no up regulation of GFP expression was observed either, ruling out any potential interaction with the hairpin acting as an ON switch. In contrast, mimics with a mismatch on the 3'-end were able to drive a significant downregulation of GFP expression both in H and L cell lines, that could be observed at protein (%GFP cells and MFI) and mRNA levels (Figure 78 and Figure 79 – 5' and 3'-missmatch) and was comparable to the one obtained with the siGFP positive control and the fully complementary mimic.

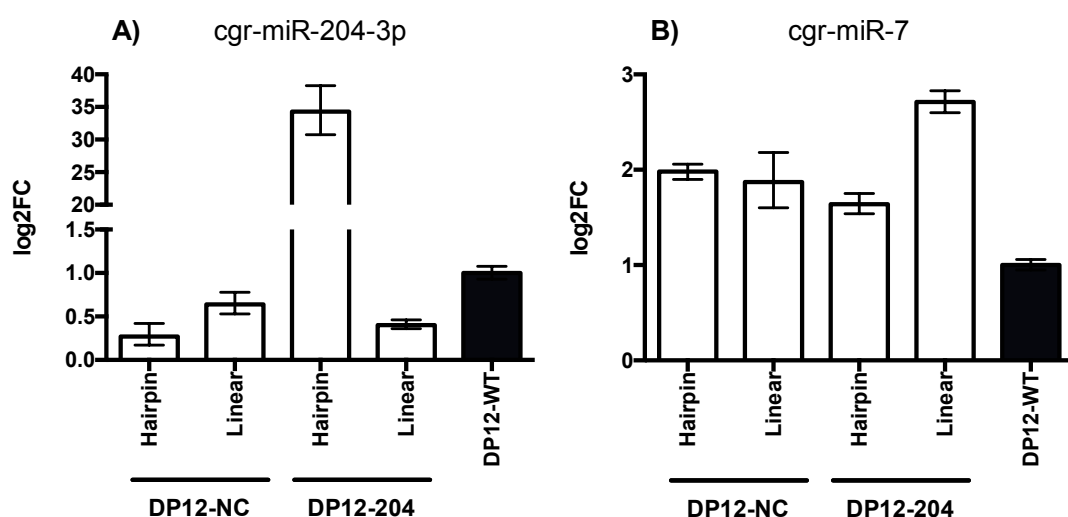


Figure 80 – Relative miRNA abundance in the four miR-toehold transgenic cell lines DP12-NC-H, DP12-NC-L, DP12-204-H and DP12-204-L: A) cgr-miR-204-3p and B) cgr-miR-7. Relative quantification was calculated using the U6 snRNA as endogenous control. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation (N=1).

The relative abundance of endogenous miR-204 was also investigated in order to evaluate possible correlations between miRNA levels and reporter d2eGFP levels and compare between the NC-H/NC-L and 204-H/204-L stable pools. Relative abundance of miR-7 was also measured as a control. U6 snRNA was used as endogenous control for relative quantification. miR-7 expression was up-regulated ca. 2-3 fold in the 4 pools when compared to miR-7 levels in wild type DP12s. Interestingly and unexpectedly, miR-204 was highly up-regulated DP12-204-H, ca. 35-fold in comparison with the WT DP12 control, while it was down-regulated in the other three tested stable pools (Figure 80).

In summary, moving stable pools to suspension changed the observed H/L differences in the NC toeholds while these were maintained in the 204 H/L toeholds. However, both tested toehold sequences, NC and 204, hairpin-like toeholds (H) showed lower mRNA expression levels than their linear counter parts, indicating that the presence or absence of a structured domain in the switch could be causing differences in the transcription of the construct or could be causing the resulting mRNA to have different stability. At this point, it was decided to focus on further investigating the 204-H/L.

2.7. *In silico* analysis and transient expression of miR-toehold variants library

Overall, the results obtained on the first generation of miRNA-toehold switches indicated that hairpin-like toeholds are able to drive downregulation of GFP expression. However, the design of first generation of toehold switches was guided by predictions of their secondary structure, without considering that differences in sequence, while necessary to drive the desired folding pattern of the switch, can also have an effect in transgene expression. In order to be able to control these changes, a second generation of toehold vectors was designed and cloned using an in-house modular system as shown in which the toehold sequence was divided in three fragments (Figure 81):

- a first fragment comprising the 5' end of the toehold sequence and the miRNA-binding site (A)
- a second fragment comprising the loop region and the Kozak-start codon sequences (B)
- a third fragment comprising the complementary region to the toehold region (or non-complementary for the linear constructs) and the linker region to maintain the sequence in frame with the downstream GFP coding sequence (C)

1st Generation (single fragment)

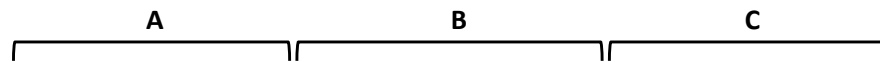
204-L-1G

GGTACCAGGCATAGGATGACAAAGGGAA---AGGAGGAAAAAAAGCCACCATGCGGGCTGTC---GGACAAAGGGAAAGGGGGGGGGGCGCGCC
 လောင်စာစတင်နေရာကိုသတ်မှတ်ပါ

204-H-1G

GGTACCAGGCATAGGATGACAAAGGGAA---GCCGCCTTTTTCGCCACCATGCGGGCTGTC---TCCCTTTGTCAAGGGGGGGGGGCGCGCC
 လောင်စာစတင်နေရာကိုသတ်မှတ်ပါ

2nd Generation (A-B-C cloning)



204-H-2G

GGTACCAGGCATAGGATGACAAAGGGAA**GTAG**CGCGCCTTTTTCGCCACCATGCGGGCTGTC**TAC**GGACAAAGGGAAAGGGGGGGGGGCGCGCC
 လောင်စာစတင်နေရာကိုသတ်မှတ်ပါ

204-H-B5-2G

GGTACCAGGCATAGGATGACAA-TTGAA**GTAG**CGCGCCTTTTTCGCCACCATGCGGGCTGTC**TAC**TTCAATTGTCAAGGGGGGGGGGCGCGCC
 လောင်စာစတင်နေရာကိုသတ်မှတ်ပါ

204-H-B11-2G

GGTACCAGGCATAGGA-CTGAAAGGGAA**GTAG**CGCGCCTTTTTCGCCACCATGCGGGCTGTC**TAC**TTCCCTTTTCAGAGGGGGGGGGGCGCGCC
 လောင်စာစတင်နေရာကိုသတ်မှတ်ပါ

Figure 81 - Modular design and aligned sequences for the second-generation miRNA-toehold sequences. First generation 204-sequences are shown above for reference. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlight in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

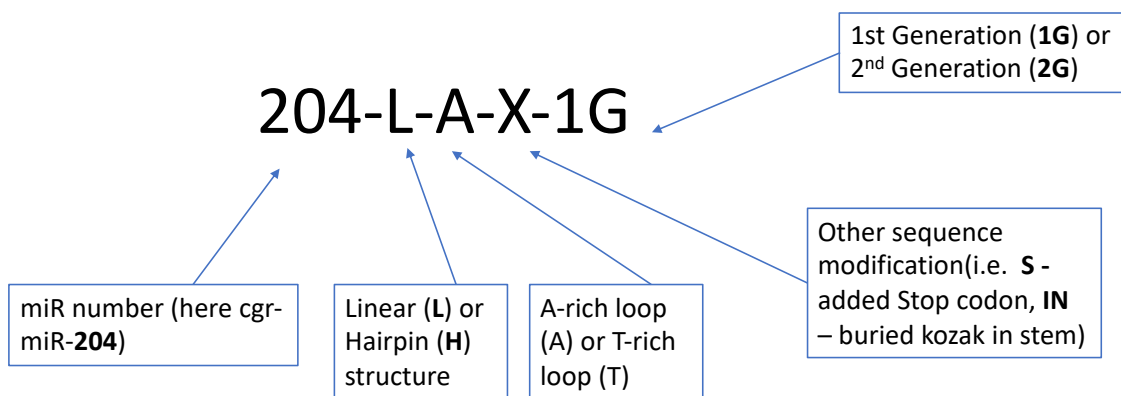


Figure 82 - Nomenclature legend for the miRNA-toeholds in this chapter.

Using the system described above, it was possible to clone toehold variants that would help us evaluate the impact of the different sequence domains on the ability of the UTR

to modulate transgene expression. By using the 204-H/L toehold switch pair as starting point and maintaining an unchanged endogenous miR switch fixed in sequence, a first small set of 2G toeholds were constructed to include mutations on the miRNA binding site introducing bulges to modulate miRNA binding/ RISC-mediated degradation.

In order to keep a consistent nomenclature to name the different miRNA-toeholds from this point, a coded naming system was defined as shown in Figure 82.

2.8. Mutations in the miRNA binding site introducing bulges to interfere with miRNA binding/ RISC-mediated degradation

In the previous section, with the incorporation of bulges at specific locations of the miRNA mimics it was possible to inhibit the downregulation of GFP expression. In order to check if having the mismatching regions in the mRNA transcript would have the same effect, miR-204 toehold vectors containing mutated miRNA binding sites in order to affect miRNA binding were cloned and evaluated. In particular, three variant toeholds were generated comprising mutations in the miRNA binding site as described below. These mutations, as shown previously in the miR-204-toehold stable pools / mimics experiments, were aimed at interfering with miRNA-seed region binding (nt. 1-8) or interfering with the slicing activity of the miRISC complex (nt. 11). See Figure 83 and Figure 84 for sequence information and Figure 85 for the secondary structure prediction for those constructs:

- 204-H-T-2G – this construct has the same sequence as the first generation miR-204 toehold construct (204-H-T-1G) with 6 more nucleotides (GTA and TAC) inserted as sticky ends for cloning of the A/B/C fragments.
- 204-H-T-B5-2G – this construct contains two mutated nucleotides and one deleted nucleotide at the 3' end of the miRNA-binding site to allow imperfect matching of the miRNA 5' seed region and the formation of a bulge.
- 204-H-T-B11-2G – this construct contains two mutated nucleotides and one deleted nucleotide around the 11-nucleotide region to allow for imperfect matching of the miRNA at the cutting site for Dicer.

204-L-A-1G GGTACCAGGCATAGGATGACAAAGGGA**A**---AGGAGGAAAAAAAGCCACC**ATG**GCGGCTGTC---GGACAAAGGGAAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

204-H-T-1G GGTACCAGGCATAGGATGACAAAGGGA**A**---GCCGCCCTTTT**TTT**GCCACC**ATG**GCGGCTGTC---TTCCTTTGTCAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

204-L-T-2G GGTACCAGGCATAGGATGACAAAGGGA**AGTAGCCGCCCTTTT****TTT**GCCACC**ATG**GCGGCTGTC**TAC**GGACAAAGGGAAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

204-H-T-2G GGTACCAGGCATAGGATGACAAAGGGA**AGTAGCCGCCCTTTT****TTT**GCCACC**ATG**GCGGCTGTC**TAC**TTCCTTTGTCAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

204-H-T-B5-2G GGTACCAGGCATAGGATGACAA-TTGA**AGTAGCCGCCCTTTT****TTT**GCCACC**ATG**GCGGCTGTC**TAC**TTCAATTGTCAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

204-H-T-B11-2G GGTACCAGGCATAGGA-CTGAAAGGGA**AGTAGCCGCCCTTTT****TTT**GCCACC**ATG**GCGGCTGTC**TAC**TTCCTTTTTCAGAGGG
 လာဝ်ဂ်မာလဲလဲလဲလဲလဲလဲလဲလဲလဲ

Figure 83 - Modular design and aligned sequences from second generation miRNA-toehold sequences with mutated miRNA-binding sites. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlight in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

204-L-A-1G G T R H R Met T K G **R L F F** **A T Met** A A V - F P L S G G G G A P V S K

204-H-T-1G G T R H R Met T K G **E E K K** **A T Met** A A V - G Q R E G G G G A P V S K

204-L-T-2G G T R H R Met T K G **S R L F F** **A T Met** A A V **S** G Q R E G G G G A P V S K

204-H-T-2G G T R H R Met T K G **S R L F F** **A T Met** A A V **S** F P L S G G G G A P V S K

204-H-T-B5-2G Y Q A Stop D D N Stop **S R L F F** **A T Met** A A V **S** F N L S G G G G A P V S K

204-H-T-B11-2G Y Q A Stop D Stop K G **S R L F F** **A T Met** A A V **S** F P F R G G G G A P V S K

Figure 84 – Amino acid sequence translated from second generation miRNA-toehold sequences with mutated miRNA-binding sites. Translated Met1 corresponding to the d2eGFP ORF is highlighted in light blue. The predicted translated sequence for the same ORF upstream of the Kozak-start codon is shown for all sequences.

In addition, a 2G vector corresponding to the 204-L-A-1G was designed. As shown in Figure 85, the 204-L-2G construct was designed so the only difference with the 204-H-2G construct was the 3'-region that is non-complementary (L) or complementary (H) to the 3'-end of the miRNA-binding site. Note that this involved not only adding the cloning

GTA and TAC sites but also to incorporate changes in the loop region between 204-L-1G and 204-L-2G, as the loop region corresponding to the 204-H constructs was used all across the 2G constructs. Namely, “AGGAGGAAAAAAAA” were replaced by “GCCGCCTTTTTTTT”. It must also be noted that this construct, named 204-L-T-2G was designed purely on nucleotide sequence similarity to the 204-H-T-2G and no secondary structure prediction was considered a priori in order to have the best sequence control.

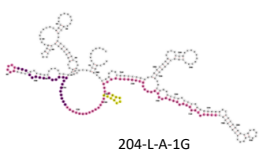
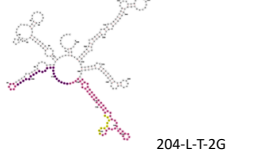
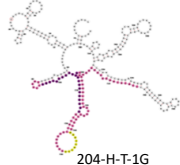
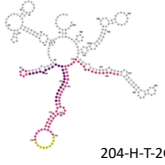
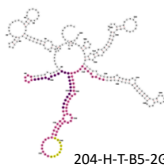
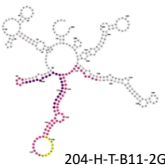
	1 st Generation (1G)	2 nd Generation (2G)
Linear (L)	 204-L-A-1G	 204-L-T-2G
Toehold (H)	 204-H-T-1G	 204-H-T-2G  204-H-T-B5-2G  204-H-T-B11-2G

Figure 85 - Predicted secondary structures using RNA fold from second generation miRNA-toehold sequences with mutated miRNA-binding sites). Predictions include the 5'UTR-region and the first 100+ nucleotides. Relevant sequences are highlighted in colour, toehold insert (pink), miRNA-binding site (violet) and Kozak-ATG sequence (yellow).

The changes introduced to obtain the 204-H-T-2G vectors did not affect the predicted secondary structure, which continued to show a hairpin-like structure around the Kozak-start codon sequence. However, the changes in the nucleotide sequence to create 204-L-T-2G resulted in a change in the predicted secondary structure, which unexpectedly changed from a “linear” conformation to a hairpin-like conformation. This made it impossible to use it as a control to assess the effect of the presence/absence of a hairpin. Despite this, nomenclature remained unchanged (this construct remained named L). In terms of amino acid sequence of the GFP ORF, the main difference between 1G and 2G toeholds was the introduction of an additional Tyrosine (Y) residue as a result of the cloning site for fragments B/C site, TAC.

The four newly constructed 2G vectors were transfected in CHO-K1 cells as previously described. GFP fluorescence levels were measured 24h after transfection and samples were taken for RNA extraction and RT-qPCR analysis to look at GFP and Hygromycin expression levels in each transient pool.

The obtained transient pools 24h after transfection exhibited different percentage of GFP positive cells ranging from 20-50% (Figure 86 – top). The four new 2nd generation (2G) toehold vectors, including the L-2G construct, showed lower GFP expression than the 1st generation linear toehold 204-L-A-1G (Figure 86 – bottom). The sequence changes included to create a bulge in the miRNA-binding site did not have a major effect on GFP expression, and only led to a modest decrease of GFP expression observed in 204-H-T-B11-2G, mutated at the 11th nucleotide of the 204-miR binding site. The absence of major differences in GFP expression between toehold with mismatch and fully complementary miRNA-binding sites indicated that the observed downregulation of GFP expression was unlikely to be related to miRNA binding and subsequent miRNA-mediated transcript degradation. In these constructs, the bulge in the seed region or around the 10th/11th nucleotide should inhibit miRNA binding and slicing activity of the miRISC. The most striking result here was observed when looking at the GFP mRNA levels in the different transient pools (Figure 87 - top). Similar levels of hygromycin resistance marker across the different pools were expected considering that this gene is expressed from an independent promoter cassette present in the expression vector as well as in the filler DNA used during transfection. It was anticipated that all pools should show similar levels of hygromycin and this would be a good control to verify equal transfection efficiencies and to normalize the data for equal amount of plasmid content. However, different levels of hygromycin mRNA were observed in each pool. These hygromycin mRNA levels correlated with the levels of GFP mRNA observed, that were also different for each transient pool. These differences could be explained by differences in transfection efficiency although all plasmids were of similar size and the same transfection protocol was used for all of them. It must be therefore assumed that transfection efficiencies should be comparable. Cell densities at 24-hours post-transfection were also comparable across different samples, with values around 1.2×10^6 cells/mL $\pm 5\%$, indicating no differences in growth between the different transfected pools. In order to correct for the apparent differences in plasmid content, it was decided to use the measured hygromycin mRNA levels to normalize the GFP mRNA data.

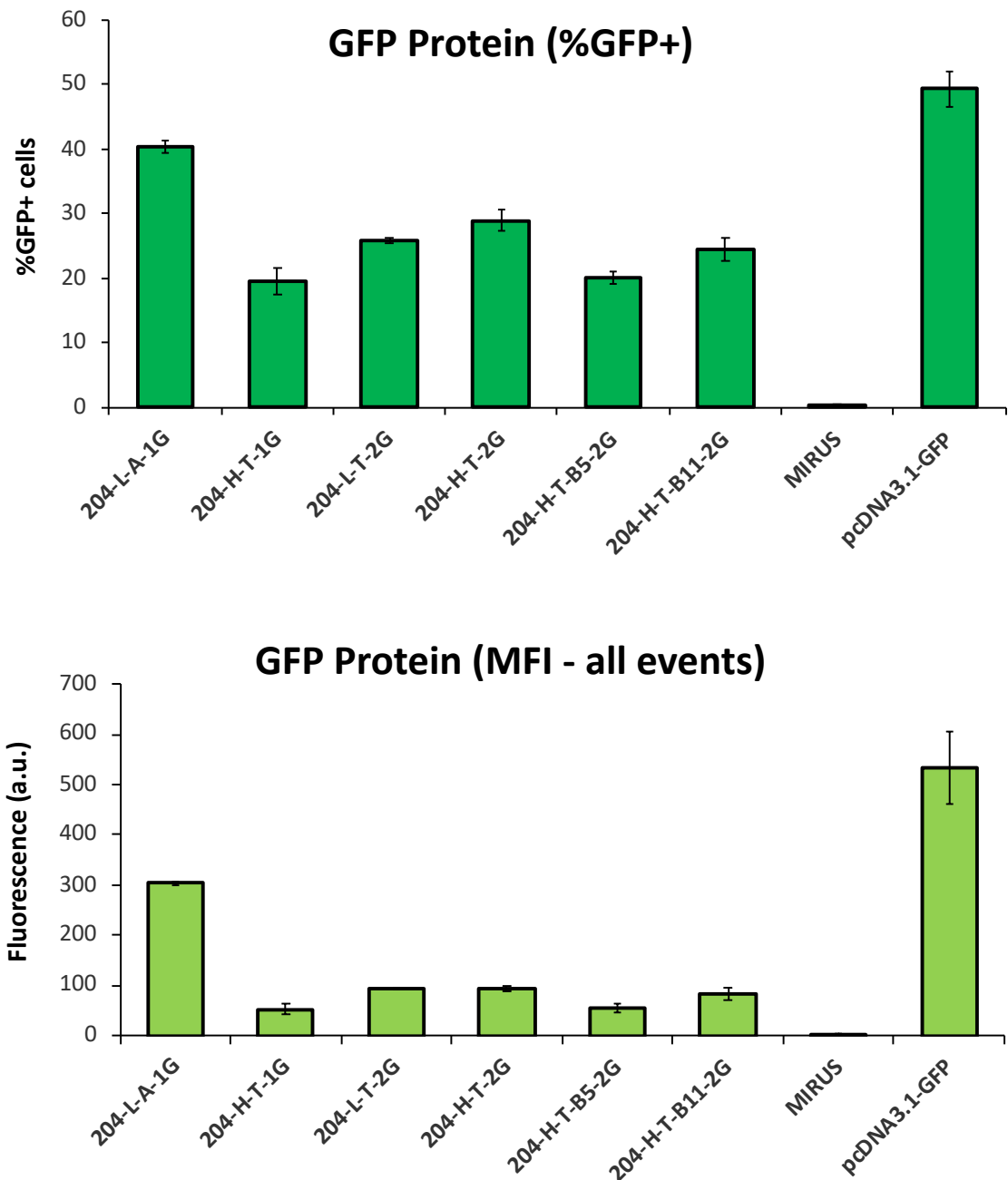


Figure 86 – Percentage of GFP positive cells and Mean Fluorescence Intensity (MFI – all events) from second generation miRNA-toehold sequences with mutated miRNA-binding sites. Cell populations were measured 24h after transfection. A pcDNA3.1 GFP plasmid and an only Mirus™ condition were used as positive and negative controls for transfection and flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

Looking at this normalized transcript levels, the different toehold plasmids showed fairly similar levels of GFP mRNA transcripts, all lower than the no toehold GFP control (Figure 87 – bottom).

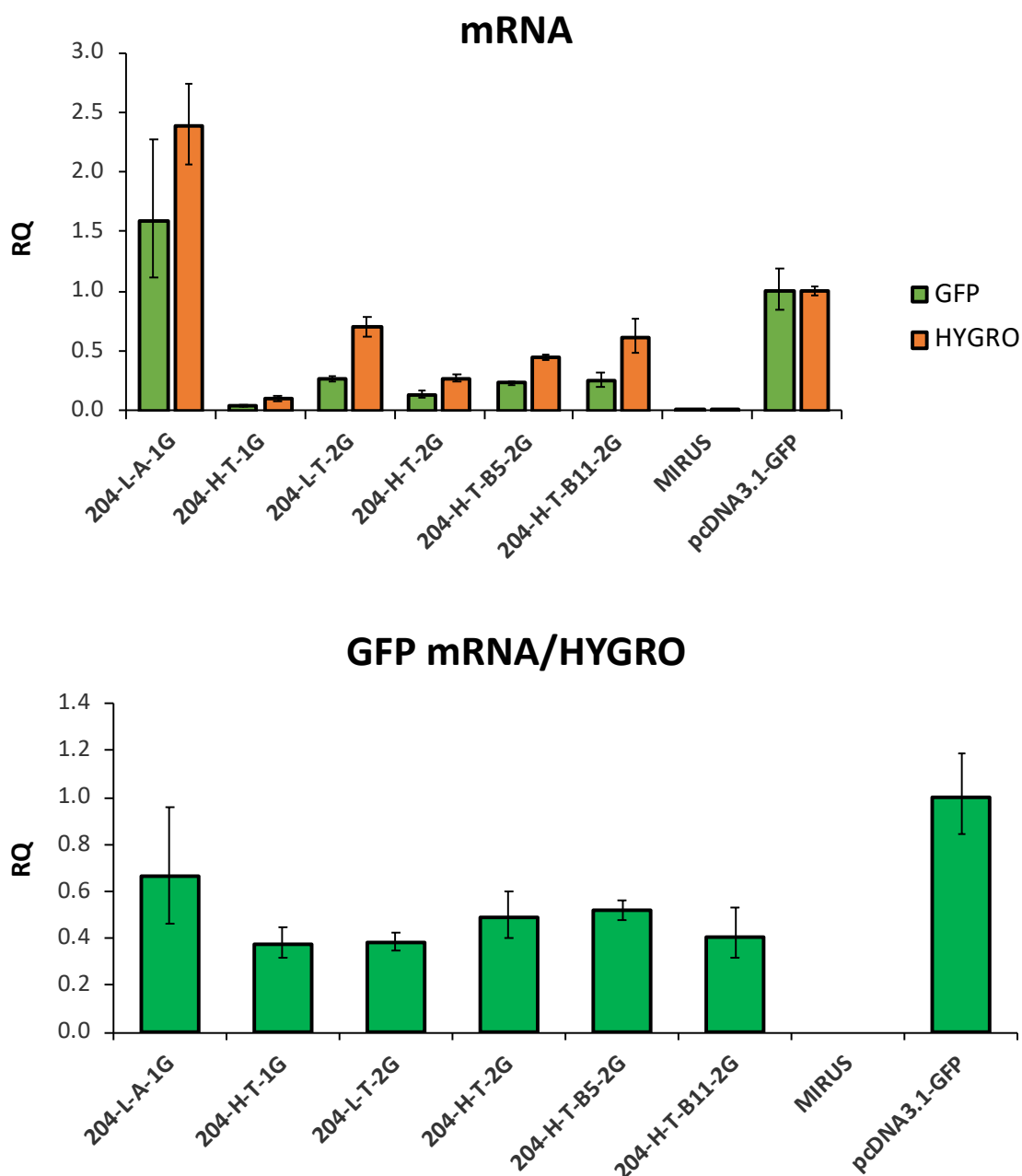


Figure 87 – GFP expression at the mRNA level second generation miRNA-toehold sequences with mutated miRNA-binding sites. Top - Relative quantification of GFP and Hygromycin mRNA transcripts. **Bottom** –Normalized GFP mRNA transcript levels using Hygromycin mRNA transcript levels to account for differences in transfection efficiency/plasmid content. An only Mirus™ condition and No Template Control (NTC) qPCR reactions were used as negative controls for cDNA synthesis / DNase treatment and qPCR. GAPDH was used as endogenous control and RQ values were calculated using the expression in the pcDNA3.1 GFP samples as reference. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation.

GFP protein expression (fluorescence) data was divided by normalized GFP mRNA levels to get an indication of the efficiency of translation for mRNA from each different toehold.

