



The
University
Of
Sheffield.

High-Throughput Platform Development for Multigene Engineering of Chinese Hamster Ovary (CHO) Cells

Claire Lucy Arnall

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Engineering
Department of Chemical and Biological Engineering

May 2018

Abstract

Within biopharmaceutical manufacturing pipelines, there are increasing numbers of “next-generation” and engineered recombinant protein products. With these come more problematic products described as “difficult-to-express” (DTE), characterised by poor production yields. With the aim of improving the performance of the Chinese hamster ovary (CHO) cell factory, cellular functions can be engineered by the simultaneous modulation of multiple product-specific components likely to influence protein production. Therefore, a tool to identify specific combinations of components and determine their optimal relative functional stoichiometry is desirable. Described here is the development of a standardised high-throughput transient transfection platform for this application. The optimisation of a multi-well electroporation plate method, enabled hundreds of high-efficiency, micro-scale co-transfections, each with multiple genes at particular ratios to be undertaken, in-parallel. This was interfaced with the rapid evaluation of growth and productivity to ascertain functional effects.

To quantitatively investigate multi-plasmid co-transfection at specified stoichiometry, co-transfections with three separate fluorescent protein plasmids were compared to transfection with a single plasmid carrying all three genes. The use of separate plasmids demonstrated that high proportions of cells conformed with close adherence in expression, although with more variability, to the targeted transfected gene ratio. Stoichiometric changes were consistently achieved by altering the levels of co-transfected plasmids, validating the utility of the platform to assess effects of multiple genes.

With the aim of improving production of an exemplar DTE IgG, the developed platform was implemented for rapid auditioning of a library of ‘effector genes’ associated with a desirable antibody-producing phenotype. Upregulation of XBP1s demonstrated the most beneficial solution, a 129% improvement to DTE IgG titre. Overexpression of individual effector genes demonstrated responses that were dependent upon their relative titration. A Design of Experiments statistical modelling approach could be implemented to investigate stoichiometry of multiple co-transfected components. The platform is an effective strategy to design novel multigene assemblies for CHO cell engineering.

Acknowledgements

The processes undertaken to reach this point have been made possible with the help and support of various people and organisations directly and indirectly. I would like to acknowledge their help, for which I am extremely grateful