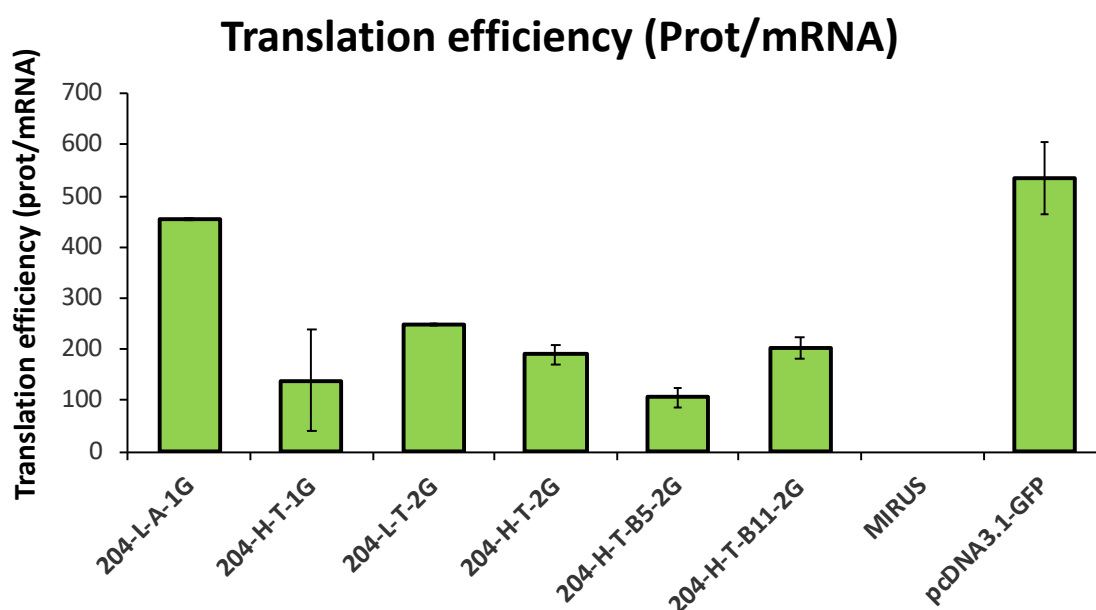


Figure 88 – Translation efficiency of GFP mRNA transcripts expressed from second generation miRNA-toehold sequences with mutated miRNA-binding sites. GFP protein levels were divided by normalized GFP mRNA transcript levels as a measure of translation efficiency of each mRNA transcript. Each bar represents the average of 3 technical replicate transfections / qPCR quantification ratios and error bars represent their standard deviation.

Transcripts originating from the 204-L-A-1G construct showed higher translation efficiencies, comparable to the translation efficiency observed for the no toehold GFP control, than the translation efficiencies obtained for transcripts coming from all the different 2G-toeholds (Figure 88). This data suggested that the presence of a hairpin-like secondary structure around the Kozak-ATG sequence, present in all constructs including the 204-L-T-2G, was responsible for the observed reduction in translation efficiency.

Differences could also come from cells growing at different rates due to differences in GFP expression. It has been shown before in this work that GFP cytotoxicity can impact growth and slower growing cells would dilute their plasmid content less than faster growing cells. Lower viabilities were measured for cells transfected with the GFP control with ca. 90% cell viability compared with 95+% observed for the remaining transfections.

From all the 2nd generation toehold tested, the most surprising reduction of GFP expression was the one observed for the toehold 204-L-T-2G. As previously mentioned, this L-toehold was initially designed to act as “linear” control for the 2nd generation of toeholds and only differed from the H-2G construct in the sequence region that matches the 3’ of the miRNA binding site allowing the formation of the hairpin stem. However,

its secondary structure was predicted to be hairpin-like and the GFP expression and the translation efficiencies observed from this construct comparable to the ones observed for the other hairpin toeholds.

Considering all the 2G toeholds in the first subset exhibited these reduced levels of GFP expression and apparent lower translation efficiencies, it was decided to investigate further changes in the toehold sequence. The following list of changes were anticipated to potentially modify GFP expression by either potentially modifying the toehold secondary structure or changing certain sequence features:

- 1) Introduction of a stop codon before the Kozak-start codon, in order to assess putative leaky expression from the uORF
- 2) Position of Kozak-start codon region in the hairpin
- 3) Mutations in the 5'-end of fragment C, corresponding the complementary region to the toehold region (or non-complementary for the linear constructs) to assess their effect on the expression of GFP and the effect of potential changed amino acids in the ORF caused by changes on the nucleotide sequence.
- 4) Sequence of the loop region

The vectors constructed to evaluate each of the above-mentioned points and the obtained GFP expression data for each of them are described in the following sections.

2.9. Position of Kozak-start codon region in the hairpin and introduction of a stop codon before the GFP ORF Kozak-start codon

Burying the Kozak-ATG sequence in a hairpin-like secondary structure seemed to cause the low expression of GFP expression observed by affecting translation efficiency of the GFP mRNA transcript. In order to attempt to achieve further repression of GFP expression, the effect of changing the position of the Kozak sequence-ATG in the hairpin was investigated. Two variants of the 204-H-T-2G toehold with the Kozak-ATG region buried further in the stem region of the hairpin were generated (Figure 89). Changing the position of the Kozak-ATG sequence caused an enlargement of the bulge present in the stem region of the hairpin for 204-H-T-IN2-2G (Figure 91). In order to “fix” this, the second variant 204-H-T-IN1-2G had the 6 nucleotides right after the GTA cloning site changed to match the first 6 nucleotides of the Kozak-ATG sequence, resulting in a tighter hairpin without a bulge (Figure 91). In addition to shifting the Kozak region

toward the 3' end of the toehold to place it deeper in the stem region, a stop codon was introduced in order to assess putative leaky expression from the upstream ORF. As mentioned before, the miR-204 binding site in the toehold contained an ATG codon that despite not being part of a Kozak-like sequence was in frame with the GFP coding sequence and could potentially recruit translational machinery and impair correct translation of the GFP mRNA transcript. It is true that this start codon was also present in the 204-H-T-1G construct, which had shown similarly high levels of GFP expression and similar translation efficiencies than the no toehold vector. While literature shows that the presence of out-of-frame upstream ORFs (uORFs) from the main ORF can impair translation (Liang *et al.*, 2016), it was considered that in the miRNA-toehold constructs the presence of an in-frame uORF could actually be driving translation of GFP, thus bypassing the repressive effect of the toehold switch. The effect of removing this upstream ATG codon was assessed by generating two toehold variants with an additional stop codon. The variant 204-H-T-S-2G was identical to 204-H-T-S-2G with the addition of the stop codon, and variant 204-H-T-INS-2G was the match to 204-H-T-IN1-2G with the additional stop codon.

The changes introduced in the IN constructs also affected the amino acid sequence of the GFP protein being expressed, with three amino acids, Alanine-Alanine-Valine being removed from the N-terminal end of the protein in the IN variants when compared with the original 204-H-T-2G construct (Figure 90). No effects on the predicted secondary structure were observed for either of the two constructs (Figure 91).

As previously described, these constructs were transfected in CHO-K1 cells and GFP expression was measured after 24h. Samples for RNA extractions and RT-qPCR of GFP and hygromycin mRNA levels were also taken.

204-H-T-1G	GGTACCAGGCATAGGATGACAAAGGGAA-- SCCGCCTTTT TTC GCCACCATG GCGGCTGTC-- <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ
204-H-T-2G	GGTACCAGGCATAGGATGACAAAGGGAA GT AGCGCCTTTT TTC GCCACCATG GCGGCTGTC TAC <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ
204-H-T-S-2G	GGTACCAGGCATAGGATGACAAAGGGAA GT AGCGCCTTT TAA TTC GCCACCATG GCGGCTGTC TAC <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ
204-H-T-IN2-2G	GGTACCAGGCATAGGATGACAAAGGGAA GT AGCGCCTTTT TTC GCGGCTGTC GCCACCATG TAC <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ
204-H-T-IN1-2G	GGTACCAGGCATAGGATGACAAAGGGAA GT AGGTGGCTTT TTC GCGGCTGTC GCCACCATG TAC <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ
204-H-T-INS-2G	GGTACCAGGCATAGGATGACAAAGGGAA GT AGGTGGCTTT TAA TTC GCGGCTGTC GCCACCATG TAC <u>TTCCTTTGTCAGGG</u> လောင်စာလောင်စာလောင်စာလောင်စာ

Figure 89 - Modular design and aligned sequences for miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlighted in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

204-H-T-2G	G T R H R Met T K G S S R L F F A T Met A A V Y F P L S G G G G A P V S K
204-H-T-S-2G	G T R H R Met T K G S S R L Stop F A T Met A A V Y F P L S G G G G A P V S K
204-H-T-IN2-2G	G T R H R Met T K G S S R L F F A A V A T Met Y F P L S G G G G A P V S K
204-H-T-IN1-2G	G T R H R Met T K G S R W L F F A A V A T Met Y F P L S G G G G A P V S K
204-H-T-INS-2G	G T R H R Met T K G S R W L Stop F A A V A T Met Y F P L S G G G G A P V S K

Figure 90 - Amino acid sequence from translated from miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. From the Kozak-start codon, translated Met1 corresponding to the d2eGFP ORF is highlighted in blue. The predicted translated sequence for the same ORF upstream of the Kozak-start codon is shown for all sequences.

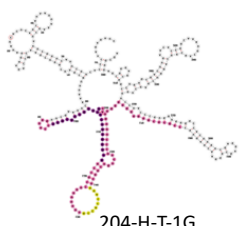
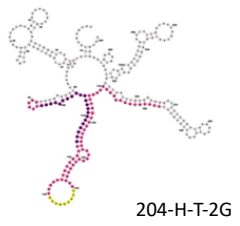
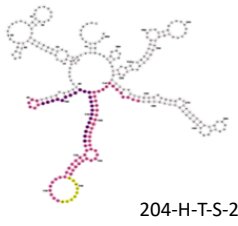
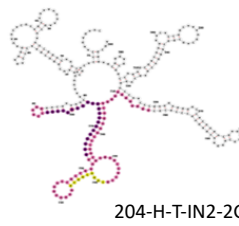
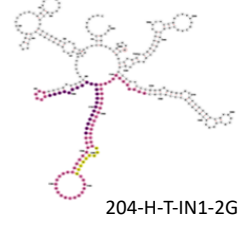
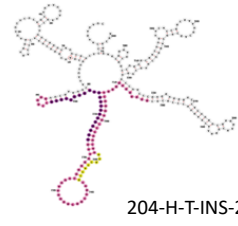
	1G	2G
H	 <p>204-H-T-1G</p>	<div>  <p>204-H-T-2G</p>  <p>204-H-T-S-2G</p> </div> <div>  <p>204-H-T-IN2-2G</p> </div> <div>  <p>204-H-T-IN1-2G</p>  <p>204-H-T-INS-2G</p> </div>

Figure 91 - Predicted secondary structures using RNA fold for the miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. Predictions include the 3'UTR-region and the first 100+ nucleotides. Relevant sequences are highlighted in colour, toehold insert (pink), miRNA-binding site (violet) and Kozak-ATG sequence (yellow).

On this occasion, the changes in the newly constructed IN toeholds had a more pronounced effect on GFP expression when compared with the original 204-H-T-2G toehold (Figure 92). The three IN-2G toeholds showed higher levels of GFP expression, while still showing similar levels of normalized GFP mRNA levels (Figure 93 - bottom).

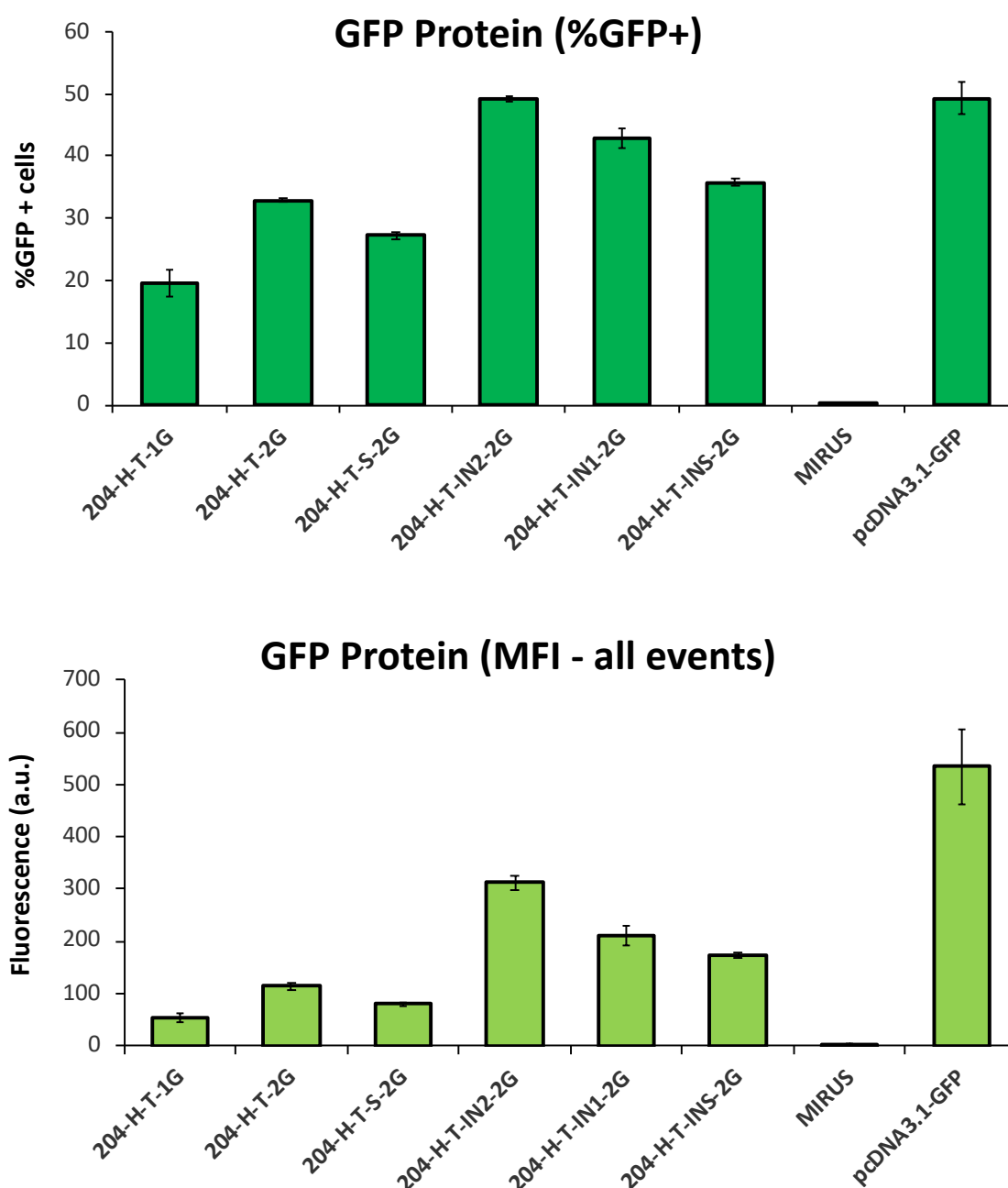


Figure 92 – Percentage of GFP positive cells (top) and Mean Fluorescence Intensity (MFI – all events) (bottom) from miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. Cell populations were measured 24h after transfection. A pcDNA3.1 GFP plasmid and an only Mirus™ condition were used as positive and negative controls for transfection and flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

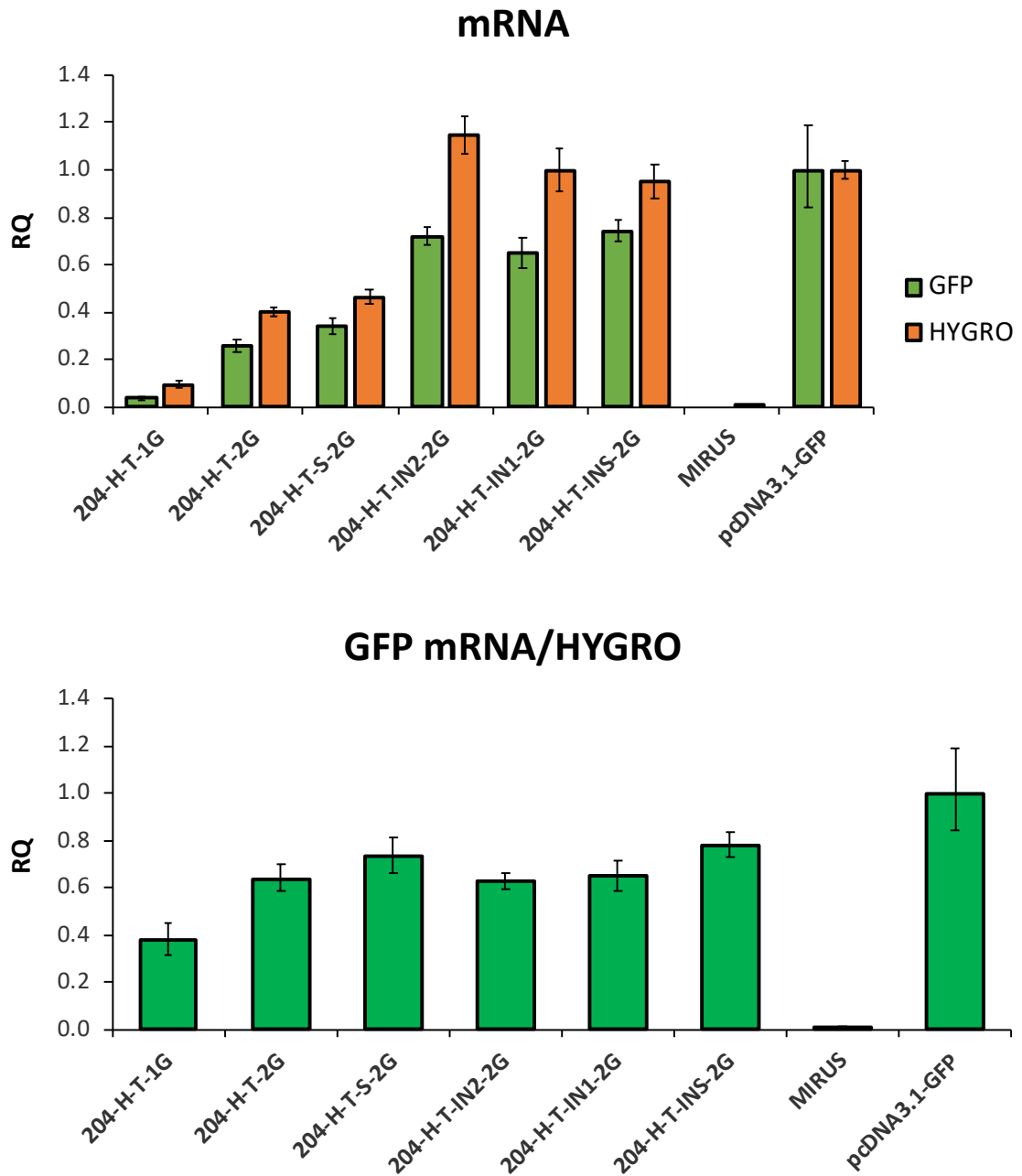


Figure 93 – GFP expression at the mRNA level for miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. Top - Relative quantification of GFP and Hygromycin mRNA transcripts. **Bottom** –Normalized GFP mRNA transcript levels using Hygromycin mRNA transcript levels to account for differences in transfection efficiency/plasmid content. An only Mirus™ condition and No Template Control (NTC) qPCR reactions were used as negative controls for cDNA synthesis / DNase treatment and qPCR. GAPDH was used as endogenous control and RQ values were calculated using the expression in the pcDNA3.1 GFP samples as reference. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation.

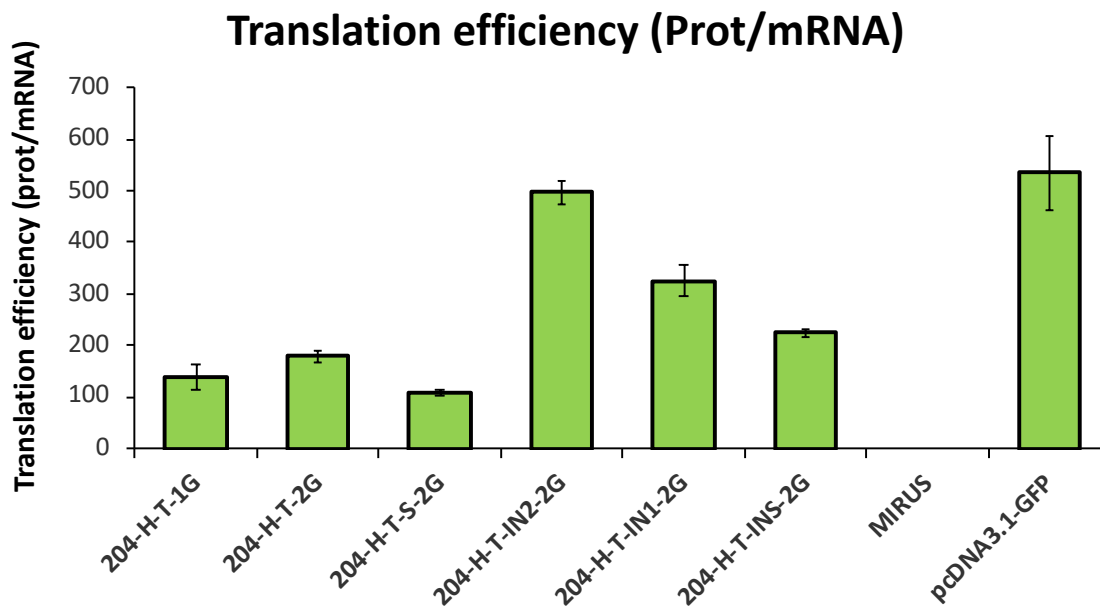


Figure 94 – Translation efficiency for the GFP mRNA transcripts expressed from miRNA-toehold sequences with variable positions for the Kozak-start codon and uORF stop codon. GFP protein levels were divided by normalized GFP mRNA transcript levels as a measure of translation efficiency of each mRNA transcript. Each bar represents the average of 3 technical replicate transfections / qPCR quantification ratios and error bars represent their standard deviation.

Translation efficiencies were also increased for transcripts produced from the IN2 and IN1 constructs when compared with the 204-H-T-2G toeholds and reached similar levels as the ones observed for transcripts from pcDNA3.1-GFP (Figure 94).

The most interesting construct here was 204-H-T-IN2-2G, for which a large increase in GFP expression was observed. This increase in GFP expression was accompanied by a large increase in translation efficiency, which reached similar levels to the no toehold controlled was observed. As previously mentioned, the 204-H-T-IN2-2G construct presented a more open predicted secondary structure environment, which might explain the observed increased translation efficiency and GFP expression. The fact that both GFP expression and translation efficiency were decreased for the construct 204-H-T-IN1-2G in which the hairpin-like structure had been made tighter would support this hypothesis.

When looking at the effect of the introduced STOP codon, in both cases the presence of the additional stop codon led to a decrease in GFP signal related to a decrease in translation efficiency of the transcript. This suggests that some translation may have been taking place starting at the upstream ATG present in the miRNA-binding site.

2.10. Mutations on the 3'-end of fragment C, corresponding the complementary/non-complementary region

It was previously observed that the secondary structure prediction for the construct 204-L-T-2G predicted the formation of a hairpin despite having a non-complementary region to the miRNA-binding site in fragment C. In order to explore if alternative non-complementary regions could lead to transcripts in which the 5'-UTR design would not fold in any hairpin-like secondary structure, emulating the 204-L-A-1G construct, variants of the 204-L-T-2G vector were constructed by changing the sequence at the 5'-end of the C fragment (Figure 85). Two different alternative sequences were tested, corresponding to the non-complementary sequences used previously to generate the linear toeholds for NC-miR and miR-409 (Figure 95). The introduction of alternative non-complementary regions on the 3' C fragment of the toehold continued to generate transcripts predicted to form different hairpin-like structures. While 204-L-T-NC-2G was predicted to form a more relaxed hairpin structure, the prediction for the 204-L-T-409-2G constructs resulted in a longer and tighter hairpin (Figure 96). In addition, the changes performed in the non-complementary regions resulted in a different amino acid sequence for the positions 5-8 of the GFP protein (Figure 97)

204-L-A-1G	GGTACCAGGCATAGGATGACAAAGGGAA---AGGAGGAAAAAAA GCCACCATG GCGGCTGTC--- GGACAAAGGGAA GGG <u>ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်</u>
204-L-T-2G	GGTACCAGGCATAGGATGACAAAGGGAA GTAGCCGCCTTTTTTT GCCACCATG GCGGCTGTC TACGGACAAAGGGAA GGG <u>ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်</u>
204-L-T-NC-2G	GGTACCAGGCATAGGATGACAAAGGGAA GTAGCCGCCTTTTTTT GCCACCATG GCGGCTGTC TACTGTCAGCTTTCT GGG <u>ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်</u>
204-L-T-409-2G	GGTACCAGGCATAGGATGACAAAGGGAA GTAGCCGCCTTTTTTT GCCACCATG GCGGCTGTC TACTCATAAGATTCT GGG <u>ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်ကော့ဂ်</u>

Figure 95 - Modular design and aligned sequences for the miRNA-toehold sequence with mutated non-complementary region. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlighted in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

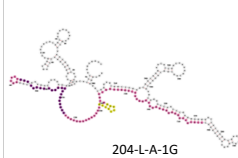
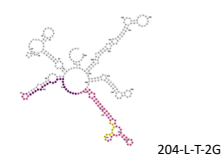
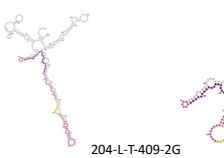
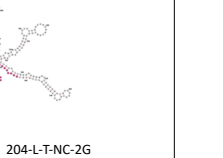
	1 st Generation (1G)	2 nd Generation (2G)			
Linear	 204-L-A-1G	 204-L-T-2G	 204-L-T-409-2G	 204-L-T-NC-2G	

Figure 96 - Predicted secondary structures using RNA fold for the miRNA-toehold sequence with mutated non-complementary region. Predictions include the 3'UTR-region and the first 100+ nucleotides. Relevant sequences are highlighted in colour, toehold insert (pink), miRNA-binding site (violet) and Kozak-ATG sequence (yellow).

204-L-A-1G	G T R H R Met T K G K E E K K A T Met A A V G Q R E G G G G A P V S K
204-L-T-2G	G T R H R Met T K G S S R L F F A T Met A A V Y G Q R E G G G A P V S K
204-L-T-NC-2G	G T R H R Met T K G S S R L F F A T Met A A V Y C Q L S G G G A P V S K
204-L-T-409-2G	G T R H R Met T K G S S R L F F A T Met A A V Y L I R F G G G A P V S K

Figure 97 - Amino acid sequence from translated miRNA-toehold sequence with mutated non-complementary region. From the Kozak-start codon, translated Met1 corresponding to the d2eGFP ORF is highlighted in light blue. The predicted translated sequence for the same ORF upstream of the Kozak-start codon is shown for all sequences.

In terms of GFP expression, the introduction of alternative non-complementary regions on the 3' C fragment of the toehold led to small changes in GFP expression (Figure 98). The percentage of GFP+ cells in the 204-L-NC-2G and 204-L-409-2G transfected populations were increase in comparison with the ones observed for the 204-L-T-2G and the 204-H-T-2G pools and were similar to those observed for the 204-L-A-1G populations. However, none of them were able to recover the GFP expression level observed for 204-L-A-1G, with MFIs for the two new constructs remaining at similar levels as the one of the 204-H-T-2G construct, namely 2-4-fold lower than the MFI obtained in the 204-L-A-1G pool.

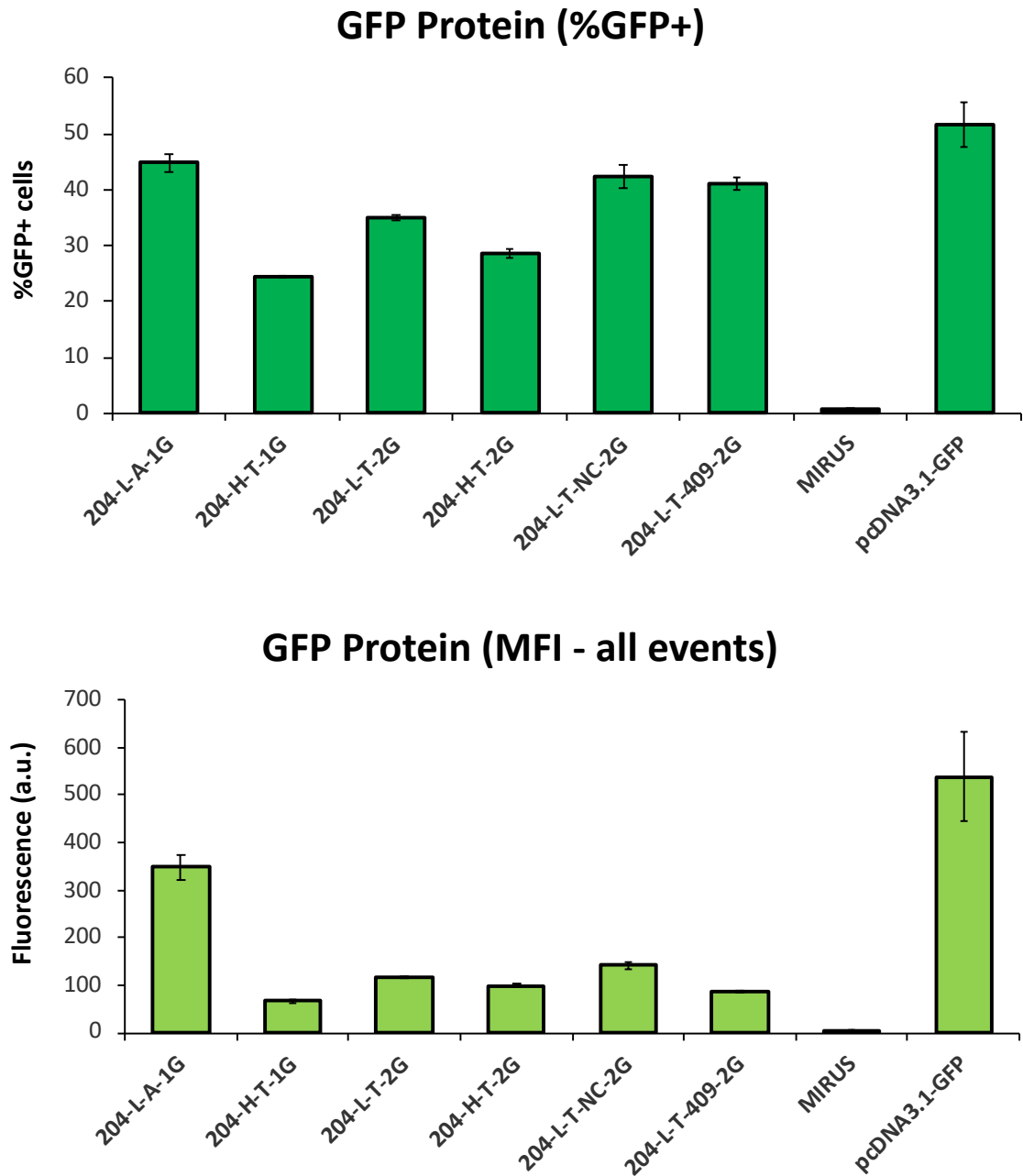


Figure 98 – Percentage of GFP positive cells (top) and Mean Fluorescence Intensity (MFI – all events) (bottom) from miRNA-toehold sequences with mutated non-complementary regions. Cell populations were measured 24h after transfection. A pcDNA3.1 GFP plasmid and an only Mirus™ condition were used as positive and negative controls for transfection and flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

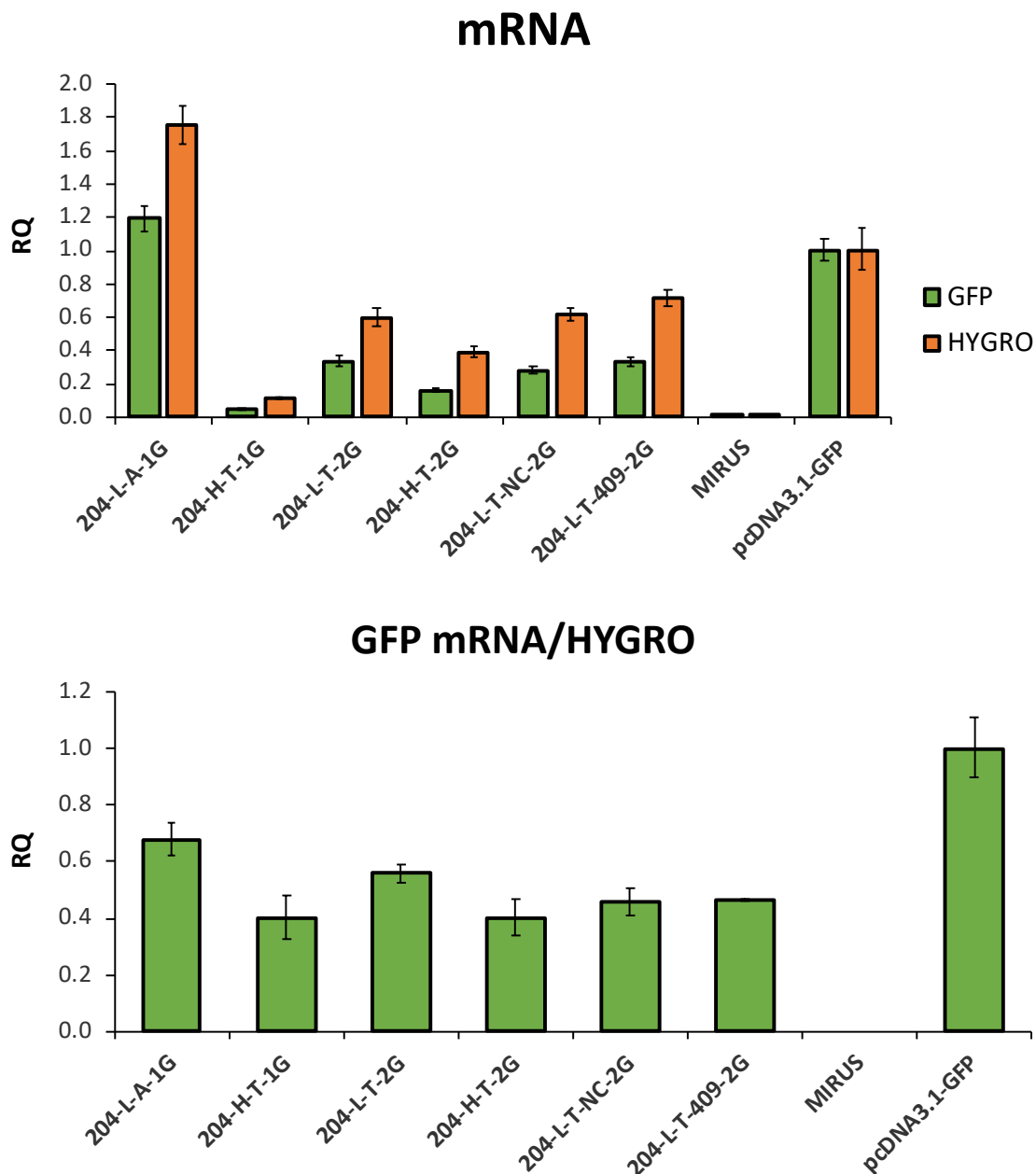


Figure 99 – GFP expression at the mRNA level for second generation miRNA-toehold sequences with mutated non-complementary regions. **Top** - Relative quantification of GFP and Hygromycin mRNA transcripts. **Bottom** –Normalized GFP mRNA transcript levels using Hygromycin mRNA transcript levels to account for differences in transfection efficiency/plasmid content. An only Mirus™ condition and No Template Control (NTC) qPCR reactions were used as negative controls for cDNA synthesis / DNase treatment and qPCR. GAPDH was used as endogenous control and RQ values were calculated using the expression in the pcDNA3.1 GFP samples as reference. Each bar represents the average of 3 technical replicate qPCR wells and error bars represent their standard deviation.

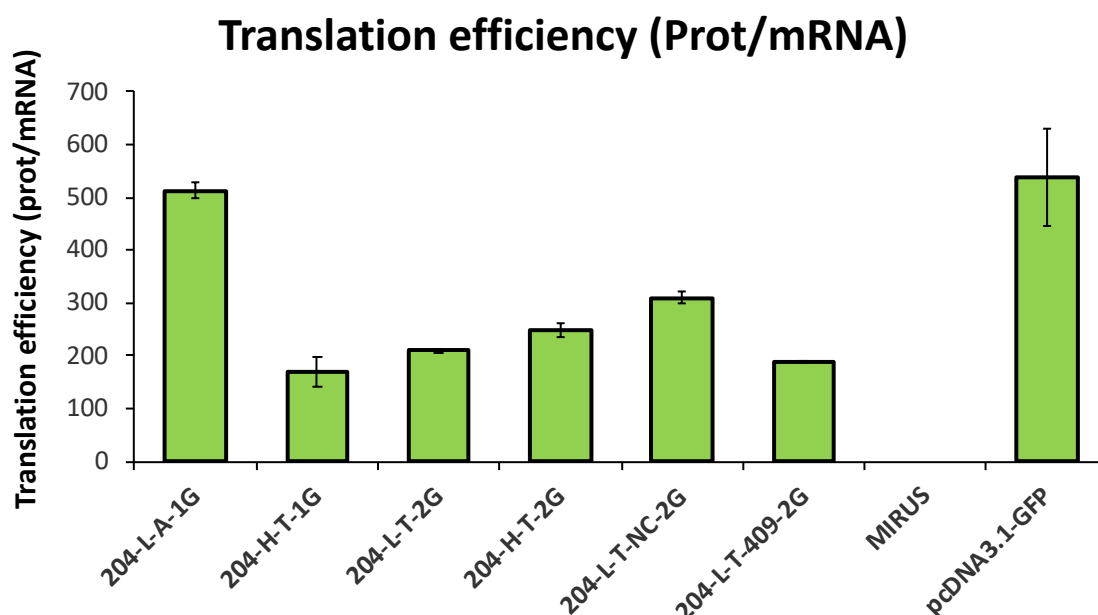


Figure 100 – Translation efficiency for the GFP mRNA transcripts expressed from second generation miRNA-toehold sequence with mutated non-complementary regions. toehold plasmid. GFP protein levels were divided by normalized GFP mRNA transcript levels as a measure of translation efficiency of each mRNA transcript. Each bar represents the average of 3 technical replicate transfections / qPCR quantification ratios and error bars represent their standard deviation.

Once more, GFP mRNA levels and hygromycin mRNA levels were measured in the transient pools obtained. Again, overall lower levels of mRNA were detected for all the 2G toeholds, with varying levels of hygromycin and GFP mRNA transcripts. After normalizing GFP mRNA with hygromycin mRNA levels as previously described, no differences were observed between the different 2G toeholds, which again showed lower levels of normalized GFP mRNA when compared with the 204-L-A-1G toehold and the no toehold GFP control plasmid (Figure 99).

As previously described, transcript translation efficiency for the different constructs tested was calculated by dividing the measured GFP protein (fluorescence) by the normalized GFP mRNA levels. As before, lower translation efficiencies were obtained from transcripts coming from 2G toeholds in comparison with transcripts expressed from 204-L-1G and transcripts coming from pcDNA3.1-GFP (Figure 100). Small differences were observed between the two new constructs. While 204-L-T-NC-2G showed higher translation efficiency than the 204-L-T-2G construct, the translation efficiency measured for the 204-L-T-409-2G was the same.

N44 ggtaccagagctcggatccaccggtc**gccaccatg**gtgagcaagggcgaggagctgttcaccggggtggtgccatcctggtcgagctggacgg

204-L-1G GGTACC**AGGCATAGGATG**ACAAAGGGAA---AGGAGGAAAAAA**GCCACCATG**GCGGCTGTC---GGACAAAGGGAAGGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

204-L-2G GGTACC**AGGCATAGGATG**ACAAAGGGAA**CTA**GCCGCCTTTT**TTC****GCCACCATG**GCGGCTGTC**TAC**GGACAAAGGGAAGGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

204-L-NC-2G GGTACC**AGGCATAGGATG**ACAAAGGGAA**CTA**GCCGCCTTTT**TTC****GCCACCATG**GCGGCTGTC**TACTGT**CAGCTTTCTGGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

204-L-NC-GCG-2G GGTACC**AGGCATAGGATG**ACAAAGGGAA**CTA**GCCGCCTTTT**TTC****GCCACCATG**GCGGCTGTC**GCGTGT**CAGCTTTCTGGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

204-L-409-2G GGTACC**AGGCATAGGATG**ACAAAGGGAA**CTA**GCCGCCTTTT**TTC****GCCACCATG**GCGGCTGTC**TAC**CTCATAAGATT**CG**GGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

204-L-409-GGA-2G GGTACC**AGGCATAGGATG**ACAAAGGGAA**CTA**GCCGCCTTTT**TTC****GCCACCATG**GCGGCTGTC**GGA**CTCATAAGATT**CG**GGGGGGGGGGCGCGCC
လည်ပတ်သောဓာတ်ပေါက်ကွဲသော

Figure 101 - Modular design and aligned sequences for the second-generation miRNA-toehold sequences with mutated tyrosine residue at the GFP N-terminal domain. Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlighted in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

N44 V P S S D P P V A T **Met** V S K G E E L F T G V V P I L V E L D

204-L-1G G T R H R **Met** T K G K E E K K A T **Met** A A V **G** Q R E G G G G A P V S K

204-L-2G G T R H R **Met** T K G S S R L F F A T **Met** A A V **Y** G Q R E G G G G A P V S K

204-L-NC-2G G T R H R **Met** T K G S S R L F F A T **Met** A A V **Y** C Q L S G G G G A P V S K

204-L-NC-GCG-2G G T R H R **Met** T K G S S R L F F A T **Met** A A V **A** C Q L S G G G G A P V S K

204-L-409-2G G T R H R **Met** T K G S S R L F F A T **Met** A A V **Y** L I R F G G G G A P V S K

204-L-409-GGA-2G G T R H R **Met** T K G S S R L F F A T **Met** A A V **G** L I R F G G G G A P V S K

Figure 102 - Amino acid sequence from translated miRNA-toehold sequences with mutated tyrosine residue at the GFP N-terminal domain. From the Kozak-start codon, translated Met1 corresponding to the d2eGFP ORF is highlighted in blue. The predicted translated sequence for the same ORF upstream of the Kozak-start codon is shown for all sequences.

At this stage, it was noted that the introduction of the TAC cloning site has resulted in the addition of a tyrosine in the ORF of GFP from 2nd generation toeholds. In order to assess the effect this additional tyrosine residue in the linker, variants for two of the L constructs, 204-L-NC-2G and 204-L-409-2G, in which the tyrosine codon was mutated to a Glycine (GCG) (204-NC-GCG-2G) or and Alanine codon (GGA) (204-L-409-GGA-2G) were designed (Figure 101 and Figure 102).

Changing the B/C cloning site from TAC to GGA in the 204-L-409-2G toehold did not alter the predicted secondary structure of the 5'-UTR region around Kozak-ATG, which remained in hairpin-like conformation (Figure 103). Small changes were observed in the secondary structure of 204-L-NC-GCG-2G when compared with 204-L-NC-2G, with a looser structure around the base of the hairpin and a shorter stem region (Figure 103).

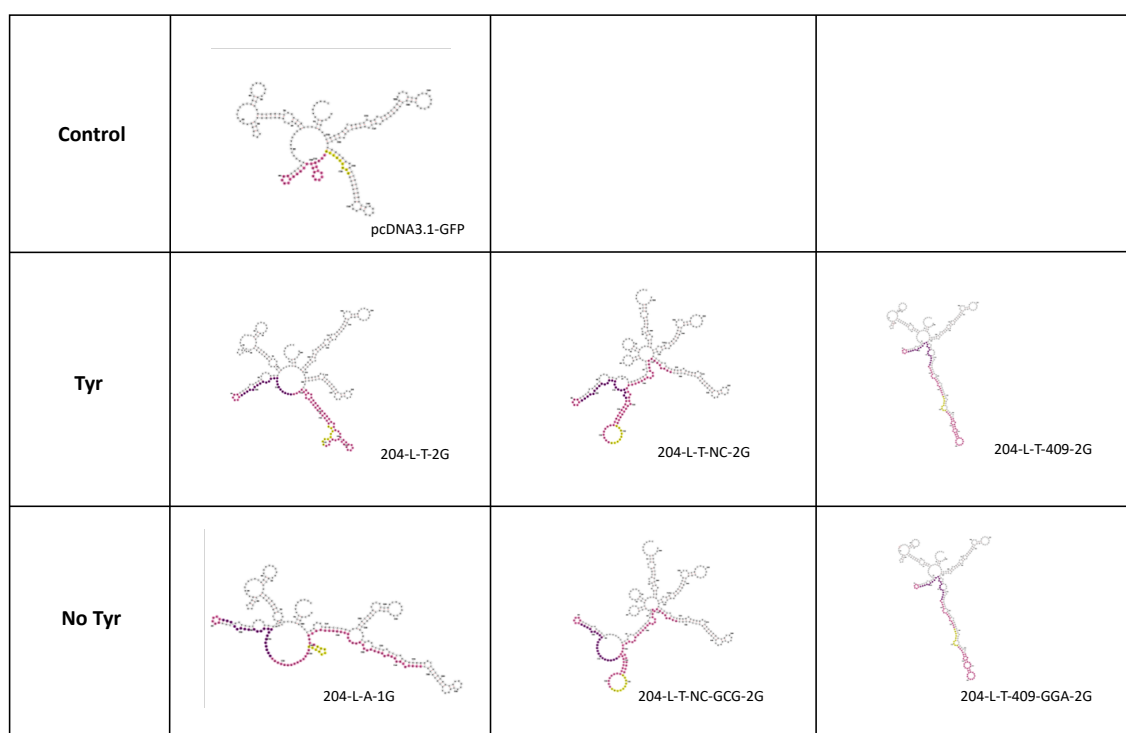


Figure 103 - Predicted secondary structures using RNA fold for the second-generation miRNA-toehold sequences with mutated tyrosine residue at the GFP N-terminal domain. Predictions include the 3'UTR-region and the first 100+ nucleotides. Relevant sequences are highlighted in colour, toehold insert (pink), miRNA-binding site (violet) and Kozak-ATG sequence (yellow).

In terms of GFP expression , mutating codons from Tyrosine (TAC) to a Glycine (GCG) or Alanine (GGA) did not show any large effect on GFP expression in comparison with the expression observed for the original 204-L-T-2G (Figure 104). These vectors continued

to express lower levels of GFP protein than the no-toehold pcDNA3.1-GFP and the 204-L-1G vectors.

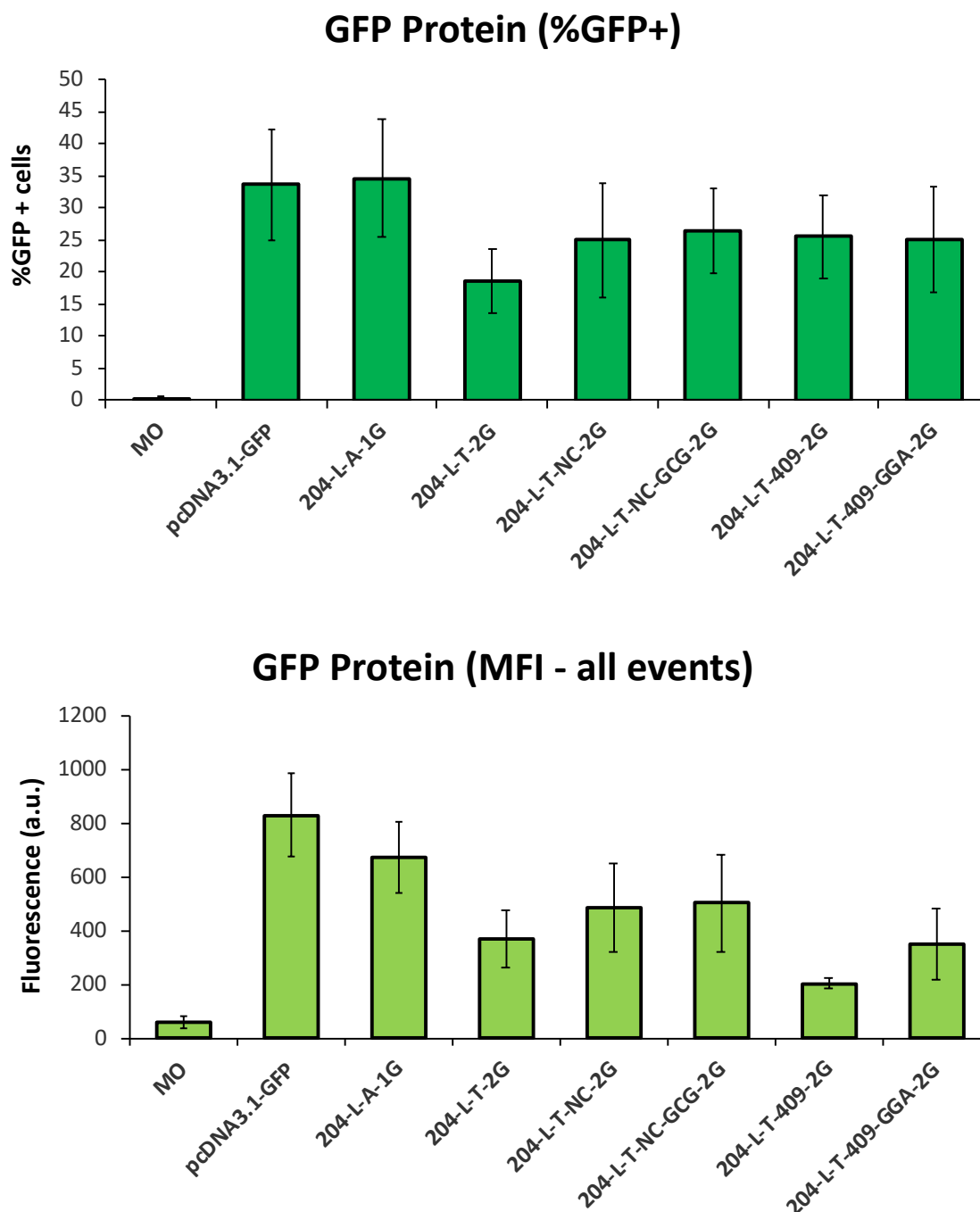


Figure 104 – Percentage of GFP positive cells and Mean Fluorescence Intensity (MFI) of the obtained cell populations (all events) 24h after transfection with different miRNA-toehold sequences with mutated tyrosine residue at the GFP N-terminal domain. A pcDNA3.1 GFP plasmid and an only Mirus™ condition were used as positive and negative controls for transfection and flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

2.11. Change of the T rich region from an A rich region in the loop

Another of the changes introduced in all the 2nd generation toeholds when compared with the 1st generation miRNA-Toeholds was the sequence in the loop region. The sequence in the loop region used in the 2nd generation toeholds was the one originally used in the 204-H-1G, which consisted of a stretch of thymidine nucleotides “TTTTTTTT”. However, some of the 1st generation toeholds, in particular 204-L-A-1A, had a different loop region consisting of an A-rich tract. To evaluate the effect of sequence changes in this region, new constructs using a stretch of T residues. To evaluate the effect of a variable loop region (fragments B), variants for the H-2G and L-2G constructs were constructed using the two different loop regions. Vectors containing the A-rich loop will be referred with the A code in contrast to the original 2G constructs that were named with the T code (Figure 105).

204-H-T-1G	GGTACCAGG CATAGGATG ACAAAGGGAA--- SCCGCCTTTTTTT GCCACCATG GCGGCTGTCTTCCTTTGTCAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-H-T-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA SCCGCCTTTTTTT GCCACCATG GCGGCTGTCT TACT TCCCTTTGTCAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-H-A-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAAAAAAA GCCACCATG GCGGCTGTCT TACT TCCCTTTGTCAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-H-A-S-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAATAAAA GCCACCATG GCGGCTGTCT TACT TCCCTTTGTCAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-H-A-INS-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAATAAAA GCGGCTGTCT GCCACCATG TACT TCCCTTTGTCAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-L-A-1G	GGTACCAGG CATAGGATG ACAAAGGGAA--- AGGAGGAAAAAAA GCCACCATG GCGGCTGTCT GGACAAAGGGAA AGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-L-T-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA SCCGCCTTTTTTT GCCACCATG GCGGCTGTCT TAC GGACAAAGGGAAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-L-A-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAAAAAAA GCCACCATG GCGGCTGTCT TAC GGACAAAGGGAAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-L-A-S-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAATAAAA GCCACCATG GCGGCTGTCT TAC GGACAAAGGGAAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>
204-L-A-INS-2G	GGTACCAGG CATAGGATG ACAAAGGGAA GTA AGGAGGAATAAAA GCGGCTGTCT GCCACCATG TAC GGACAAAGGGAAAGGGGGGGGGGCGCGCC <i>လောင်စာသေစာသေစာသေစာသေစာ</i>

Figure 105 - Modular design and aligned sequences for the second-generation miRNA-toehold sequences with variant loop regions (green). Underlined regions correspond to the miRNA binding site (5'-end) and complementary/non-complementary region (3'-end). The endogenous miRNA region and its degree of complementarity (bold for mismatching nucleotides) is shown in pink. The Kozak-start codon sequence for each construct is highlighted in light blue. Nucleotides inserted for cloning purposes are shown in bold black.

This was done anticipating that this region could be responsible for the changes observed in secondary structure. T-variants with a Kozak-ATG region shifted toward the 3' and T-variants containing a Stop codon in the loop region were also constructed (Figure 105).

The tested variants chosen to test the effect of swapping the loop region from T to A were the H and L constructs containing a stop codon to stop leaky translation from the in frame upstream ORF (uORF) starting at the start codon part of the cgr-miR-204-3p binding site sequence (204-L/H-A-S-2G), and the H and an L construct variant in which the Kozak-ATG was shifted to the 3' end of the B fragment so it would lie buried deeper into the stem region in the event of a predicted hairpin-like secondary structure (204-L/H-A-INS-2G).

Secondary structure predictions for the 2G "T" variants compared with the original 2G "A" are shown in Figure 106. Changing the loop region upstream of the Kozak-ATG sequence from "T"'s to "A"'s led to changes in the predicted secondary structures for the L constructs. Whereas H-A-2G constructs continued to have predicted hairpin-like secondary structures, the L-A-2G constructs showed "linear" secondary structures such as the ones displayed by predictions done on the first generation 204-L-1G vectors.

In terms of GFP expression, swapping the T-rich region by an A-rich drastically increased the detected GFP signal independently of their hairpin-like or linear secondary structure (204-X-T-2G vs 204-X-A-2G in Figure 107). All constructs except 204-L-T-2G showed similar %GFP positive cells. 204-L-A-2G showed an MFI comparable to 204-L-A-1G and pcDNA3.1-GFP. Surprisingly, 204-H-A-2G also showed a high MFI comparable to the no toehold and the 204-L-A-1G and 204-L-A-2G constructs. This indicates that the T-rich loop region might be the cause of the previously observed GFP repression rather than the secondary structure predicted for the toehold region.

Regarding the position effect of the Kozak-ATG sequence and the presence/absence of the upstream stop codon, among the "A" variants, the stop codon constructs S-2G and INS-2G showed lower GFP expression, once more suggesting that on one uORF (that starts in the miRNA binding site) is in part responsible for the observed GFP signal. The A-INS-2G constructs showed even lower GFP expression than the A-S-2G constructs independently of the secondary structure (H or L) that was predicted. This indicates the absence of a relationship between secondary structure and Kozak-ATG position. In other

words, rather than an effect of burying the Kozak-ATG sequence in a hairpin being at the source of the observed downregulation of GFP expression, it would seem that the sequence around the Kozak-ATG region, which is what also changed with the position swap is responsible for the observed decrease in expression.

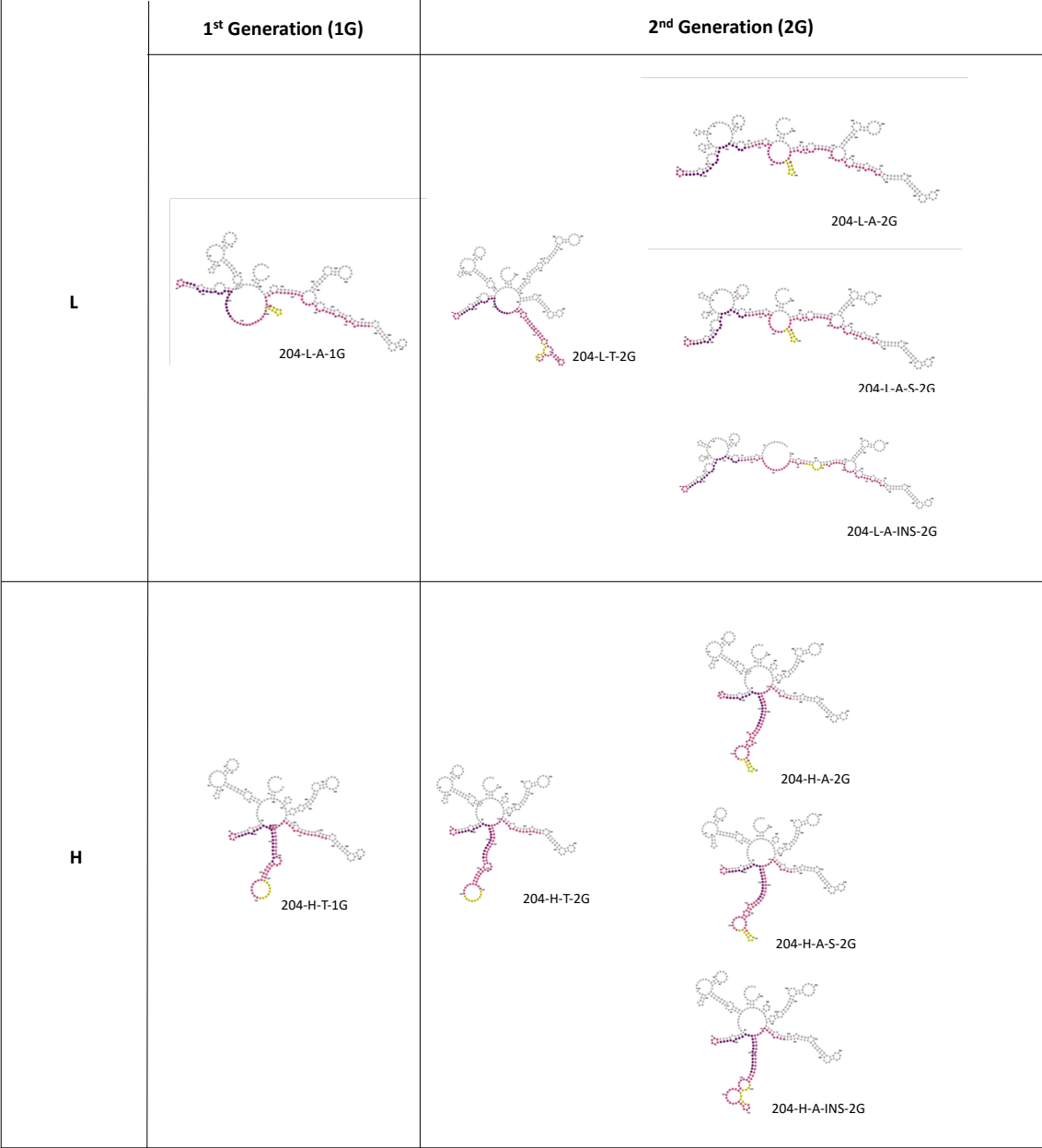


Figure 106 - Predicted secondary structures using RNA fold for the miRNA-toehold sequences with variant loop regions. Predictions include the 3'UTR-region and the first 100+ nucleotides. Relevant sequences are highlighted in colour, toehold insert (pink), miRNA-binding site (violet) and Kozak-ATG sequence (yellow).

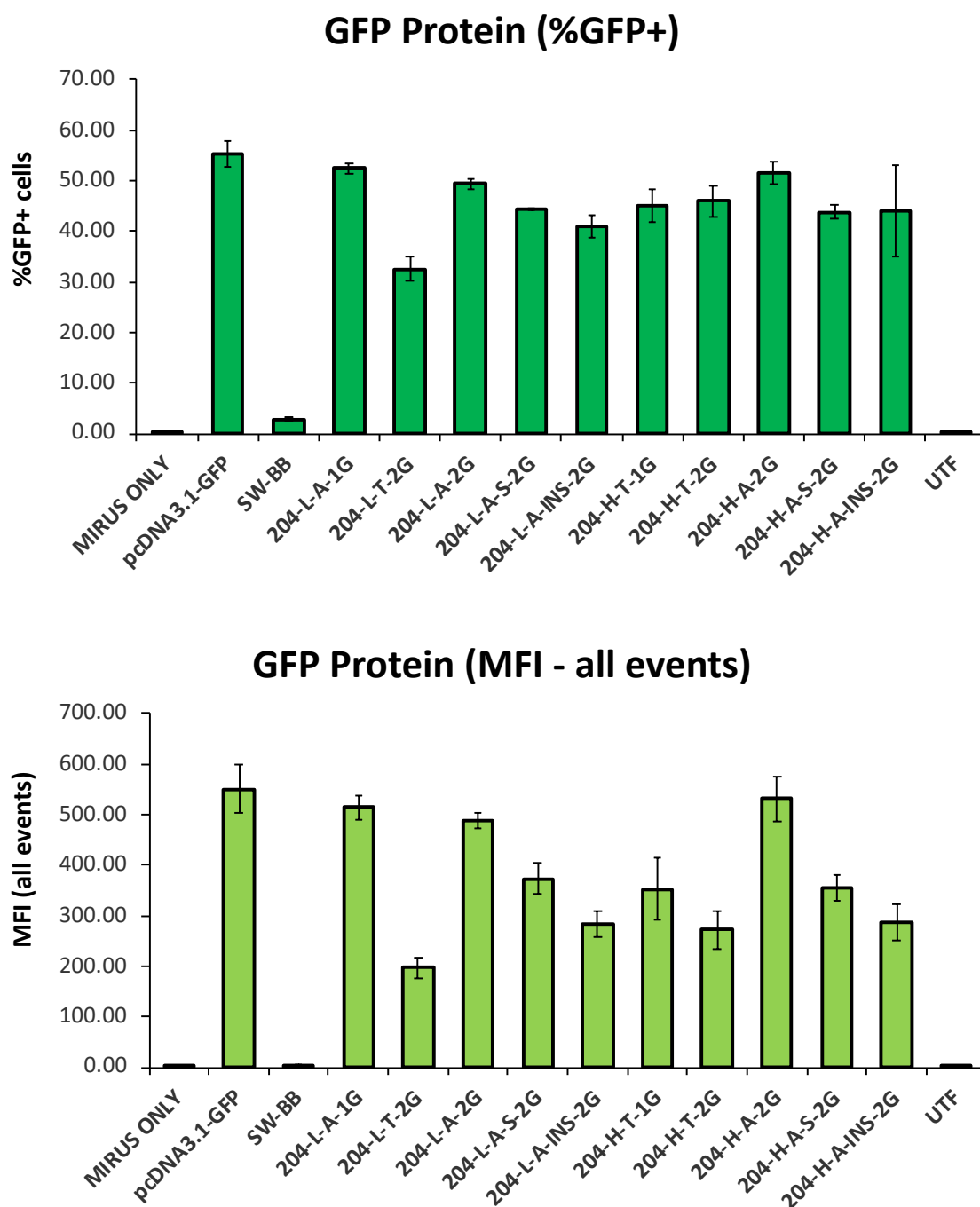


Figure 107 – Percentage of GFP positive cells and Mean Fluorescence Intensity (MFI) of the obtained cell populations (all events) 24h after transfection with miRNA-toehold sequences with variant loop regions. A pcDNA3.1 GFP plasmid (showed in small chart) and an only Mirus™ condition were used as positive and negative controls for transfection and flow cytometry gating. Each bar represents the average of 3 technical replicate transfections and error bars represent their standard deviation.

2.12. Are changes in the 5-UTR affecting mRNA, thus decreasing GFP protein expression?

5'-UTRs play various roles in the regulation of translation of mRNA transcripts, one of the being their potential effect on mRNA transcript stability. In order to investigate mRNA stability from transcripts expressed from the different toehold containing transcripts, an mRNA stability assay was optimised and used to compare transcript stability from the no toehold UTR control expressed from the pcDNA3.1-GFP vector and the 204-H-T-2G vector, as this plasmid was the one showing lower levels of GFP expression from the 2G constructs.

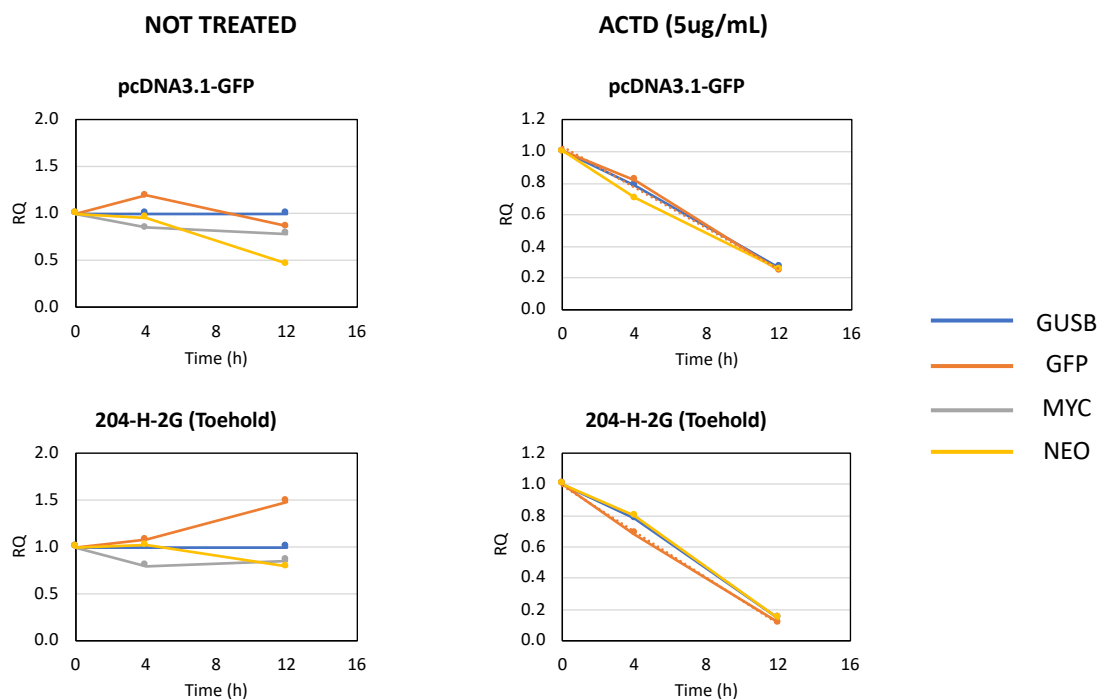


Figure 108 - RNA decay assay on 204-H-T-2G and pcDNA3.1-GFP to evaluate differences in mRNA stability derived from the addition of the toehold 5-UTR. mRNA levels for GFP, the selection marker in the filler plasmid Neomycin and the endogenous GAPDH and c-MYC from untreated samples and samples treated with 5ug/mL of actinomycin D were analysed by RT-qPCR. Relative quantification for each mRNA transcript was performed using the $2^{-\Delta Ct}$ method: $RQ = 2^{-\Delta Ct}$ where $-\Delta Ct = (Ct_{target})_{Time\ 1} - (Ct_{target})_{Time\ 0}$. Each datapoint represents the average of 3 technical replicate transfections pools and analysed as a single sample by qPCR.

In short, CHO-K1 cells were seeded at 250.000 cells/ well in duplicate sets of 6-well plates and incubated overnight at 37°C. The following day, cells were transfected as previously described using the test plasmids and incubated for another 24h. Duplicate transfections were done per plasmid. 24h post-transfection, media was replaced with

fresh DMEM:F12 media +5% FBS + 5 ug/mL Actinomycin D. A parallel set of 6-well plates was set up with media without Actinomycin D as an untreated control.

Cells (different wells) were sampled at different time points (4-12h) by removing the actinomycin D containing media and directly adding Trizol® reagent to the well to extract RNA. RNA samples were purified and mRNA levels for GFP, the selection marker in the filler plasmid Neomycin and the endogenous GAPDH and c-MYC were analysed by RT-qPCR. Relative quantification for each mRNA transcript was performed using the $2^{-\Delta Ct}$ method as described in Methods Section 1.3.4.

GFP mRNA transcripts from the two different plasmids did not exhibit differences in terms of their mRNA decay rates, indicating that the presence of the toehold did not have an impact on mRNA stability (**Figure 108**). Based on this, differences observed in GFP expression from the two tested vectors were most likely not related with differences in the stability of the mRNA transcripts.

3. Discussion

The original idea of this project was to develop a novel miRNA-responsive element, the miRNA-toehold switch, which could repress transgene expression and de-repress it in the presence a particular miRNA. The first prototype of these novel miRNA-switches was based on previously described toehold switches (Green *et al.*, 2014). It must be acknowledged that while the design principles outlined in Green *et al.* 2014 were followed, those had been defined in a bacterial system and little was done to optimize the original constructs. Early experiments were done more as a proof of concept than a real optimization study aiming to define the best toeholds one could make to inhibit the expression of the reporter gene. After testing the initial constructs in various cell lines, early data indicated that different toehold miRNA-switches were able to drive different levels of GFP repression, providing a promising starting point for the project to move forward.

After testing the possibility of using miRNA mimics to switch on transgene expression, it became apparent that while toehold switches were specifically targeted by complementary miRNAs, this further repressed transgene expression instead of up-regulating it. RT-qPCR analysis showed that miRNAs targeting the binding site present in the miRNA-toehold switch were driving degradation of the mRNA transcript

independently of the secondary structure predicted to form around it and most likely through the canonical pathway, as mRNA degradation was shown to be dependent on perfect matching of both the seed region and the 10/11nt of the miRNA with the miRNA binding site. Without ruling out the possibility that secondary structure would have an effect on translation, it was also considered that the presence of a miRNA-binding site on the 5'UTR could result in the binding of loaded miRISC and the subsequent steric blocking of the ribosome scanning mechanism. miRNA-toehold targeting by miRNA mimics was shown with the downregulation of GFP expression both at protein and mRNA transcript levels. Assuming that RISC loaded with bulged miRNA mimics are targeting the GFP transcripts in the same way, the fact that GFP expression levels did not change between untransfected pools and pools transfected with bulged mimic controls indicates that miRISC binding at the miRNA-binding site is not enough to repress translation. Still guided by the hypothesis that secondary structure, namely the presence of a hairpin around the Kozak-start codon region, was in some way repressing GFP expression, further sequence changes were investigated in a second generation of miRNA-toeholds. This time, in addition to monitoring protein levels by flow cytometry, mRNA levels were also measured. The analysis of GFP expression from a variety of miRNA-toehold sequences, all of them predicted to form hairpin-like structures, showed the high complexity of the tested system. All the newly generated miRNA-toeholds showed reduced expression of GFP. Combined data from RT-qPCR and actinomycin D treatment showed that the lower GFP expression was more likely due to less efficient translation of the mRNA transcripts carrying a miRNA-toehold UTR rather than the latter affecting mRNA stability, another factor that could be reducing expression of GFP. Comparing expression from hairpin constructs versus linear constructs, the presence of the hairpin itself did not seem to be the main driver of the observed GFP downregulation. Instead certain sequence features were shown to have a greater impact on GFP expression. For example, the presence of a particular sequence in the loop region led to a strong up-regulation of GFP expression independently of other sequence modifications. In addition, the presence of an upstream ORF frame was also shown to be contributing to the detected GFP signal, with the presence of upstream stop codon leading to a reduction of GFP protein but no apparent effect on GFP mRNA levels.

In the context of translation in eukaryotic systems such as CHO cells the obtained results are not so surprising. Cap-dependent translation initiation in eukaryotic cells is a complex process involving several steps, starting with the recruitment of the 40S small ribosomal subunit and its associated eukaryotic initiation factors (EIFs) to the 5' cap of the mRNA transcript. The EIF4F cap-binding complex consists of eIF4E, which interacts with the scaffolding initiation factor eIF4G and the RNA helicase eIF4A (Leppek, Das and Barna, 2018). Precisely, the presence of this RNA helicase can be seen as the perfect countermeasure to avoid the possibility of repressing translation by means of secondary structures near the start codon. It is true that previous studies in yeast have shown that hairpins placed in the 5'-UTR can be used to predictably modulate transgene expression (Babiskin and Smolke, 2011; Lamping, Niimi and Cannon, 2013; Weenink *et al.*, 2018). In mammalian systems, it has been shown that different levels of translation inhibition can be obtained by placing hairpins in the 5'-UTR (Babendure, 2006). However, and consistently with the findings in this thesis, all these studies show that transgene expression modulation is dependent on the thermodynamic properties of the hairpins as well as to their position in relation to the 5'-cap. Hairpins in close proximity of the 5'-cap were able to inhibit protein expression while hairpins in more distant UTR positions from the 5'-cap failed to do so (Babendure, 2006). In this study, miRNA-toeholds were cloned 88 nucleotides from the putative transcription start site, a much greater distance than those described in literature.

In addition to structural features affecting translation, several studies have shown the existence of sequence-determined motifs able to recruit specific RNA binding proteins that can modify RNA molecules and affect translation. A well-studied example of those are the Iron Regulatory Elements (IRE), which are found in the UTRs of several genes encoding proteins involved in iron metabolism such as the iron storage protein ferritin, and recruit RNA binding proteins that sterically block the assembly of the initiation complex to the cap region thus inhibiting translation (Gray and Hentze, 1994; Muckenthaler, Gray and Hentze, 1998). Another example of a sequence-dependent mechanism that regulates translation of mRNA transcripts are N⁶-methyladenosines (m6A) that require specific sequence motifs. These RNA modifications are the most abundant modifications in mammalian mRNAs (Wei, Gershowitz and Moss, 1975) and several studies have shown that m6A modifications on the 5'-UTR of genes have large effects on protein expression either by regulating RNA protein interactions (Meyer *et*

et al., 2015) or by affecting transcript stability (Costello, Lao, Barron, *et al.*, 2019). None of those motifs were considered during the design of the miRNA-toehold sequences used in this work and therefore, if present, could be inadvertently modified by some of the changes performed on the miRNA-toeholds thus impacting GFP expression. Further work would be necessary to understand how the tested miRNA-toehold switch affected GFP expression and define the actual mechanisms at play. However, one of the changes that showed a more dramatic effect on GFP expression, namely the swap of a T-rich sequence in the loop by an A-rich region deserves some attention. A priori, it was hypothesized that a sequence change from a T-rich to an A-rich region *would have a major impact* based on the rationale that both sequences have a very low GC content, and both sequences were chosen to allow the formation or not of a hairpin secondary structure. However, the presence of AU-rich motifs needs to be discussed in a broader context than simply looking at what secondary structure could be formed. AU-rich motifs, also known as AU-rich elements (ARE) are a well-known class of RNA motifs that have been shown to affect mRNA stability and translation efficiency. While AREs were first described in the 3'UTR of mRNA transcripts and as being involved in promoting mRNA degradation (Chen and Shyu, 1994), more recent studies have described the role of AREs in the 5'-UTR in enhancing translation of different genes (Millard *et al.*, 2000; Ang *et al.*, 2019). An interesting factor known to bind to AREs is the HuR protein. This protein was first described as a decay delaying factor (Peng *et al.*, 1998), and has also been shown to mediate control of miRNA-mediated repression of translation (Meisner and Filipowicz, 2010). U-rich RNA motifs have also been shown to be targets of the heterogeneous nuclear ribonucleoprotein C proteins (Görlach, Burd and Dreyfuss, 1994). These proteins are also very relevant in the context of the results observed with the miRNA-toehold switches as they act as nucleating particles for the 40S ribosomal subunit during translation initiation (Beyer *et al.*, 1977). Poly-adenylated sequences are also very important in the regulation of translation when present in the 3'-UTR of mRNA transcripts. Indeed, they are essential for the circularization of the mRNA transcript by the interaction of the Poly(A) binding protein (PABP), which binds the 3' poly(A) tail, and the eIF4G at the cap (Leppek, Das and Barna, 2018). While bound to the 3'-UTR, PABP has been shown to be critical for efficient translation initiation (Kahvejian *et al.*, 2005), reports have also shown that the binding of PABP to a poly(A)-rich element in the 5'-UTR can inhibit translation (Bag, 2001; Melo, De Melo Neto and Martins De Sá, 2003).

While it is reasonable to hypothesize that regulation by AU-rich regions and their related factors could contribute to explain, at least partially, some of the observed effects of miRNA-toehold switches, a more systematic investigation of miRNA-toehold mutants would be needed to start digging deeper into potential RBP interactions happening there.

The other modification that showed marked effects on GFP expression between the different toeholds tested was the stop codon that putatively interferes with the upstream ORF detected in the binding site for miR-204, in frame with the ORF of the GFP. The reduction in GFP expression observed upon addition of a stop codon in all the different tested combinations strongly suggests that translation was occurring from this uAUG codon. Approximately half of human mRNAs have AUGs upstream of the primary AUG (pAUG) and these uORFs have previously been shown to regulate translation and have been associated with reduced protein expression (Calvo, Pagliarini and Mootha, 2009). As reviewed in the introduction, transgene expression control systems have been developed combining uORF and specific antisense oligonucleotides that target them and “switch ON” expression by blocking translation from these upstream sites (Liang *et al.*, 2016). While uORFs have been widely associated with downregulation of translation leading to lower protein yields, the presence of uORF in the 5-UTR can also enhance translation by promote re-initiation (Young and Wek, 2016). While in the system investigated in this work, the presence of an uORF in-frame with the reporter gene was shown to cause the up-regulation of GFP expression with an apparent increase in translation efficiency, there are probably many other factors involved, some of them already discussed, that make it difficult to identify a single, specific reason for the observed results. Clearly what seem to be minor changes in a short sequence region can have major implications in the molecular events at play.

4. Conclusion

In summary, what started as an attempt to develop a miRNA-toehold switch able to repress transgene expression and be “switched on” by the presence of a specific miRNA, ended up highlighting several factors that must be taken into account in future attempts to develop 5'-UTR switches in mammalian systems. While the impact of the secondary structure of mRNA translation must be taken into account, the data presented in this chapter shows that the sequence context both upstream and downstream of the Kozak-

ATG needs also to be carefully considered. Moreover, if miRNA-responsive elements are to be used in the 5'-UTR, both the sequence and structural environment of these has to be taken into account in the design. Considering that miRNA-responsive elements will necessarily be linked to a useful endogenous miRNA trigger, as previously discussed in Chapter 1, it is unlikely that a one-fits-all MRE design will be achieved, rather each element will need to be bespoke for a particular application.

CHAPTER 6

Concluding Remarks and Future Work

The work performed for this thesis aimed to develop new tools for miRNA-based control of transgene expression in Chinese Hamster Ovary cells in the context of the production of biopharmaceuticals. Under this main goal, three research projects, one per chapter of this thesis, were carried out.

The first one aimed to identify endogenous miRNA expression profiles that could be used as effector elements for controlled gene expression. Two different datasets were analysed in order to identify miRNAs with expression profiles that could be related to specific culture conditions or phases in order to use them as effector molecules in a transgene expression control system.

The investigation, initially of a pre-existing miRNA expression dataset obtained from collaborators aimed to identify miRNAs, whose expression could be related to ammonia metabolism, by comparing differential expression profiles between cells growing in glutamine containing media and cells adapted to glutamine free conditions. The ultimate goal of this project was to find a subset of miRNAs that could be controlled by addition/removal of glutamine, or whose expression profile could be correlated to the production of ammonia in the bioreactor. Two miRNA candidates, hsa-miR-1275 and cgr-miR-1195, were successfully identified and validated. However, their low level of expression made them unlikely to be effective switches. Some of the challenges encountered during the analysis of this first dataset were related to the difficulty in determining a suitable target mining strategy that would result in robust identifications to move forward to the validation stage without having to test a large number of candidates.

Collectively, the results of this work show that while available data might still have potential and could still be used to identify miRNAs useful for sensor applications, new tailor-made profiling studies will be necessary in order to generate data to serve that purpose. Not only to avoid situations such as the ones encountered here, but also to guarantee that any potential findings could be readily transfers to an industrial setup.

Some of the lessons learned during this work should be incorporated in future work involving the use of pre-existing microarray-based datasets. Carefully merging microarray miRNA expression data with the most up to date annotation as well as cross checking it with other profiling datasets (e.g. RNASeq) should be the first steps in future target mining studies. Carefully considering expression levels at an early stage would

also be recommended so as to discard candidates for which evidence of low expression is available. As a result of this study, a glutamine-free model with CHO-K1 cells adapted to glutamine free media and matched glutamine dependent parental cell line, was generated. Further work could be performed using those cells with the aim of generating an RNASeq based miRNA profiling dataset that could complement or even replace the microarray based one. While the comparison of glutamine dependent cells and glutamine-free cells in a time course experiment would be useful, more specific experimental interventions such as glutamine or ammonia spikes could be added. This would allow further exploration of the relationship between miRNA changes and glutamine/ammonia changes. Further work should prioritise the use of RNASeq over other technologies, at least for the initial identification of targets as made apparent during the investigation of the second dataset presented in Chapter 3. While it was only possible to perform a preliminary data analysis, due to the late arrival of the data, it was possible to identify three main clusters of miRNAs with expression patterns that could easily be matched to distinct growth phases. Further work on this dataset could involve the individual investigation of each of the miRNA candidates included in each of the three identified clusters as well as the validation of the most interesting miRNA or subsets of miRNAs. From a bioprocessing perspective, miRNAs included in the cluster showing timed up regulation at the interface of exponential and stationary phase should be prioritised. As mentioned before, the availability of abundance data and the possibility to reliably cross check data with other published datasets would most likely increase the success rate of identification/validation.

The work presented here has also highlighted the importance of considering the biological context for the different miRNAs or miRNA clusters investigated. While this type of investigation usually relies on target prediction and pathway analysis, as performed in this work, it will be essential to identify the endogenous function that potential miRNA switches have within the cell regulatory networks. Future studies towards developing miRNA-based control systems should include a pre-assessment of the function of particular miRNAs or groups of miRNAs in order to determine their suitability as molecular switches. In the context of biopharmaceutical production, miRNAs that affect desirable phenotypic traits such as growth, productivity or product quality should be deprioritized, while miRNAs that have no apparent links with these critical cell culture attributes should be moved forward and investigated first. Once

more, RNASeq technologies are probably the best choice, as these allow the sequencing matched miRNA and mRNA samples, which would provide further evidence supporting the predicted miRNA functions, thus allowing to perform a more informed filtering of miRNA candidates to test as switches.

The next steps after having identified attractive miRNA candidates that could be used as switches under specific conditions would be to link their expression to the expression of a transgene of interest. In theory this would be achievable by including miRNA responsive elements in the genetic construct used to express the transgene. The work included in Chapter 4 aimed precisely to do this by investigating the use of miRNA sponges as miRNA responsive elements that can be used to translate miRNA expression changes into controlled transgene expression. With the aim to develop a temperature induced miRNA-based transgene control system, miRNA expression upon temperature shift, a commonly used intervention in CHO-based bioprocesses used to improve cell productivity and extend cell viability, was investigated. Besides its current use in the biopharmaceutical industry context, this model system was deemed very interesting due to the simplicity of changing the temperature of the reactor, and the extensive research that has been carried out around it. In a first stage, miRNA sponges were tested by using a combination of synthetic and endogenous miRNA targets that allowed the optimisation of the expression system and the methods required to test a potential miRNA-based transgene expression control system. In a second stage, using an RNASeq profiling experiment as starting point, several miRNA candidates that showed expression changes upon temperature shift were validated and used to test the use of miRNA sponges in a temperature dependent transgene expression control system. Collectively, these results showed that, while the miRNAs can modulate transgene expression in a constitutive manner, as several previous studies had already proved, translating dynamic changes of miRNA expression into changes of mRNA expression or protein expression will require more than simply having miRNA responsive elements linking the two.

In addition, all the preliminary work performed in Chapter 3 highlighted certain methodological challenges, for example, the difficulty of working with a transient expression system and then moving things to the more relevant stable expression conditions, or the challenges arising when attempting to control experiments for things

like transfection efficiency and accurately established links between miRNA expression changes and transgene expression changes. While the systemic effect of temperature shift was clearly one of the sources of the challenges encountered, the work presented here shows that future work should include modifications in the vector and experimental designs should be including elements aiming to overcome them. For instant, the use of a double reporter vector system, with one of the transgenes acting as an endogenous control, free of repression, and being used to normalize the data on a per cell basis, could be a suitable strategy to bypass the difficulties related to controlling the different transient experiments. In the stably expressing pools, that double reporter could also act as internal control and help to evaluate if the observed changes in expression are indeed linked to the changes in miRNA expression or are a side effect of other changes caused by the reduction of temperature.

Additionally, the learnings from the work presented here show that rather than approaching the problem using a traditional bottom-up approach, the development of miRNA genetic switches might be more effectively tackled using a top-down approach. In other words, rather than investigating each component of a potential miRNA switch on its own then combining the parts to build and evaluate a more complex system, it would now be possible to start with a complete prototypic switch build for a large number of miRNA candidates and test it in a high-throughput fashion in a bioprocess, aiming to identified the ones that deliver functional outputs that can be selected and refined. In this direction, a more ambitious but interesting follow-up project, intertwining the profiling work and the sensor optimization work could involve the generation and use of a CHO specific library of miRNA sensors. This approach would allow testing the possibility of obtaining detectable miRNA/miRNA sensor interactions and rapidly identify pairs that lead to detectable changes in expression from the start. Moreover, this library approach could not only be tested in transient, but transfected pools could be selected with antibiotic up to the stage were stable pools would be obtained. Moving to stable conditions would also allow testing for transgene copy number and ensure that the same number of integrations per cell per construct, ideally one, is obtained. This would avoid effects derived from different gene doses and would allow a more accurate determination of the actual sensing capacity of each single miRNA candidate. These libraries of CHO-stable pools expressing miRNA sensors could then be used to screen potential interventions such as temperature shift, addition of

metabolites, time or others and readily identify those pairs of miRs/miR-sensors that are able to respond as required in a particular context. An alternative and possibly complementary approach for that would be the use of well characterized landing pads for stable and robust expression of transgene, in line with the current developments in the field of targeted integration discussed in Chapter 4. As mentioned already, while this might provide the opportunity to test a large number of miRNA sensors in a very controlled context, it is possible that this limits their applicability in the real world, as no platform “controllable” cell line could be built if the specific requirements for proper miRNA based control clash with other cell desirable phenotypes such as fast growth and high productivity.

The third part of this thesis attempted the development of a novel miRNA responsive element, the miRNA-toehold switch, which should be able to constitutively repress transgene expression and be de-repressed in the presence of a particular endogenous miRNA. Whilst the addition of these miRNA-toehold switches in the 5'UTR region proved useful to obtain reduced levels of transgene expression, no additional ON/OFF functionality was obtained. This work, together with the findings already discussed before, reinforce the idea that while the presence of miRNA responsive elements in the UTR region allows to link miRNA and mRNA/protein expression in CHO cells, there are still many molecular mechanisms at play that are not well understood, and therefore, some knowledge is missing to fully interpret results and use them to build functional molecular tools for specific application. Collectively, the results presented in this thesis show this might not be achievable due to the increased complexity of the eukaryotic translational machinery.

The take home message from the work performed in this project is that subtle changes in the UTR region can drive drastic effects in transgene expression, while local secondary structure, at least in our hands, seem to play a minor role. Some of the work in yeast and other mammalian systems has demonstrated the possibility to tune gene expression engineering the secondary structure of the 5'-UTR. However, to my knowledge UTR engineering has only allowed the generation of populations with defined expression levels of transgene, but no work to date has achieved “proper control”. Further work should explore the possibility of combining secondary structure-based transgene expression modulation with miRNAs building on the concept of

toehold-like switches. It would be interesting to explore if by modifying the structural environment around the miRNA binding site it is possible to obtain enhanced or impaired miRNA binding, thus creating, for example, more or less efficient miRNA sponges depending on their secondary structure. Future work deriving in particular from this third project should definitely explore further the different elements that were shown to have a strong effect on transgene expression such as the modifications performed around the loop region, as well as the use of upstream ORFs to design switch elements, although the latter has already been investigated. It would also be interesting to further explore the apparent up regulation of cgr-miR-204 in cells stably expressing the 1G hairpin-like 204 construct. Several unsuccessful attempts were made to detect both the endogenous cgr-pre-miR 204 and the putative synthetic 204-pre-miR resulting from the hypothetical processing of the toehold hairpin. Further future work could involve the generation of larger libraries of miRNA-toehold switches, including a broader range of sequence and structural features, thus reproducing the already discussed studies performed in yeast cells. These studies would allow to systematically identify functional sequence and structural feature in a CHO cell-based system.

In summary, the projects undertaken for this thesis have advanced three of the important streams that need to converge for the development of miRNA-based transgene expression control systems in the future. More miRNA profiles and signatures specific for particular CHO-bioprocess relevant situations will surely be identified as more and more efforts are put towards sequencing industry relevant experimental setups like the one presented here. Regarding the molecular biology tools and methods necessary to build these miRNA expression profiles into industrially applicable genetic circuits, collectively the results provided in this thesis indicate that most likely, miRNAs on their own will not be useful switch molecules and approaches combining them with other genetic elements are most likely required in order to build robust switch systems for controlled gene expression in the future. More generally, and as seen in the last project, while more rational engineering should be tested, further work to dissect and understand the basic endogenous molecular mechanisms controlling gene expression should be encouraged even in a very applied field such as CHO cell-based research. This will be critical to be able to develop novel synthetic systems without accidentally colliding with what nature has already put in place.

BIBLIOGRAPHY

- Al-Fageeh, M. B. and Smales, C. M. (2009) 'Cold-inducible RNA binding protein (CIRP) expression is modulated by alternative mRNAs', *Rna*, 15(6), pp. 1165–1176. doi: 10.1261/rna.1179109.
- Al-Fageeh, M. B. and Smales, C. M. (2013) 'Alternative promoters regulate cold inducible RNA-binding (CIRP) gene expression and enhance transgene expression in mammalian cells', *Molecular Biotechnology*, 54(2), pp. 238–249. doi: 10.1007/s12033-013-9649-5.
- Altamirano, C. *et al.* (2000) 'Improvement of CHO Cell Culture Medium Formulation: Simultaneous Substitution of Glucose and Glutamine', *Biotechnology Progress*, 16(1), pp. 69–75. doi: 10.1021/bp990124j.
- Amann, T. *et al.* (2019) 'Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles', *Metabolic Engineering*, 52, pp. 143–152. doi: 10.1016/j.ymben.2018.11.014.
- An, C., Trinh, V. B. and Yokobayashi, Y. (2006) 'Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer – small molecule interaction Artificial control of gene expression in mammalian cells by modulating RNA interference through aptamer – small molecule i', *Rna*, 12, pp. 710–716. doi: 10.1261/rna.2299306.The.
- Ang, Z. *et al.* (2019) 'Novel AU-rich proximal UTR sequences (APS) enhance CXCL8 synthesis upon the induction of rpS6 phosphorylation', *PLoS Genetics*, 15(4), pp. 1–28. doi: 10.1371/journal.pgen.1008077.
- Ausländer, S. *et al.* (2014) 'A general design strategy for protein-responsive riboswitches in mammalian cells', *Nature Methods*, 11(11), pp. 1154–1160. doi: 10.1038/nmeth.3136.
- Babendure, J. R. (2006) 'Control of mammalian translation by mRNA structure near caps', *RNA*, 12(5), pp. 851–861. doi: 10.1261/rna.2309906.
- Babiskin, A. H. and Smolke, C. D. (2011) 'A synthetic library of RNA control modules for predictable tuning of gene expression in yeast', *Molecular Systems Biology*. Nature Publishing Group, 7(471), pp. 1–15. doi: 10.1038/msb.2011.4.
- Bag, J. (2001) 'Feedback inhibition of poly(A)-binding protein mRNA translation: A possible mechanism of translation arrest by stalled 40 S ribosomal subunits', *Journal of Biological Chemistry*, 276(50), pp. 47352–47360. doi: 10.1074/jbc.M107676200.
- Barron, N. *et al.* (2011) 'Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7', *Journal of Biotechnology*. Elsevier B.V., 151(2), pp. 204–211. doi: 10.1016/j.jbiotec.2010.12.005.
- Barron, Niall *et al.* (2011) 'MicroRNAs: Tiny targets for engineering CHO cell phenotypes?', *Biotechnology Letters*, 33(1), pp. 11–21. doi: 10.1007/s10529-010-0415-5.
- Bartel, D. P. (2009) 'MicroRNAs: Target Recognition and Regulatory Functions', *Cell*, 136(2), pp. 215–233. doi: 10.1016/j.cell.2009.01.002.

- Bartel, D. P. (2018) 'Metazoan MicroRNAs', *Cell*. doi: 10.1016/j.cell.2018.03.006.
- Beilstein, K. *et al.* (2015) 'Conditional Control of Mammalian Gene Expression by Tetracycline-Dependent Hammerhead Ribozymes', *ACS Synthetic Biology*, 4(5), pp. 526–534. doi: 10.1021/sb500270h.
- Beisel, C. L. *et al.* (2008) 'Model-guided design of ligand-regulated RNAi for programmable control of gene expression', *Molecular Systems Biology*, 4(1), p. 224. doi: 10.1038/msb.2008.62.
- Beisel, C. L. *et al.* (2011) 'Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing', *Nucleic Acids Research*, 39(7), pp. 2981–2994. doi: 10.1093/nar/gkq954.
- Beyer, A. L. *et al.* (1977) 'Identification and characterization of the packaging proteins of core 40S hnRNP particles', *Cell*, 11(1), pp. 127–138. doi: 10.1016/0092-8674(77)90323-3.
- Biggar, K. K. and Storey, K. B. (2014) 'Insight into temperature-dependent microRNA function in mammalian hibernators', *Temperature*, 1(2), pp. 84–86. doi: 10.4161/temp.29656.
- Bloom, R. J., Winkler, S. M. and Smolke, C. D. (2015) 'Synthetic feedback control using an RNAi-based gene-regulatory device', *Journal of Biological Engineering*, 9(1), p. 5. doi: 10.1186/s13036-015-0002-3.
- Bort, J. A. H., Stern, B. and Borth, N. (2010) 'CHO-K1 host cells adapted to growth in glutamine-free medium by FACS-assisted evolution', *Biotechnology Journal*, 5(10), pp. 1090–1097. doi: 10.1002/biot.201000095.
- Bort, J. a H. *et al.* (2012) 'Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures', *Biotechnology Journal*, 7(4), pp. 500–515. doi: 10.1002/biot.201100143.
- Brown, B. D. *et al.* (2006a) 'Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer', *Nature Medicine*, 12(5), pp. 585–591. doi: 10.1038/nm1398.
- Brown, B. D. *et al.* (2006b) 'Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer', *Nature Medicine*, 12(5), pp. 585–591. doi: 10.1038/nm1398.
- Brown, B. D. *et al.* (2007) 'Endogenous microRNA can be broadly exploited to regulate transgene expression according to tissue, lineage and differentiation state', *Nature Biotechnology*, 25(12), pp. 1457–1467. doi: 10.1038/nbt1372.
- Bustin, S. A. *et al.* (2009) 'The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments', *Clinical Chemistry*, 55(4), pp. 611–622. doi: 10.1373/clinchem.2008.112797.
- Butler, M. and Meneses-Acosta, A. (2012) 'Recent advances in technology supporting biopharmaceutical production from mammalian cells', *Applied Microbiology and Biotechnology*, 96(4), pp. 885–894. doi: 10.1007/s00253-012-4451-z.

- Cabodevilla, A. G. *et al.* (2013) 'Cell survival during complete nutrient deprivation depends on lipid droplet-fueled β -oxidation of fatty acids', *Journal of Biological Chemistry*, 288(39), pp. 27777–27788. doi: 10.1074/jbc.M113.466656.
- Calvo, S. E., Pagliarini, D. J. and Mootha, V. K. (2009) 'Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans', *Proceedings of the National Academy of Sciences of the United States of America*, 106(18), pp. 7507–7512. doi: 10.1073/pnas.0810916106.
- Chan, K. F. *et al.* (2016) 'Inactivation of GDP-fucose transporter gene (Slc35c1) in CHO cells by ZFNs, TALENs and CRISPR-Cas9 for production of fucose-free antibodies', *Biotechnology Journal*. John Wiley & Sons, Ltd, 11(3), pp. 399–414. doi: 10.1002/biot.201500331.
- Chang, T.-C. *et al.* (2008) 'Widespread microRNA repression by Myc contributes to tumorigenesis', *Nature Genetics*, 40(1), pp. 43–50. doi: 10.1038/ng.2007.30.
- Chen, C. Y. and Shyu, A. B. (1994) 'Selective degradation of early-response-gene mRNAs: functional analyses of sequence features of the AU-rich elements.', *Molecular and Cellular Biology*, 14(12), pp. 8471–8482. doi: 10.1128/mcb.14.12.8471.
- Chen, Z. *et al.* (2016) 'Theophylline controllable RNAi-based genetic switches regulate expression of lncRNA TINCR and malignant phenotypes in bladder cancer cells', *Scientific Reports*. Nature Publishing Group, 6(April), p. 30798. doi: 10.1038/srep30798.
- Chiang, G. G. and Sisk, W. P. (2005) 'Bcl-xL mediates increased production of humanized monoclonal antibodies in Chinese hamster ovary cells', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 91(7), pp. 779–792. doi: 10.1002/bit.20551.
- Chipman, L. B. and Pasquinelli, A. E. (2019) 'miRNA Targeting: Growing beyond the Seed', *Trends in Genetics*, 35(3), pp. 215–222. doi: 10.1016/j.tig.2018.12.005.
- Chu, L. and Robinson, D. K. (2001) 'Industrial choices for protein production by large-scale cell culture', *Current Opinion in Biotechnology*, 12(2), pp. 180–187. doi: 10.1016/S0958-1669(00)00197-X.
- Clarke, C. *et al.* (2012) 'Integrated miRNA, mRNA and protein expression analysis reveals the role of post-transcriptional regulation in controlling CHO cell growth rate', *BMC Genomics*. BMC Genomics, 13(1), p. 656. doi: 10.1186/1471-2164-13-656.
- Corcoran, D. L. *et al.* (2009) 'Features of Mammalian microRNA Promoters Emerge from Polymerase II Chromatin Immunoprecipitation Data', *PLoS ONE*. Edited by C. K. Patil. Public Library of Science, 4(4), p. e5279. doi: 10.1371/journal.pone.0005279.
- Cost, G. J. *et al.* (2010) 'BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 105(2), pp. 330–340. doi: 10.1002/bit.22541.
- Costello, A. *et al.* (2017) 'Conditional Knockdown of Endogenous MicroRNAs in CHO Cells Using TET-ON-SanDI Sponge Vectors', in Meleady, P. (ed.). New York, NY: Springer New York, pp. 87–100. doi: 10.1007/978-1-4939-6972-2_6.

- Costello, A. *et al.* (2018) 'Depletion of endogenous miRNA-378-3p increases peak cell density of CHO DP12 cells and is correlated with elevated levels of ubiquitin carboxyl-terminal hydrolase 14', *Journal of Biotechnology*. Elsevier, 288(October), pp. 30–40. doi: 10.1016/j.jbiotec.2018.10.008.
- Costello, A., Lao, N. T., Barron, N., *et al.* (2019) 'Improved yield of rhEPO in CHO cells with synthetic 5' UTR', *Biotechnology Letters*. Springer Netherlands, 41(2), pp. 231–239. doi: 10.1007/s10529-018-2632-2.
- Costello, A., Lao, N. T., Gallagher, C., *et al.* (2019) 'Leaky Expression of the TET-On System Hinders Control of Endogenous miRNA Abundance', *Biotechnology Journal*, 14(3), p. 1800219. doi: 10.1002/biot.201800219.
- Cullen, L. M. and Arndt, G. M. (2005) 'Genome-wide screening for gene function using RNAi in mammalian cells', *Immunology and Cell Biology*, 83(3), pp. 217–223. doi: 10.1111/j.1440-1711.2005.01332.x.
- Dahodwala, H. *et al.* (2019) 'Increased mAb production in amplified CHO cell lines is associated with increased interaction of CREB1 with transgene promoter', *Current Research in Biotechnology*. The Authors, 1, pp. 49–57. doi: 10.1016/j.crbiot.2019.09.001.
- Datta, P., Linhardt, R. J. and Sharfstein, S. T. (2013) 'An 'omics approach towards CHO cell engineering', *Biotechnology and Bioengineering*, 110(5), pp. 1255–1271. doi: 10.1002/bit.24841.
- Davis, B. N. and Hata, A. (2010) 'Mechanisms of control of microRNA biogenesis', *Journal of Biochemistry*, 148(4), pp. 381–392. doi: 10.1093/jb/mvq096.
- Diendorfer, A. B. *et al.* (2015) 'Annotation of additional evolutionary conserved microRNAs in CHO cells from updated genomic data', *Biotechnology and Bioengineering*, 112(7), pp. 1488–1493. doi: 10.1002/bit.25539.
- Doyle, F. *et al.* (2017) 'Engineering Structurally Interacting RNA (sxRNA)', *Scientific Reports*. Nature Publishing Group, 7(February), p. 45393. doi: 10.1038/srep45393.
- Doyle, F. and Tenenbaum, S. A. (2014) 'Trans-regulation of RNA-binding protein motifs by microRNA', *Frontiers in Genetics*, 5(APR), pp. 1–10. doi: 10.3389/fgene.2014.00079.
- Dreesen, I. A. J. and Fussenegger, M. (2011) 'Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of chinese hamster ovary cells', *Biotechnology and Bioengineering*, 108(4), pp. 853–866. doi: 10.1002/bit.22990.
- Druz, A. *et al.* (2013) 'Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells', *Metabolic Engineering*. Elsevier, 16(1), pp. 87–94. doi: 10.1016/j.ymben.2012.12.004.
- Duarte, T. M. *et al.* (2014) 'Metabolic responses of CHO cells to limitation of key amino acids', *Biotechnology and Bioengineering*, 111(10), pp. 2095–2106. doi: 10.1002/bit.25266.

- Duportet, X. *et al.* (2014) 'A platform for rapid prototyping of synthetic gene networks in mammalian cells.', *Nucleic acids research*, 42(21), pp. 13440–13451. doi: 10.1093/nar/gku1082.
- Ebert, M. S., Neilson, J. R. and Sharp, P. A. (2007) 'MicroRNA sponges: Competitive inhibitors of small RNAs in mammalian cells', *Nature Methods*, 4(9), pp. 721–726. doi: 10.1038/nmeth1079.
- Edros, R., McDonnell, S. and Al-Rubeai, M. (2014) 'The relationship between mTOR signalling pathway and recombinant antibody productivity in CHO cell lines', *BMC Biotechnology*, 14, pp. 14–16. doi: 10.1186/1472-6750-14-15.
- Eichhorn, S. W. *et al.* (2014) 'mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAs by the Time Substantial Repression Ensues', *Molecular Cell*, 56(1), pp. 104–115. doi: 10.1016/j.molcel.2014.08.028.
- Emmerling, V. V. *et al.* (2016) 'Temperature-sensitive miR-483 is a conserved regulator of recombinant protein and viral vector production in mammalian cells', *Biotechnology and Bioengineering*, 113(4), pp. 830–841. doi: 10.1002/bit.25853.
- Figuerola, B. *et al.* (2007) 'Enhanced cell culture performance using inducible anti-apoptotic genes E1B-19K and Aven in the production of a monoclonal antibody with Chinese hamster ovary cells', *Biotechnology and Bioengineering*, 97(4), pp. 877–892. doi: 10.1002/bit.21222.
- Fischer, S., Mathias, S., *et al.* (2015) 'Enhanced protein production by microRNA-30 family in CHO cells is mediated by the modulation of the ubiquitin pathway', *Journal of Biotechnology*. Elsevier B.V. doi: 10.1016/j.jbiotec.2015.08.002.
- Fischer, S., Paul, A. J., *et al.* (2015) 'miR-2861 as novel HDAC5 inhibitor in CHO cells enhances productivity while maintaining product quality', *Biotechnology and Bioengineering*, 9999(xxx), p. n/a-n/a. doi: 10.1002/bit.25626.
- Fischer, S., Handrick, R. and Otte, K. (2015) 'The art of CHO cell engineering: A comprehensive retrospect and future perspectives', *Biotechnology Advances*. Elsevier B.V., 33(8), pp. 1878–1896. doi: 10.1016/j.biotechadv.2015.10.015.
- Flintoff, W. F., Davidson, S. V. and Siminovitch, L. (1976) 'Isolation and partial characterization of three methotrexate-resistant phenotypes from Chinese hamster ovary cells', *Somatic Cell Genetics*, 2(3), pp. 245–261. doi: 10.1007/BF01538963.
- Flintoff, W. F., Spindler, S. M. and Siminovitch, L. (1976) 'Genetic characterization of methotrexate-resistant chinese hamster ovary cells', *In Vitro*, 12(11), pp. 749–757. doi: 10.1007/BF02835450.
- Fomina-Yadlin, D. *et al.* (2014) 'Cellular responses to individual amino-acid depletion in antibody-expressing and parental CHO cell lines', *Biotechnology and Bioengineering*, 111(5), pp. 965–979. doi: 10.1002/bit.25155.
- Fox, S. R. *et al.* (2005) 'A detailed understanding of the enhanced hypothermic productivity of interferon- γ by Chinese-hamster ovary cells', *Biotechnology and Applied*

Biochemistry, 41(3), p. 255. doi: 10.1042/ba20040066.

Gaidukov, L. *et al.* (2018) 'A multi-landing pad DNA integration platform for mammalian cell engineering', *Nucleic Acids Research*. Oxford University Press, 46(8), pp. 4072–4086. doi: 10.1093/nar/gky216.

Gam, J. J., Babb, J. and Weiss, R. (2018) 'A mixed antagonistic/synergistic miRNA repression model enables accurate predictions of multi-input miRNA sensor activity', *Nature Communications*. Springer US, 9(1), pp. 1–12. doi: 10.1038/s41467-018-04575-0.

Gameiro, P. A. and Struhl, K. (2018) 'Nutrient Deprivation Elicits a Transcriptional and Translational Inflammatory Response Coupled to Decreased Protein Synthesis', *Cell Reports*, 24(6), pp. 1415–1424. doi: 10.1016/j.celrep.2018.07.021.

Gammell, P. *et al.* (2007) 'Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells', *Journal of Biotechnology*, 130(3), pp. 213–218. doi: 10.1016/j.jbiotec.2007.04.020.

Genzel, Y. *et al.* (2005) 'Substitution of glutamine by pyruvate to reduce ammonia formation and growth inhibition of mammalian cells', *Biotechnology Progress*, 21(1), pp. 58–69. doi: 10.1021/bp049827d.

Georgakilas, G. *et al.* (2014) 'microTSS: accurate microRNA transcription start site identification reveals a significant number of divergent pri-miRNAs', *Nature Communications*, 5(1), p. 5700. doi: 10.1038/ncomms6700.

Görlach, M., Burd, C. G. and Dreyfuss, G. (1994) 'The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins', *Journal of Biological Chemistry*, 269(37), pp. 23074–23078.

Gotic, I. *et al.* (2016) 'Temperature regulates splicing efficiency of the cold-inducible RNA-binding protein gene Cirbp', *Genes & Development*, 30(17), pp. 2005–2017. doi: 10.1101/gad.287094.116.

Gottesman, M. M. (1987) '[1] Chinese hamster ovary cells', in *Methods in Enzymology*, pp. 3–8. doi: 10.1016/S0076-6879(87)51004-7.

Gray, N. K. and Hentze, M. W. (1994) 'Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs.', *The EMBO Journal*, 13(16), pp. 3882–3891. doi: 10.1002/j.1460-2075.1994.tb06699.x.

Greber, D., El-Baba, M. D. and Fussenegger, M. (2008) 'Intronic encoded siRNAs improve dynamic range of mammalian gene regulation systems and toggle switch', *Nucleic Acids Research*, 36(16), pp. e101–e101. doi: 10.1093/nar/gkn443.

Green, A. A. *et al.* (2017) 'Complex cellular logic computation using ribocomputing devices', *Nature*. Nature Publishing Group, 548(7665), pp. 117–121. doi: 10.1038/nature23271.

Green, A. a *et al.* (2014) 'Toehold Switches : De-Novo-Designed Regulators of Gene Expression', *Cell*. Elsevier Inc., 159(4), pp. 925–939. doi: 10.1016/j.cell.2014.10.002.

- Ha, M. and Kim, V. N. (2014) 'Regulation of microRNA biogenesis', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 15(8), pp. 509–524. doi: 10.1038/nrm3838.
- Ha, T. K. and Lee, G. M. (2014) 'Effect of glutamine substitution by TCA cycle intermediates on the production and sialylation of Fc-fusion protein in Chinese hamster ovary cell culture', *Journal of Biotechnology*. Elsevier B.V., 180, pp. 23–29. doi: 10.1016/j.jbiotec.2014.04.002.
- Hackl, M. *et al.* (2011) 'Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering', *Journal of Biotechnology*. Elsevier B.V., 153(1–2), pp. 62–75. doi: 10.1016/j.jbiotec.2011.02.011.
- Han, Y. K. *et al.* (2011) 'Autophagy and apoptosis of recombinant Chinese hamster ovary cells during fed-batch culture: Effect of nutrient supplementation', *Biotechnology and Bioengineering*, 108(9), pp. 2182–2192. doi: 10.1002/bit.23165.
- Harcum, S. W. and Lee, K. H. (2016) 'CHO Cells Can Make More Protein', *Cell Systems*. Elsevier Inc., 3(5), pp. 412–413. doi: 10.1016/j.cels.2016.11.007.
- Harreither, E. *et al.* (2015) 'Microarray profiling of preselected CHO host cell subclones identifies gene expression patterns associated with increased production capacity', *Biotechnology Journal*, 10(10), pp. 1625–1638. doi: 10.1002/biot.201400857.
- Herzig, S. and Shaw, R. J. (2018) 'AMPK: guardian of metabolism and mitochondrial homeostasis', *Nature Reviews Molecular Cell Biology*, 19(2), pp. 121–135. doi: 10.1038/nrm.2017.95.
- Hirosawa, M. *et al.* (2017) 'Cell-type-specific genome editing with a microRNA-responsive CRISPR–Cas9 switch', *Nucleic Acids Research*, 45(13), pp. e118–e118. doi: 10.1093/nar/gkx309.
- Hong, J. K., Cho, S. M. and Yoon, S. K. (2010) 'Substitution of glutamine by glutamate enhances production and galactosylation of recombinant IgG in Chinese hamster ovary cells', *Applied Microbiology and Biotechnology*, 88(4), pp. 869–876. doi: 10.1007/s00253-010-2790-1.
- Hosios, A. M. *et al.* (2016) 'Amino Acids Rather than Glucose Account for the Majority of Cell Mass in Proliferating Mammalian Cells', *Developmental Cell*. Elsevier Inc., 36(5), pp. 540–549. doi: 10.1016/j.devcel.2016.02.012.
- Huang, E. P., Marquis, C. P. and Gray, P. P. (2004) 'Process development for a recombinant Chinese Hamster Ovary (CHO) cell line utilizing a metal induced and amplified metallothionein expression system', *Biotechnology and Bioengineering*, 88(4), pp. 437–450. doi: 10.1002/bit.20194.
- Hwang, S. O. and Lee, G. M. (2008) 'Nutrient deprivation induces autophagy as well as apoptosis in Chinese hamster ovary cell culture', *Biotechnology and Bioengineering*, 99(3), pp. 678–685. doi: 10.1002/bit.21589.

Inwood, S., Betenbaugh, M. J. and Shiloach, J. (2018) 'Methods for using small non-coding rnas to improve recombinant protein expression in mammalian cells', *Genes*, 9(1). doi: 10.3390/genes9010025.

Jadhav, V. *et al.* (2014) 'Stable overexpression of miR-17 enhances recombinant protein production of CHO cells', *Journal of Biotechnology*. Elsevier B.V., 175(1), pp. 38–44. doi: 10.1016/j.jbiotec.2014.01.032.

Jardon, M. A. *et al.* (2012) 'Inhibition of glutamine-dependent autophagy increases t-PA production in CHO Cell fed-batch processes', *Biotechnology and Bioengineering*, 109(5), pp. 1228–1238. doi: 10.1002/bit.24393.

Jo, M. H. *et al.* (2015) 'Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs', *Molecular Cell*, 59(1), pp. 117–124. doi: 10.1016/j.molcel.2015.04.027.

Jonas, S. and Izaurralde, E. (2015) 'Towards a molecular understanding of microRNA-mediated gene silencing', *Nature Reviews Genetics*, 16(7), pp. 421–433. doi: 10.1038/nrg3965.

Jossé, L., Zhang, L. and Smales, C. M. (2018) 'Application of microRNA Targeted 3'UTRs to Repress DHFR Selection Marker Expression for Development of Recombinant Antibody Expressing CHO Cell Pools', *Biotechnology Journal*, 13(10), pp. 1–8. doi: 10.1002/biot.201800129.

Kaas, C. S. *et al.* (2014) 'Toward genome-scale models of the Chinese hamster ovary cells: incentives, status and perspectives', *Pharmaceutical Bioprocessing*, 2(5), pp. 437–448. doi: 10.4155/pbp.14.54.

Kahvejian, A. *et al.* (2005) 'Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms', *Genes and Development*, 19(1), pp. 104–113. doi: 10.1101/gad.1262905.

Kao, F. T. and Puck, T. T. (1967) 'Genetics of somatic mammalian cells. IV. Properties of Chinese hamster cell mutants with respect to the requirement for proline.', *Genetics*, 55(3), pp. 513–524.

Kao, F. T. and Puck, T. T. (1968) 'Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells.', *Proceedings of the National Academy of Sciences of the United States of America*, 60(4), pp. 1275–81. doi: 10.1073/pnas.60.4.1275.

Karali, M. *et al.* (2011) 'MicroRNA-Restricted Transgene Expression in the Retina', *PLoS ONE*. Edited by A. Urtti, 6(7), p. e22166. doi: 10.1371/journal.pone.0022166.

Kawamata, T., Seitz, H. and Tomari, Y. (2009) 'Structural determinants of miRNAs for RISC loading and slicer-independent unwinding', *Nature Structural & Molecular Biology*, 16(9), pp. 953–960. doi: 10.1038/nsmb.1630.

Kelly, P. S. *et al.* (2014) 'Conserved microRNA function as a basis for Chinese hamster ovary cell engineering', *Biotechnology Letters*, 37(4), pp. 787–798. doi: 10.1007/s10529-014-1751-7.

- Kelly, P. S. *et al.* (2015) 'Re-programming CHO cell metabolism using miR-23 tips the balance towards a highly productive phenotype', *Biotechnology Journal*, 10(7), pp. 1029–1040. doi: 10.1002/biot.201500101.
- Kennedy, A. B. *et al.* (2014) 'Protein-responsive ribozyme switches in eukaryotic cells', *Nucleic Acids Research*, 42(19), pp. 12306–12321. doi: 10.1093/nar/gku875.
- Kim, J. *et al.* (2019) 'De novo-designed translation-repressing riboregulators for multi-input cellular logic', *Nature Chemical Biology*. Springer US, 15(12), pp. 1173–1182. doi: 10.1038/s41589-019-0388-1.
- Kim, J. and Guan, K. L. (2019) 'mTOR as a central hub of nutrient signalling and cell growth', *Nature Cell Biology*. Springer US, 21(1), pp. 63–71. doi: 10.1038/s41556-018-0205-1.
- Kim, J. Y., Kim, Y.-G. and Lee, G. M. (2012) 'CHO cells in biotechnology for production of recombinant proteins: current state and further potential', *Applied Microbiology and Biotechnology*, 93(3), pp. 917–930. doi: 10.1007/s00253-011-3758-5.
- Kim, K. *et al.* (2018) 'SRSF3 recruits DROSHA to the basal junction of primary microRNAs', *RNA*, 24(7), pp. 892–898. doi: 10.1261/rna.065862.118.
- Kim, V. N., Han, J. and Siomi, M. C. (2009) 'Biogenesis of small RNAs in animals', *Nature Reviews Molecular Cell Biology*, 10(2), pp. 126–139. doi: 10.1038/nrm2632.
- Klanert, G. *et al.* (2016) 'A signature of 12 microRNAs is robustly associated with growth rate in a variety of CHO cell lines', *Journal of Biotechnology*. Elsevier B.V., 235, pp. 150–161. doi: 10.1016/j.jbiotec.2016.03.022.
- Kluiver, J. *et al.* (2012) 'Rapid generation of microRNA sponges for microRNA inhibition', *PLoS ONE*, 7(1), pp. 14–21. doi: 10.1371/journal.pone.0029275.
- Knuckles, P. *et al.* (2017) 'RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding', *Nature Structural & Molecular Biology*, 24(7), pp. 561–569. doi: 10.1038/nsmb.3419.
- Koressaar, T. and Remm, M. (2007) 'Enhancements and modifications of primer design program Primer3', *Bioinformatics*, 23(10), pp. 1289–1291. doi: 10.1093/bioinformatics/btm091.
- Kretz, M. *et al.* (2013) 'Control of somatic tissue differentiation by the long non-coding RNA TINCR', *Nature*. Nature Publishing Group, 493(7431), pp. 231–235. doi: 10.1038/nature11661.
- Lalonde, M.-E. and Durocher, Y. (2017) 'Therapeutic glycoprotein production in mammalian cells', *Journal of Biotechnology*. Elsevier, 251(April), pp. 128–140. doi: 10.1016/j.jbiotec.2017.04.028.
- Lamping, E., Niimi, M. and Cannon, R. D. (2013) 'Small, synthetic, GC-rich mRNA stem-loop modules 5' proximal to the AUG start-codon predictably tune gene expression in yeast', *Microbial Cell Factories*. Microbial Cell Factories, 12(1), p. 1. doi: 10.1186/1475-2859-12-74.

Lapique, N. and Benenson, Y. (2014) 'Digital switching in a biosensor circuit via programmable timing of gene availability', *Nature Chemical Biology*. Nature Publishing Group, 10(12), pp. 1020–1027. doi: 10.1038/nchembio.1680.

Lee, E. U., Roth, J. and Paulson, J. C. (1989) 'Alteration of terminal glycosylation sequences on N-linked oligosaccharides of Chinese hamster ovary cells by expression of beta-galactoside alpha 2,6-sialyltransferase.', *The Journal of biological chemistry*, 264(23), pp. 13848–55. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2668274>.

Lee, J. S. and Lee, G. M. (2012) 'Rapamycin treatment inhibits CHO cell death in a serum-free suspension culture by autophagy induction', *Biotechnology and Bioengineering*, 109(12), pp. 3093–3102. doi: 10.1002/bit.24567.

Lee, Y. *et al.* (2004) 'MicroRNA genes are transcribed by RNA polymerase II', *The EMBO Journal*, 23(20), pp. 4051–4060. doi: 10.1038/sj.emboj.7600385.

Leisner, M. *et al.* (2010) 'Rationally designed logic integration of regulatory signals in mammalian cells', *Nature Nanotechnology*, 5(9), pp. 666–670. doi: 10.1038/nnano.2010.135.

Leppek, K., Das, R. and Barna, M. (2018) 'Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them', *Nature Reviews Molecular Cell Biology*, 19(3), pp. 158–174. doi: 10.1038/nrm.2017.103.

Ley, D. *et al.* (2015) 'Multi-omic profiling -of EPO-producing Chinese hamster ovary cell panel reveals metabolic adaptation to heterologous protein production', *Biotechnology and Bioengineering*, 112(11), pp. 2373–2387. doi: 10.1002/bit.25652.

Ley, D. *et al.* (2019) 'Reprogramming AA catabolism in CHO cells with CRISPR/Cas9 genome editing improves cell growth and reduces byproduct secretion', *Metabolic Engineering*, 56, pp. 120–129. doi: 10.1016/j.ymben.2019.09.005.

Li, C.-H. *et al.* (2017) 'EZH2 coupled with HOTAIR to silence MicroRNA-34a by the induction of heterochromatin formation in human pancreatic ductal adenocarcinoma', *International Journal of Cancer*. John Wiley & Sons, Ltd, 140(1), pp. 120–129. doi: 10.1002/ijc.30414.

Li, Y. *et al.* (2015) 'Modular construction of mammalian gene circuits using TALE transcriptional repressors', *Nature Chemical Biology*. Nature Publishing Group, 11(3), pp. 207–213. doi: 10.1038/nchembio.1736.

Liang, X. *et al.* (2016) 'Translation efficiency of mRNAs is increased by antisense oligonucleotides targeting upstream open reading frames', *Nature Biotechnology*. Nature Publishing Group, 34(8), pp. 875–880. doi: 10.1038/nbt.3589.

Lin, N. *et al.* (2011) 'Profiling highly conserved microrna expression in recombinant IgG-producing and parental Chinese hamster ovary cells', *Biotechnology Progress*, 27(4), pp. 1163–1171. doi: 10.1002/btpr.556.

Lipscomb, M. L., Mowry, M. C. and Kompala, D. S. (2004) 'Production of a Secreted Glycoprotein from an Inducible Promoter System in a Perfusion Bioreactor',

Biotechnology Progress, 20(5), pp. 1402–1407. doi: 10.1021/bp049973j.

Lisec, J. *et al.* (2019) 'Cancer cell lipid class homeostasis is altered under nutrient-deprivation but stable under hypoxia', *BMC Cancer*. BMC Cancer, 19(1), pp. 1–11. doi: 10.1186/s12885-019-5733-y.

Liu, J. *et al.* (2005) 'MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies', *Nature Cell Biology*, 7(7), pp. 719–723. doi: 10.1038/ncb1274.

Livak, K. J. and Schmittgen, T. D. (2001) 'Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method', *Methods*, 25(4), pp. 402–408. doi: 10.1006/meth.2001.1262.

Lytle, J. R., Yario, T. A. and Steitz, J. A. (2007) 'Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR', *Proceedings of the National Academy of Sciences*, 104(23), pp. 9667–9672. doi: 10.1073/pnas.0703820104.

Maccani, A. *et al.* (2014) 'Identification of microRNAs specific for high producer CHO cell lines using steady-state cultivation', *Applied Microbiology and Biotechnology*, 98(17), pp. 7535–7548. doi: 10.1007/s00253-014-5911-4.

MacRae, I. J. (2006) 'Structural Basis for Double-Stranded RNA Processing by Dicer', *Science*, 311(5758), pp. 195–198. doi: 10.1126/science.1121638.

Majors, B. S. *et al.* (2012) 'Directed evolution of mammalian anti-apoptosis proteins by somatic hypermutation', *Protein Engineering Design and Selection*, 25(1), pp. 27–38. doi: 10.1093/protein/gzr052.

Malphettes, L. and Fussenegger, M. (2006) 'Improved transgene expression fine-tuning in mammalian cells using a novel transcription?translation network', *Journal of Biotechnology*, 124(4), pp. 732–746. doi: 10.1016/j.jbiotec.2006.01.003.

Mansfield, J. H. *et al.* (2004) 'MicroRNA-responsive "sensor" transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression', *Nature Genetics*, 36(10), pp. 1079–1083. doi: 10.1038/ng1421.

Maralingannavar, V. *et al.* (2019) 'Superfluous glutamine synthetase activity in Chinese Hamster Ovary cells selected under glutamine limitation is growth limiting in glutamine-replete conditions and can be inhibited by serine', *Biotechnology Progress*, 35(5), pp. 1–11. doi: 10.1002/btpr.2856.

Masterton, R. J. *et al.* (2010) 'Post-translational events of a model reporter protein proceed with higher fidelity and accuracy upon mild hypothermic culturing of Chinese hamster ovary cells', *Biotechnology and Bioengineering*, 105(1), pp. 215–220. doi: 10.1002/bit.22533.

Masterton, R. J. and Smales, C. M. (2014) 'The impact of process temperature on mammalian cell lines and the implications for the production of recombinant proteins in CHO cells', *Pharmaceutical Bioprocessing*, 2(1), pp. 49–61. doi: 10.4155/pbp.14.3.

Meisner, N. C. and Filipowicz, W. (2010) 'Properties of the regulatory RNA-binding protein HuR and its role in controlling miRNA repression', *Advances in Experimental*

Medicine and Biology, 700, pp. 106–123. doi: 10.1007/978-1-4419-7823-3_10.

Melo, E. O., De Melo Neto, O. P. and Martins De Sá, C. (2003) 'Adenosine-rich elements present in the 5'-untranslated region of PABP mRNA can selectively reduce the abundance and translation of CAT mRNAs in vivo', *FEBS Letters*, 546(2–3), pp. 329–334. doi: 10.1016/S0014-5793(03)00620-3.

Meyer, K. D. *et al.* (2015) '5' UTR m6A Promotes Cap-Independent Translation', *Cell*, 163(4), pp. 999–1010. doi: 10.1016/j.cell.2015.10.012.

Michlewski, G. and Cáceres, J. F. (2019) 'Post-transcriptional control of miRNA biogenesis', *RNA*, 25(1), pp. 1–16. doi: 10.1261/rna.068692.118.

Miki, K. *et al.* (2015) 'Efficient Detection and Purification of Cell Populations Using Synthetic MicroRNA Switches', *Cell Stem Cell*. Elsevier Inc., 16(6), pp. 699–711. doi: 10.1016/j.stem.2015.04.005.

Millard, S. S. *et al.* (2000) 'A U-Rich Element in the 5' Untranslated Region Is Necessary for the Translation of p27 mRNA', *Molecular and Cellular Biology*, 20(16), pp. 5947–5959. doi: 10.1128/mcb.20.16.5947-5959.2000.

Mohammadi, P., Beerenwinkel, N. and Benenson, Y. (2017) 'Automated Design of Synthetic Cell Classifier Circuits Using a Two-Step Optimization Strategy', *Cell Systems*. Elsevier Inc., 4(2), pp. 207–218.e14. doi: 10.1016/j.cels.2017.01.003.

Morales, S., Monzo, M. and Navarro, A. (2017) 'Epigenetic regulation mechanisms of microRNA expression', *Biomolecular Concepts*. Berlin, Boston: De Gruyter, 8(5–6), pp. 203–212. doi: 10.1515/bmc-2017-0024.

Motheramgari, K. *et al.* (2019) 'Expanding the Chinese hamster ovary cell long non-coding RNA transcriptome using RNASeq', pp. 1–12. doi: 10.1101/863241.

Muckenthaler, M., Gray, N. K. and Hentze, M. W. (1998) 'IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F', *Molecular Cell*, 2(3), pp. 383–388. doi: 10.1016/S1097-2765(00)80282-8.

Nguyen, L. N. *et al.* (2019) 'Novel Promoters Derived from Chinese Hamster Ovary Cells via In Silico and In Vitro Analysis', *Biotechnology Journal*, 14(11), pp. 0–2. doi: 10.1002/biot.201900125.

Nguyen, T. A. *et al.* (2015) 'Functional Anatomy of the Human Microprocessor', *Cell*, 161(6), pp. 1374–1387. doi: 10.1016/j.cell.2015.05.010.

Ozsolak, F. *et al.* (2008) 'Chromatin structure analyses identify miRNA promoters', *Genes & Development*, 22(22), pp. 3172–3183. doi: 10.1101/gad.1706508.

Papageorgiou, A. *et al.* (2015) 'A Genome-Wide siRNA Screen in Mammalian Cells for Regulators of S6 Phosphorylation', *PLOS ONE*. Edited by D. Cota, 10(3), p. e0116096. doi: 10.1371/journal.pone.0116096.

Parker, R. and Sheth, U. (2007) 'P Bodies and the Control of mRNA Translation and Degradation', *Molecular Cell*, 25(5), pp. 635–646. doi: 10.1016/j.molcel.2007.02.011.

- Pavlova, N. N. *et al.* (2018) 'As Extracellular Glutamine Levels Decline, Asparagine Becomes an Essential Amino Acid', *Cell Metabolism*. Elsevier Inc., 27(2), pp. 428–438.e5. doi: 10.1016/j.cmet.2017.12.006.
- Peng, S. S. Y. *et al.* (1998) 'RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein', *EMBO Journal*, 17(12), pp. 3461–3470. doi: 10.1093/emboj/17.12.3461.
- Potla, R. *et al.* (2015) 'Shifts in temperature within the physiologic range modify strand-specific expression of select human microRNAs', *Rna*, 21(7), pp. 1261–1273. doi: 10.1261/rna.049122.114.
- Pritchard, C. C., Cheng, H. H. and Tewari, M. (2012) 'MicroRNA profiling: Approaches and considerations', *Nature Reviews Genetics*. Nature Publishing Group, 13(5), pp. 358–369. doi: 10.1038/nrg3198.
- Prochazka, L. *et al.* (2014) 'Highly modular bow-tie gene circuits with programmable dynamic behaviour.', *Nature communications*, 5, p. 4729. doi: 10.1038/ncomms5729.
- Puck, T. T. and Fisher, H. W. (1956) 'GENETICS OF SOMATIC MAMMALIAN CELLS', *The Journal of Experimental Medicine*, 104(3), pp. 427–434. doi: 10.1084/jem.104.3.427.
- Puck, T. T. and Kao, F.-T. (1982) 'SOMATIC CELL GENETICS AND ITS APPLICATION TO MEDICINE', *Annual Review of Genetics*. Annual Reviews, 16(1), pp. 225–271. doi: 10.1146/annurev.ge.16.120182.001301.
- Purnick, P. E. M. and Weiss, R. (2009) 'The second wave of synthetic biology: from modules to systems', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 10(6), pp. 410–422. doi: 10.1038/nrm2698.
- Rajendra, Y. *et al.* (2012) 'Reduced glutamine concentration improves protein production in growth-arrested CHO-DG44 and HEK-293E cells', *Biotechnology Letters*, 34(4), pp. 619–626. doi: 10.1007/s10529-011-0809-z.
- Rashid, H. O. *et al.* (2015) 'ER stress: Autophagy induction, inhibition and selection', *Autophagy*, 11(11), pp. 1956–1977. doi: 10.1080/15548627.2015.1091141.
- Ratnadiwakara, M., Mohenska, M. and Änkö, M.-L. (2018) 'Splicing factors as regulators of miRNA biogenesis – links to human disease', *Seminars in Cell & Developmental Biology*, 79, pp. 113–122. doi: 10.1016/j.semcdb.2017.10.008.
- de Rie, D. *et al.* (2017) 'An integrated expression atlas of miRNAs and their promoters in human and mouse', *Nature Biotechnology*, 35(9), pp. 872–878. doi: 10.1038/nbt.3947.
- Rinaudo, K. *et al.* (2007) 'A universal RNAi-based logic evaluator that operates in mammalian cells', *Nature Biotechnology*, 25(7), pp. 795–801. doi: 10.1038/nbt1307.
- Sanchez, N. *et al.* (2013) 'MiR-7 Triggers Cell Cycle Arrest at the G1/S Transition by Targeting Multiple Genes Including Skp2 and Psme3', *PLoS ONE*. Edited by B. Bardoni, 8(6), p. e65671. doi: 10.1371/journal.pone.0065671.
- Santos, A. K., Parreira, R. C. and Resende, R. R. (2016) 'Expression System Based on an

MTIIa Promoter to Produce hPSA in Mammalian Cell Cultures', *Frontiers in Microbiology*, 7(AUG), pp. 1–10. doi: 10.3389/fmicb.2016.01280.

Sayeg, M. K. *et al.* (2015) 'Rationally Designed MicroRNA-Based Genetic Classifiers Target Specific Neurons in the Brain', *ACS Synthetic Biology*, 4(7), pp. 788–795. doi: 10.1021/acssynbio.5b00040.

Schirle, N. T., Sheu-Gruttadauria, J. and MacRae, I. J. (2014) 'Structural basis for microRNA targeting', *Science*, 346(6209), pp. 608–613. doi: 10.1126/science.1258040.

Serganov, A. and Nudler, E. (2013) 'A Decade of Riboswitches', *Cell*, 152(1–2), pp. 17–24. doi: 10.1016/j.cell.2012.12.024.

Shridhar, S. *et al.* (2017) 'Transcriptomic changes in CHO cells after adaptation to suspension growth in protein-free medium analysed by a species-specific microarray', *Journal of Biotechnology*, 257, pp. 13–21. doi: 10.1016/j.jbiotec.2017.03.012.

Starega-Roslan, J. *et al.* (2011) 'Structural basis of microRNA length variety', *Nucleic Acids Research*, 39(1), pp. 257–268. doi: 10.1093/nar/gkq727.

Starega-Roslan, J., Galka-Marciniak, P. and Krzyzosiak, W. J. (2015) 'Nucleotide sequence of miRNA precursor contributes to cleavage site selection by Dicer', *Nucleic Acids Research*, 43(22), pp. 10939–10951. doi: 10.1093/nar/gkv968.

Steiner, F. A. *et al.* (2007) 'Structural features of small RNA precursors determine Argonaute loading in *Caenorhabditis elegans*', *Nature Structural & Molecular Biology*, 14(10), pp. 927–933. doi: 10.1038/nsmb1308.

Strotbek, M. *et al.* (2013) 'Stable microRNA expression enhances therapeutic antibody productivity of Chinese hamster ovary cells', *Metabolic Engineering*. Elsevier, 20, pp. 157–166. doi: 10.1016/j.ymben.2013.10.005.

Sumitomo, Y. *et al.* (2012) 'Identification of a novel enhancer that binds Sp1 and contributes to induction of cold-inducible RNA-binding protein (cirp) expression in mammalian cells', *BMC Biotechnology*. BMC Biotechnology, 12(1), p. 72. doi: 10.1186/1472-6750-12-72.

Tam, O. H. *et al.* (2008) 'Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes', *Nature*, 453(7194), pp. 534–538. doi: 10.1038/nature06904.

Tamošaitis, L. and Smales, C. M. (2018) 'Meta-Analysis of Publicly Available Chinese Hamster Ovary (CHO) Cell Transcriptomic Datasets for Identifying Engineering Targets to Enhance Recombinant Protein Yields', *Biotechnology Journal*, 13(10), p. 1800066. doi: 10.1002/biot.201800066.

Tan, J. G. L. *et al.* (2015) 'Heat shock protein 27 overexpression in CHO cells modulates apoptosis pathways and delays activation of caspases to improve recombinant monoclonal antibody titre in fed-batch bioreactors', *Biotechnology Journal*. John Wiley & Sons, Ltd, 10(5), pp. 790–800. doi: 10.1002/biot.201400764.

Taschwer, M. *et al.* (2012) 'Growth, productivity and protein glycosylation in a CHO

EpoFc producer cell line adapted to glutamine-free growth', *Journal of Biotechnology*. Elsevier B.V., 157(2), pp. 295–303. doi: 10.1016/j.jbiotec.2011.11.014.

Thompson, L. H. and Baker, R. M. (1973) 'Isolation of mutants of cultured mammalian cells.', *Methods in cell biology*, 6, pp. 209–81. doi: 10.1016/s0091-679x(08)60052-7.

Thomson, D. W. and Dinger, M. E. (2016) 'Endogenous microRNA sponges: evidence and controversy', *Nature Reviews Genetics*. Nature Publishing Group, 17(5), pp. 272–283. doi: 10.1038/nrg.2016.20.

Tzani, I. *et al.* (2020) 'Subphysiological temperature induces pervasive alternative splicing in Chinese hamster ovary cells', *Biotechnology and Bioengineering*, 117(8), pp. 2489–2503. doi: 10.1002/bit.27365.

Untergasser, A. *et al.* (2012) 'Primer3—new capabilities and interfaces', *Nucleic Acids Research*, 40(15), pp. e115–e115. doi: 10.1093/nar/gks596.

Urlaub, G. *et al.* (1983) 'Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells', *Cell*, 33(2), pp. 405–412. doi: 10.1016/0092-8674(83)90422-1.

Urlaub, G. and Chasin, L. A. (1980) 'Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity.', *Proceedings of the National Academy of Sciences*, 77(7), pp. 4216–4220. doi: 10.1073/pnas.77.7.4216.

Valdés-Bango Curell, R. and Barron, N. (2018) 'Exploring the Potential Application of Short Non-Coding RNA-Based Genetic Circuits in Chinese Hamster Ovary Cells', *Biotechnology Journal*, 13(10), pp. 1–8. doi: 10.1002/biot.201700220.

Vogl, T. and Glieder, A. (2013) 'Regulation of *Pichia pastoris* promoters and its consequences for protein production', *New Biotechnology*. Elsevier B.V., 30(4), pp. 385–404. doi: 10.1016/j.nbt.2012.11.010.

Wachsmuth, M. *et al.* (2013) 'De novo design of a synthetic riboswitch that regulates transcription termination', *Nucleic Acids Research*, 41(4), pp. 2541–2551. doi: 10.1093/nar/gks1330.

Wahrheit, J., Nicolae, A. and Heinzle, E. (2014) 'Dynamics of growth and metabolism controlled by glutamine availability in Chinese hamster ovary cells', *Applied Microbiology and Biotechnology*, 98(4), pp. 1771–1783. doi: 10.1007/s00253-013-5452-2.

Walsh, G. (2010) 'Biopharmaceutical benchmarks 2010', *Nature Biotechnology*, 28(9), pp. 917–924. doi: 10.1038/nbt0910-917.

Walsh, G. (2014) 'Biopharmaceutical benchmarks 2014', *Nature Biotechnology*, 32(10), pp. 992–1000. doi: 10.1038/nbt.3040.

Walsh, G. (2018) 'Biopharmaceutical benchmarks 2018', *Nature Biotechnology*. Nature Publishing Group, 36(12), pp. 1136–1145. doi: 10.1038/nbt.4305.

Walsh, G. and Jefferis, R. (2006) 'Post-translational modifications in the context of

therapeutic proteins.', *Nature biotechnology*, 24(10), pp. 1241–1252. doi: 10.1038/nbt1252.

Wang, T. *et al.* (2012) 'Overview of Regulatory Strategies and Molecular Elements in Metabolic Engineering of Bacteria', *Molecular Biotechnology*, 52(3), pp. 300–308. doi: 10.1007/s12033-012-9514-y.

Weber, W. *et al.* (2009) 'A biotin-triggered genetic switch in mammalian cells and mice', *Metabolic Engineering*, 11(2), pp. 117–124. doi: 10.1016/j.ymben.2008.12.001.

Weenink, T. *et al.* (2018) 'Design of RNA Hairpin Modules that Predictably Tune Translation in Yeast', *Synthetic Biology*, 3(di), pp. 1–9. doi: 10.1093/synbio/ysy019.

Wei, C. M., Gershowitz, A. and Moss, B. (1975) 'Methylated nucleotides block 5' terminus of HeLa cell messenger RNA', *Cell*, 4(4), pp. 379–386. doi: 10.1016/0092-8674(75)90158-0.

Weinhandl, K. *et al.* (2014) 'Carbon source dependent promoters in yeasts', *Microbial Cell Factories*, 13(1), p. 5. doi: 10.1186/1475-2859-13-5.

Wilkens, C. a. and Gerdtzen, Z. P. (2015) 'Comparative Metabolic Analysis of CHO Cell Clones Obtained through Cell Engineering, for IgG Productivity, Growth and Cell Longevity', *PLOS ONE*, 10(3), p. e0119053. doi: 10.1371/journal.pone.0119053.

Wroblewska, L. *et al.* (2015) 'Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery', *Nature Biotechnology*. Nature Publishing Group, 33(8), pp. 839–841. doi: 10.1038/nbt.3301.

Wurm, F. M. (2013) 'CHO Quasispecies—Implications for Manufacturing Processes', *Processes*, 1(3). doi: 10.3390/pr1030296.

Xie, Z. *et al.* (2011) 'Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells', *Science*, 333(6047), pp. 1307–1311. doi: 10.1126/science.1205527.

Xu, J. *et al.* (2019) 'Systematic development of temperature shift strategies for Chinese hamster ovary cells based on short duration cultures and kinetic modeling', *mAbs*. Taylor & Francis, 11(1), pp. 191–204. doi: 10.1080/19420862.2018.1525262.

Xu, T. *et al.* (2015) 'SP1-induced upregulation of the long noncoding RNA TINCR regulates cell proliferation and apoptosis by affecting KLF2 mRNA stability in gastric cancer', *Oncogene*, 34(45), pp. 5648–5661. doi: 10.1038/onc.2015.18.

Xu, X. *et al.* (2011) 'The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line.', *Nature biotechnology*. Nature Publishing Group, 29(8), pp. 735–741. doi: 10.1038/nbt.1932.

Yi, R. *et al.* (2003) 'Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs', *Genes and Development*, 17(24), pp. 3011–3016. doi: 10.1101/gad.1158803.

Yoda, M. *et al.* (2010) 'ATP-dependent human RISC assembly pathways', *Nature Structural & Molecular Biology*, 17(1), pp. 17–23. doi: 10.1038/nsmb.1733.

Young, S. K. and Wek, R. C. (2016) 'Upstream open reading frames differentially regulate genespecific translation in the integrated stress response', *Journal of Biological Chemistry*, 291(33), pp. 16927–16935. doi: 10.1074/jbc.R116.733899.

Zeng, Y. and Cullen, B. R. (2005) 'Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences', *Journal of Biological Chemistry*, 280(30), pp. 27595–27603. doi: 10.1074/jbc.M504714200.

Zhu, M. M., Mollet, M. and Hubert, R. S. (2007) 'Industrial Production of Therapeutic Proteins: Cell Lines, Cell Culture, and Purification', in Kent, J. A. (ed.) *Kent and Riegel's Handbook of Industrial Chemistry and Biotechnology*. Boston, MA: Springer US, pp. 1421–1448. doi: 10.1007/978-0-387-27843-8_32.

APPENDICES

APPENDIX A – SUPPLEMENTARY MATERIAL FOR CHAPTER 3

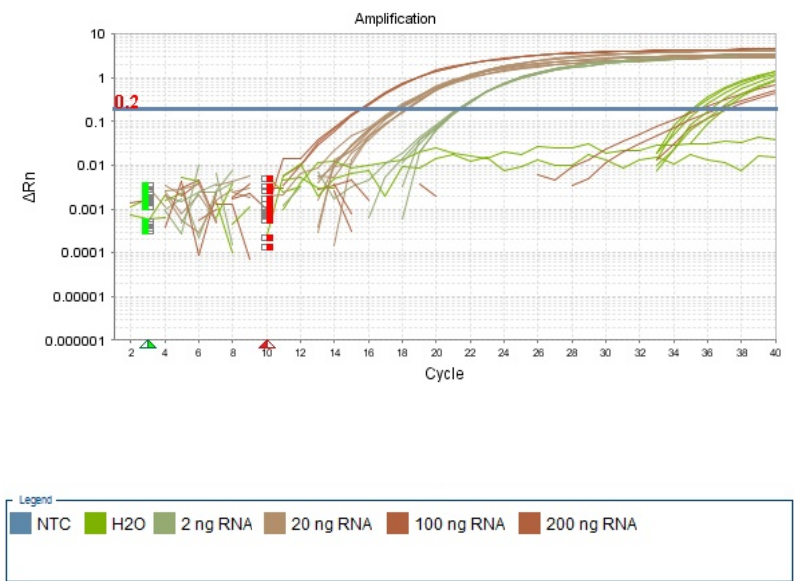
Supplementary Table 1

Cluster 1	Cluster 2	Cluster 3
cgr-miR-29a-5p	cgr-miR-16-5p	cgr-miR-199a-5p
cgr-miR-30d	cgr-miR-29b	cgr-miR-34c-5p
cgr-miR-26a-2	cgr-miR-93-5p	cgr-miR-23a-5p
cgr-let-7a	cgr-miR-25-5p	cgr-miR-146b-5p
cgr-miR-10b-5p	cgr-miR-26b-5p	cgr-miR-140-5p
cgr-let-7b	cgr-miR-20a	cgr-miR-34b-5p
cgr-miR-378-5p	cgr-miR-92a-5p	cgr-miR-27b-5p
cgr-miR-30a-5p	cgr-miR-30e-5p	cgr-miR-425-5p
cgr-let-7d-5p	cgr-miR-19b-5p	cgr-miR-674
cgr-miR-29c-5p	cgr-miR-7a	cgr-miR-320a
cgr-miR-29c-3p	cgr-miR-92a-3p	cgr-miR-199a
cgr-miR-181a-5p	cgr-miR-99a-5p	cgr-miR-196a-5p
cgr-miR-28-5p	cgr-miR-872-5p	cgr-miR-338
cgr-miR-365-5p	cgr-miR-30b-5p	cgr-miR-344
cgr-miR-423-5p	cgr-miR-350-5p	cgr-miR-298-5p
cgr-miR-32-5p	cgr-miR-17-5p	cgr-miR-598
cgr-miR-222-5p	cgr-miR-15b-5p	cgr-miR-34a
cgr-miR-3068	cgr-miR-106b-5p	cgr-miR-1260
cgr-miR-142-3p	cgr-miR-130a-5p	cgr-miR-6092
	cgr-miR-130a-3p	cgr-miR-184
	cgr-miR-503	cgr-miR-214-5p
	cgr-miR-190a	cgr-miR-455-5p
	cgr-miR-146a	cgr-miR-296
	cgr-miR-542-5p	cgr-miR-653
	cgr-miR-18a-5p	cgr-miR-328
	cgr-miR-301a-5p	cgr-miR-1843
	cgr-miR-15a-5p	cgr-miR-143
	cgr-miR-19a	
	cgr-miR-195	
	cgr-miR-107	
	cgr-miR-671-5p	
	cgr-miR-409-5p	
	cgr-miR-450b-5p	
	cgr-miR-497-5p	
	cgr-miR-497b	
	cgr-miR-374-5p	
	cgr-miR-137-5p	
	cgr-miR-505-5p	
	cgr-let-7c	
	cgr-miR-154-5p	
	cgr-miR-499-5p	
	cgr-miR-339	
	cgr-miR-744-5p	
	cgr-miR-190b	
	cgr-miR-33	
	cgr-miR-29b-5p	
	cgr-miR-9	

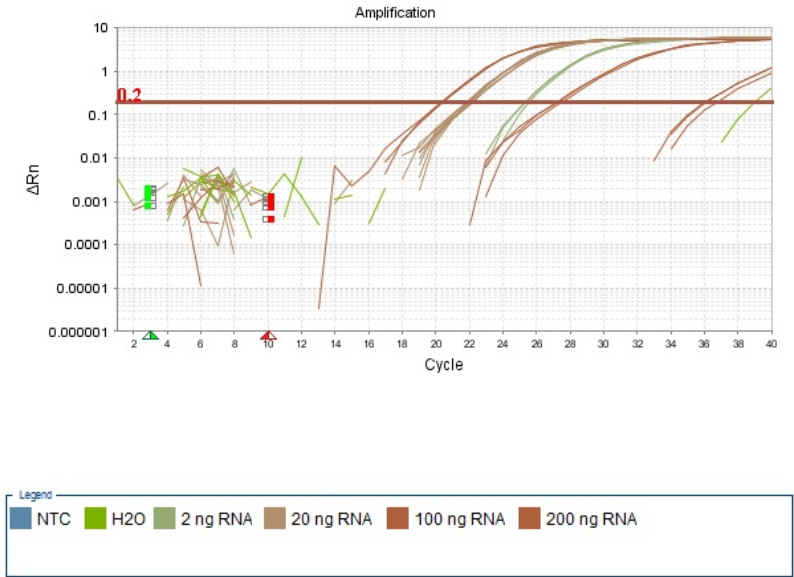
Supplementary Figure 2

Amplification curves corresponding to the attempt to validate ammonia correlated miRNAs – Chapter 3

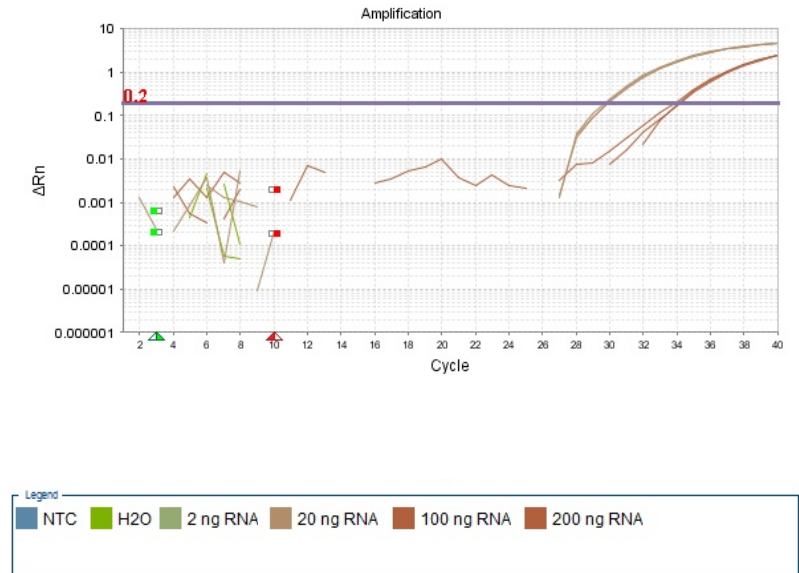
U6-snRNA



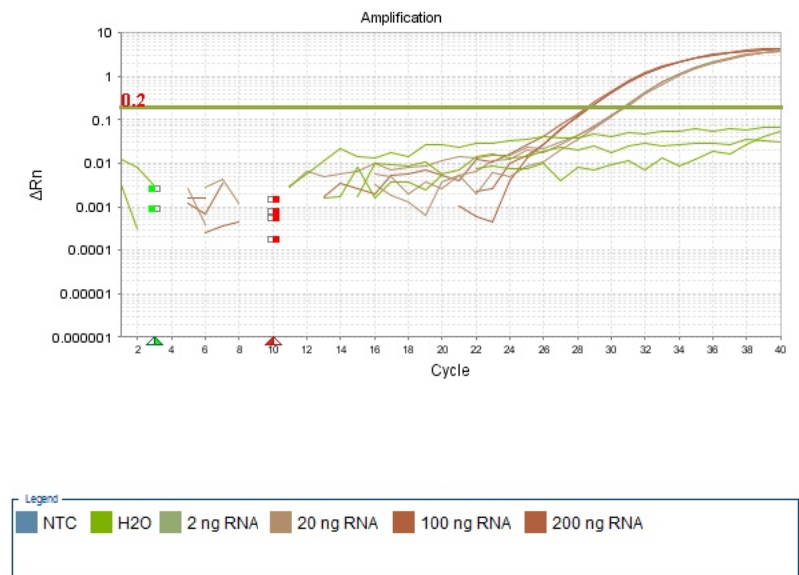
hsa-miR-378



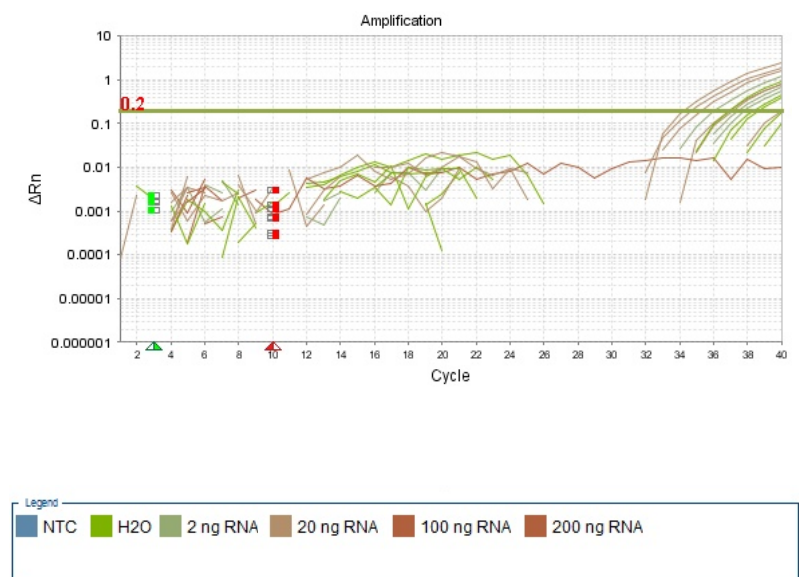
hsa-miR-338



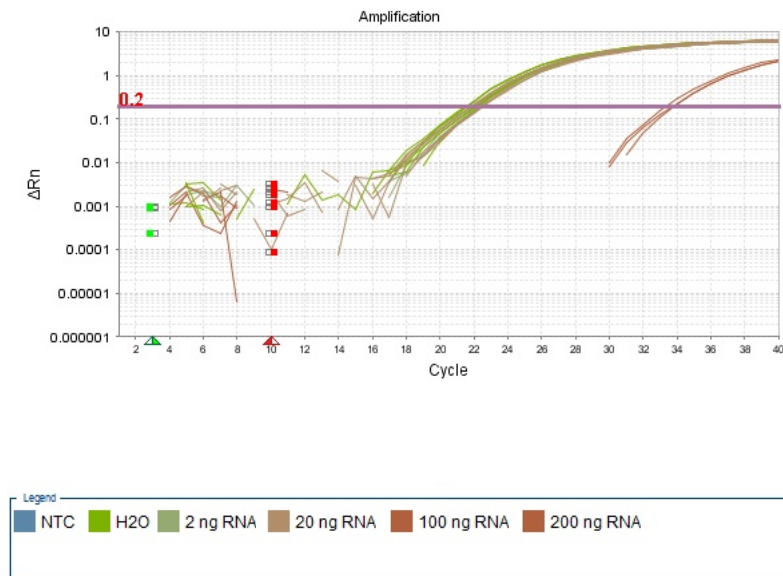
hsa-miR-409



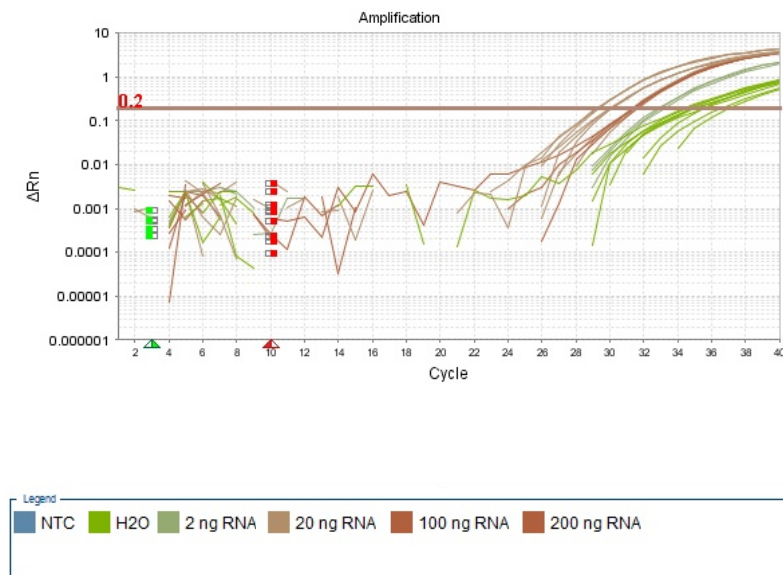
hsa-miR-187*



rno-miR-187*



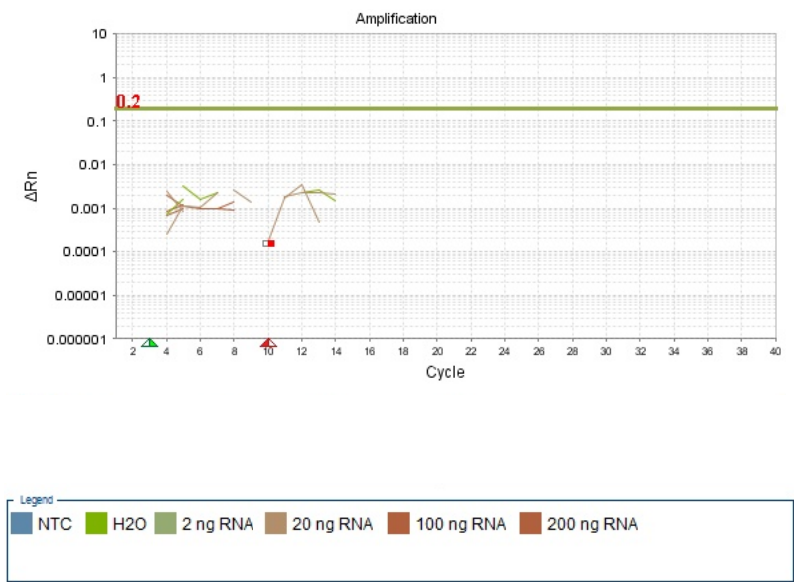
hsa-miR-638



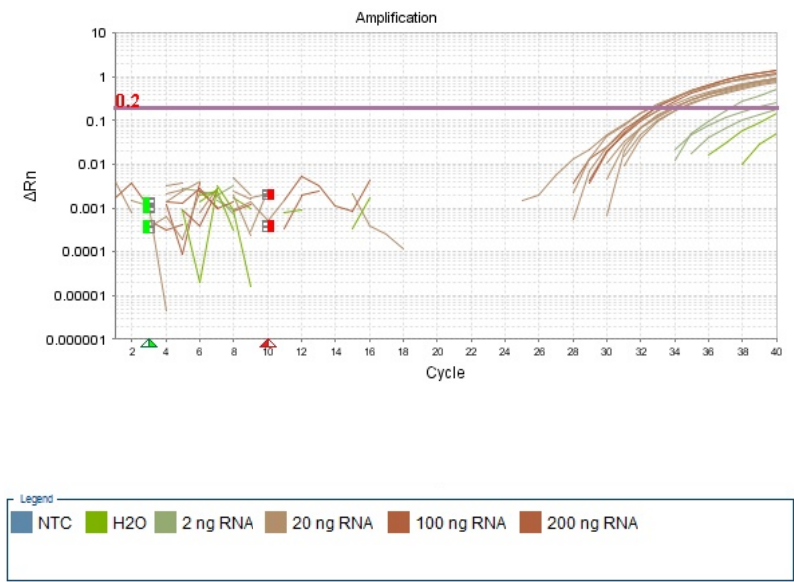
hsa-miR-3613



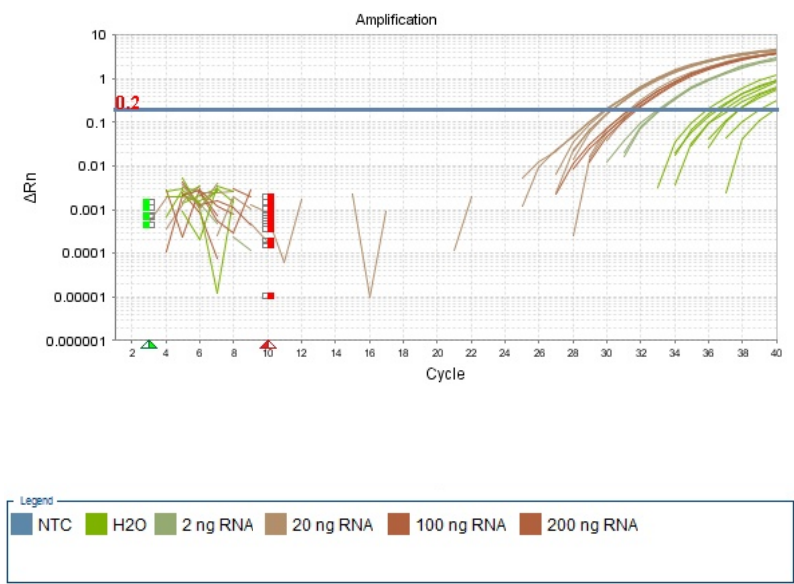
hsa-miR-3685



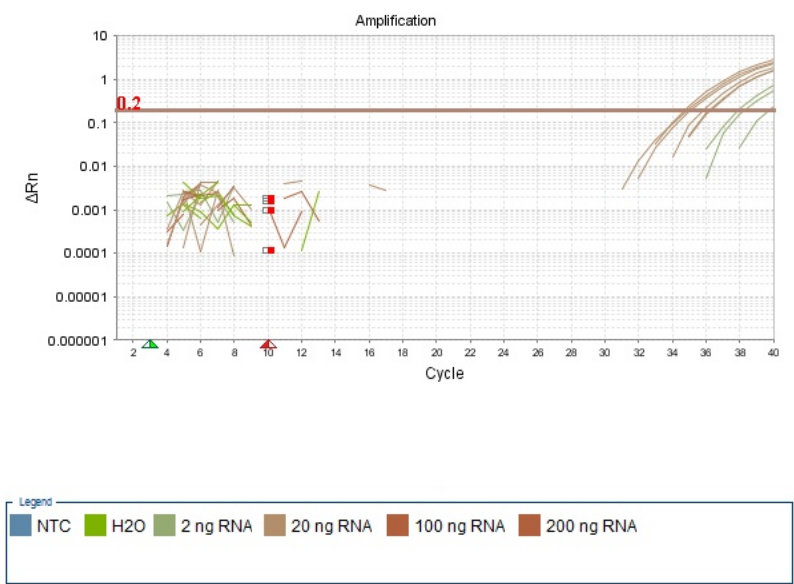
mmu-miR-1195



mmu-miR-1275



mmu-miR-3103



Supplementary Table 3

Raw data (Ct values, ΔC_t calculations and RQ values) for the validation of hsa-miR-1275 and mmu-miR-1195 in CHO-K1 glutamine dependent (+G) and glutamine-free (-G) by RT-qPCR using miRNA TaqMan assays - Chapter 3

Target Name	Sample Name	Time (h)	RQ	RQ Min	RQ Max	C _T Mean	ΔC_t Mean	ΔC_t SE	$\Delta \Delta C_t$
U6-snRNA	K1 +G D2	48				16.61			
U6-snRNA	K1 +G D4	96				16.41			
U6-snRNA	K1 +G D6	144				16.58			
U6-snRNA	K1 +G D8	192				20.01			
U6-snRNA	K1 -G D2	48				16.58			
U6-snRNA	K1 -G D4	96				16.57			
U6-snRNA	K1 -G D6	144				16.37			
U6-snRNA	K1 -G D8	192				17.35			
mmu-miR-1195	K1 +G D2	48	1.00	0.93	1.07	31.24	14.45	0.05	0.00
mmu-miR-1195	K1 +G D4	96	0.55	0.51	0.59	31.79	15.31	0.05	0.87
mmu-miR-1195	K1 +G D6	144	0.48	0.44	0.51	32.24	15.52	0.05	1.07
mmu-miR-1195	K1 +G D8	192	10.82	9.53	12.28	32.38	11.01	0.09	-3.44
mmu-miR-1195	K1 -G D2	48	0.51	0.47	0.55	32.16	15.42	0.05	0.97
mmu-miR-1195	K1 -G D4	96	0.56	0.53	0.60	32.05	15.27	0.05	0.83
mmu-miR-1195	K1 -G D6	144	0.37	0.34	0.41	32.32	15.87	0.06	1.43

mmu-miR-1195	K1 -G D8	192	0.83	0.71	0.97	32.43	14.71	0.11	0.26
hsa-miR-1275	K1 +G D2	48	1.00	0.81	1.23	29.22	13.10	0.11	0.00
hsa-miR-1275	K1 +G D4	96	1.31	1.16	1.47	28.90	12.72	0.06	-0.39
hsa-miR-1275	K1 +G D6	144	1.56	1.24	1.95	28.66	12.47	0.12	-0.64
hsa-miR-1275	K1 +G D8	192	0.02	0.01	0.03	35.45	18.88	0.34	5.78
hsa-miR-1275	K1 -G D2	48	1.04	0.95	1.14	29.04	13.04	0.05	-0.06
hsa-miR-1275	K1 -G D4	96	0.69	0.58	0.81	29.67	13.65	0.09	0.54
hsa-miR-1275	K1 -G D6	144	1.02	0.88	1.18	29.19	13.07	0.08	-0.03
hsa-miR-1275	K1 -G D8	192	0.03	0.01	0.05	34.73	18.39	0.36	5.29

APPENDIX B – SUPPLEMENTARY MATERIAL FOR CHAPTER 4

Supplementary Table 1

Raw data corresponding to Figure 36, which includes Mean Fluorescence Intensity (MFI) of GFP+ events, %GFP+ cells, viable cell density, Viability, MFI for all events and a normalized measure of MFI/VCD

Mean Fluorescence Intensity (MFI)					
	AVERAGE			STDEV	
	37C	31C	31C/37C	37C	31C
UBC	89.475	127.635	143%	0.78	18.52
EF1a	1156.15	1275.345	110%	4.65	21.12
CMV	157.875	191.565	121%	16.34	13.51
Only Mirus	31.72	15.895		12.37	1.66
UTF	23.875	25.725		4.31	12.03
%GFP+ cells					
	AVERAGE			STDEV	
	37C	31C	31C/37C	37C	31C
UBC	16.75	26.975	161%	1.51	1.21
EF1a	54.925	53.34	97%	1.32	1.57
CMV	20.1	28.07	140%	0.88	0.59
Only Mirus	0.41	0.375		0.03	0.05
UTF	0.29	0.315		0.03	0.19
Viable Cell Density					
	AVERAGE			STDEV	
	37C	31C	31C/37C	37C	31C
UBC	2.89E+06	2.55E+06	88%	1.20E+05	6.55E+04
EF1a	2.62E+06	2.30E+06	88%	7.03E+03	7.40E+04
CMV	2.91E+06	2.72E+06	93%	5.48E+04	4.14E+04
Only Mirus	4.49E+06	2.92E+06		1.49E+06	1.94E+05
UTF	3.88E+06	3.15E+06		3.85E+05	2.95E+04
Viability					
	AVERAGE			STDEV	
	37C	31C		37C	31C
UBC	98.4	98.225		0.21	0.25
EF1a	93.6	92.975		0.99	0.04
CMV	97.675	98.025		0.04	0.39
Only Mirus	98.525	98.775		0.04	0.04
UTF	97.45	98.9		0.35	0.57

(continues next page)

MFI (all events)					
	AVERAGE			STDEV	
	37C	31C	31C/37C	37C	31C
UBC	17.67	36.74	208%	1.24	3.51
EF1a	636.59	681.68	107%	12.69	8.75
CMV	34.31	56.21	164%	1.92	4.89
Only Mirus	2.41	2.32		0.04	0.08
UTF	2.14	2.16		0.02	0.04
MFI / VCD					
	AVERAGE				
	37C	31C	31C/37C		
UBC	6.11E-06	1.44E-05	235%		
EF1a	2.43E-04	2.97E-04	122%		
CMV	1.18E-05	2.07E-05	175%		
Only Mirus	5.37E-07	7.94E-07			
UTF	5.50E-07	6.86E-07			

Supplementary Table 2

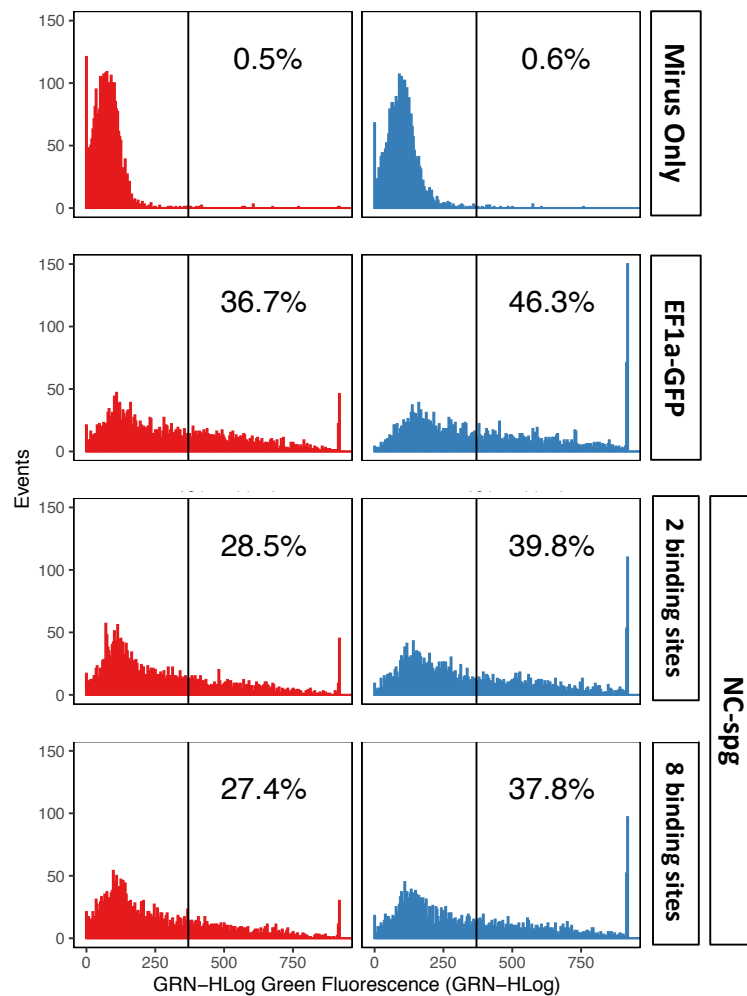
Viable cell density, Viability, Mean Fluorescence Intensity and percentage of GFP positive cells for CHO-K1 cells transfected with a CMV-driven GFP plasmid (N44) or an EF1a-driven GFP plasmid (EF1a-GFP). A Mirus™ only (MO) mock transfection was used as negative control and for flow cytometry gating.

Viable Cell Density (VCD)		24h after TF	T0	24h at 37C	+24h TS	+48h TS
Temperature	Time (h)	-24	0	24	48	72
37C	MO	2.82E+06	2.00E+05	6.93E+05	2.06E+06	1.62E+06
37C	N44	2.17E+06	2.00E+05	5.98E+05	1.82E+06	1.39E+06
37C	EF1a-GFP	2.34E+06	2.00E+05	5.77E+05	1.75E+06	1.40E+06
31C	MO		2.00E+05	6.93E+05	1.56E+06	1.21E+06
31C	N44		2.00E+05	5.98E+05	1.35E+06	1.17E+06
31C	EF1a-GFP		2.00E+05	5.77E+05	1.06E+06	1.14E+06
Viability (%)		24h after TF	T0	24h at 37C	+24h TS	+48h TS
Temperature	Time (h)	-24	0	24	48	72
37C	MO	97.54	97.54	96.73	98.33	95.33
37C	N44	82.18	82.18	90.235	97.78	93.42
37C	EF1a-GFP	85.06	85.06	90.835	97.7	94.25
31C	MO		97.54	96.73	98.53	94.64
31C	N44		82.18	90.235	96.51	94.57
31C	EF1a-GFP		85.06	90.835	96.59	94.29
VCD (STDEV)		24h after TF	T0	24h at 37C	+24h TS	+48h TS
Temperature	Time (h)	-24	0	24	48	72
37C	MO			3.59E+04	1.67E+04	7.66E+04
37C	N44			2.19E+04	7.88E+04	8.37E+04
37C	EF1a-GFP			5.05E+03	3.50E+04	2.67E+05
31C	MO			3.59E+04	5.25E+04	7.65E+04
31C	N44			2.19E+04	5.03E+04	6.50E+04
31C	EF1a-GFP			5.05E+03	4.98E+04	7.76E+04
Viability (STDEV)		24h after TF	T0	24h at 37C	+24h TS	+48h TS
Temperature	Time (h)	-24	0	24	48	72
37C	MO			0.57	0.10	0.38
37C	N44			0.66	0.45	0.25
37C	EF1a-GFP			0.23	0.57	0.38
31C	MO			0.57	0.13	0.57
31C	N44			0.66	0.16	0.35
31C	EF1a-GFP			0.23	0.18	0.35

MFI (Y-Mean)						
Temperature	Time (h)	-24	0	24h at 37C	+24h TS	+48h TS
37C	MO	34.11		152.17	70.53	34.125
37C	N44	1570.27		987.175	391.74	181.13
37C	EF1a-GFP	1748.98		995.39	355.75	178.48
31C	MO			152.17	33.36	33.64
31C	N44			987.175	482.045	282.865
31C	EF1a-GFP			995.39	427.97	239.46
%GFP+ cells						
Temperature	Time (h)	-24	0	24h at 37C	+24h TS	+48h TS
37C	MO	0.3		0.06	0.29	1.12
37C	N44	76.2		48.44	29.16	10.77
37C	EF1a-GFP	69.34		47.7	26.14	9.16
31C	MO			0.06	0.19	0.74
31C	N44			48.44	36.27	19.84
31C	EF1a-GFP			47.7	32.23	16.63
I (Y-Mean) (STDEV)						
Temperature	Time (h)	-24	0	24h at 37C	+24h TS	+48h TS
37C	MO			59.99	55.80	1.53
37C	N44			4.77	18.95	15.40
37C	EF1a-GFP			33.63	12.52	12.13
31C	MO			59.99	13.44	0.41
31C	N44			4.77	46.39	11.22
31C	EF1a-GFP			33.63	10.41	32.92
GFP+ cells (STDEV)						
Temperature	Time (h)	-24	0	24h at 37C	+24h TS	+48h TS
37C	MO			0.06	0.04	0.06
37C	N44			0.17	0.48	0.64
37C	EF1a-GFP			1.22	1.02	0.25
31C	MO			0.06	0.07	0.14
31C	N44			0.17	0.35	0.65
31C	EF1a-GFP			1.22	0.35	0.21

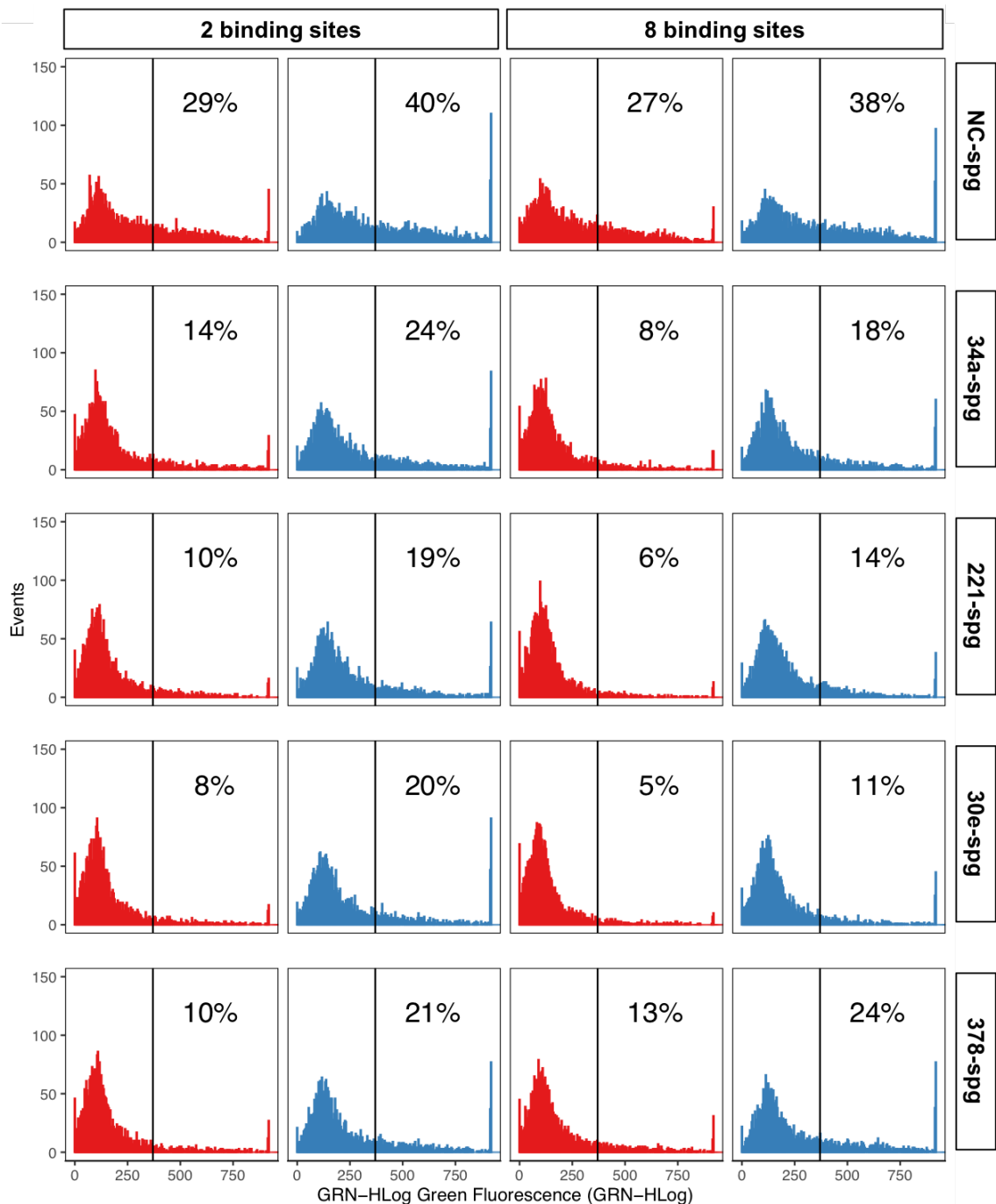
Supplementary Figure 3

Density plots showing flow cytometry data (Green channel – HLog) for transiently transfected populations using only transfection reagent (Mirus™ only), an EF1a-GFP positive control and the two negative control NC-sponges with 2 and 8 binding sites. Transfected populations were subjected to temperature shift (37C – red / 31C blue). A total of 5000 events were recorded for each sample and triplicate samples were analysed (representative case showed).



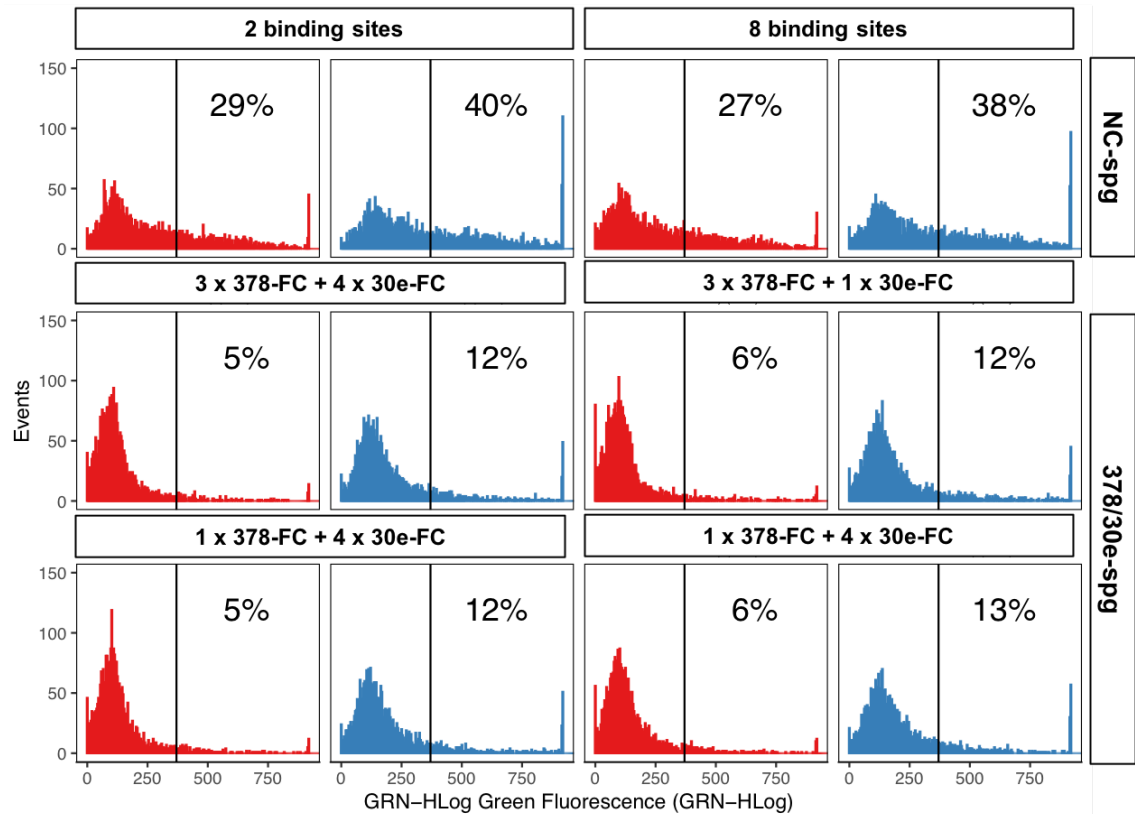
Supplementary Figure 4

Histogram plots showing flow cytometry data (Green channel – HLog) for transiently transfected populations using the different miRNA-sponges. Transfected populations were subjected to temperature shift (37C – red / 31C blue). A total of 5000 events were recorded for each sample and triplicate samples were analysed (representative case showed).



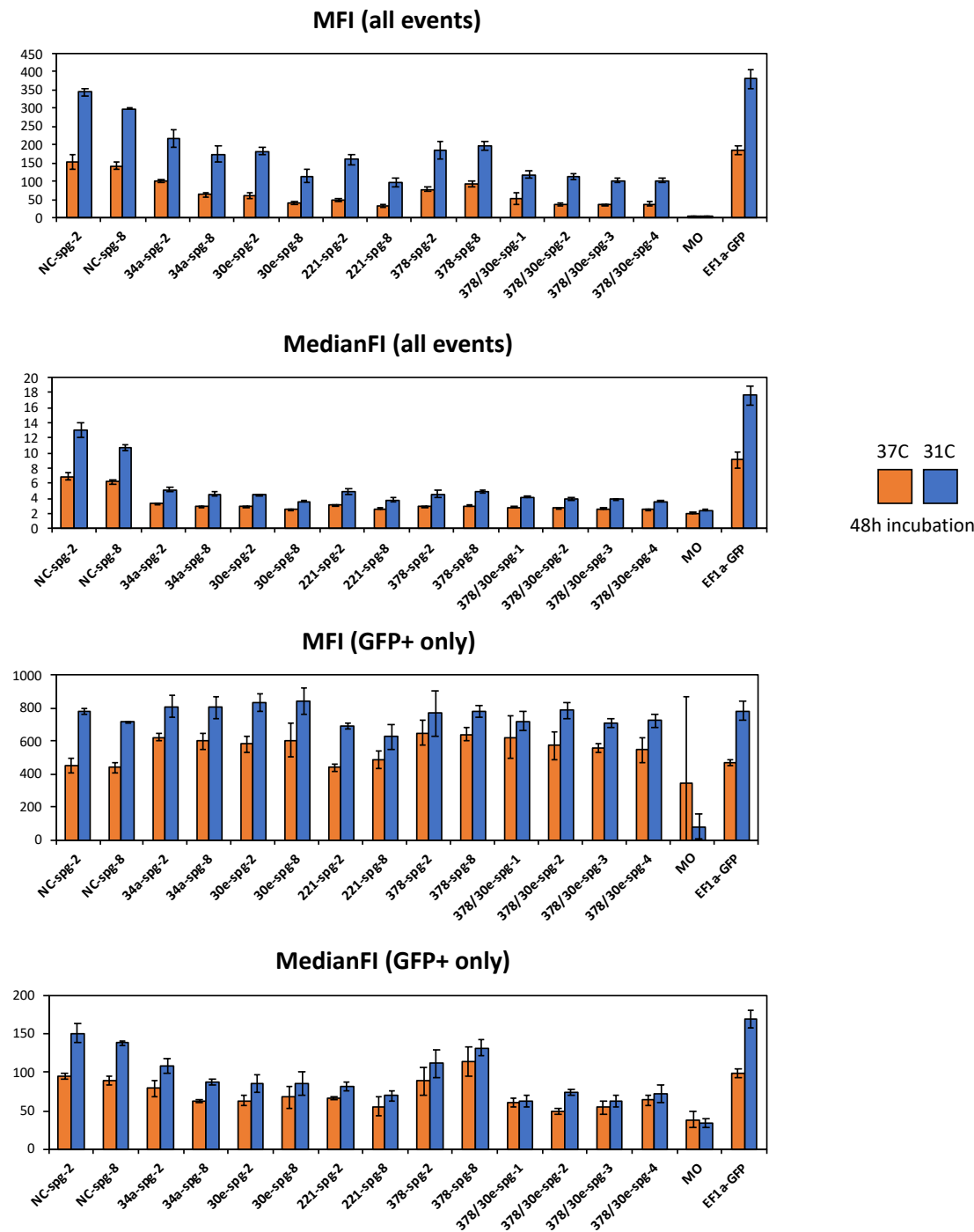
Supplementary Figure 5

Histogram plots showing flow cytometry data (Green channel – HLog) for stable pools expression the different hybrid miRNA-sponges (cgr-miR-378-3p / cgr-miR-30e-5p). A total of 5000+ events were recorded for each.



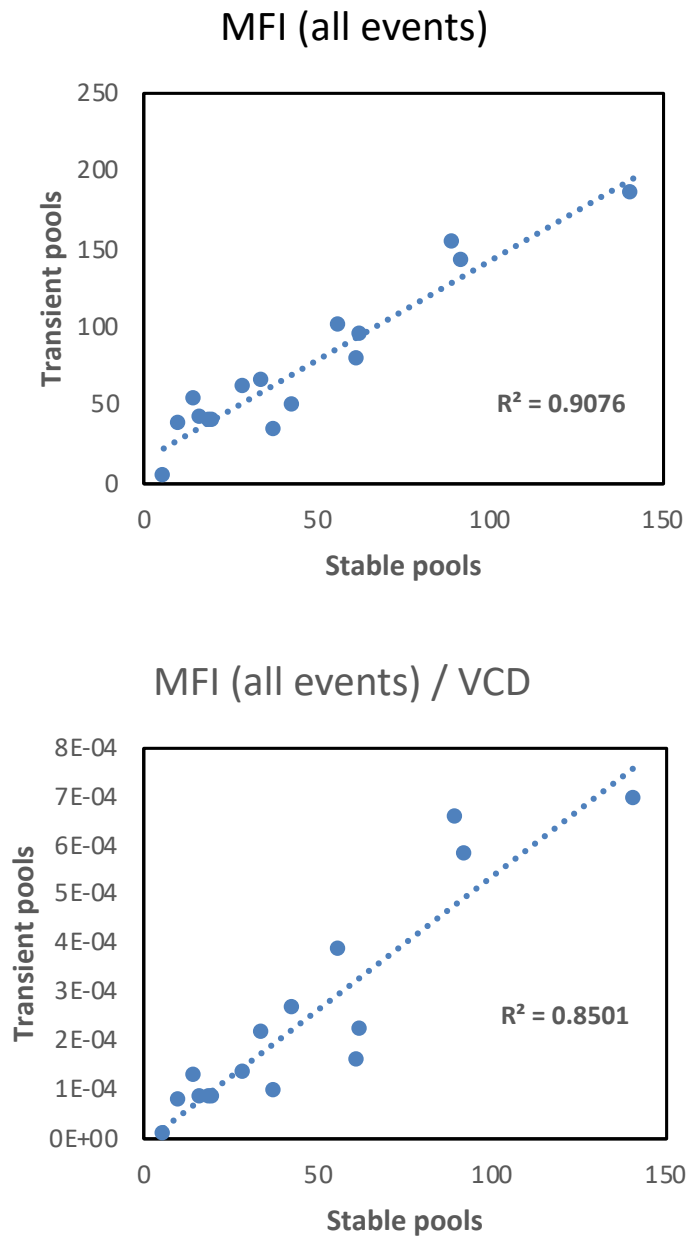
Supplementary Figure 6

Different available parameters to represent GFP expression of the different populations transiently transfected with sponges.



Supplementary Figure 7

Correlation between GFP expression from the different miRNA-sponges observed in transiently transfected cells versus GFP expression observed in the respective stable mixed pools. Correlations were calculated for both raw MFI (all events) and MFI (all events normalised by VCD).



Supplementary Figure 8

Correlation between the GFP reduction observed driven by the different miRNA-sponges in transiently transfected cells versus GFP reduction observed in the respective stable mixed pools. Top plot shows correlations calculated between a representative passage (P4) and data from 3 different time points in the transient pools. Bottom plot shows correlations between miRNA-sponge driven GFP reductions observed for different passages of the stable pools (P4 to P6).

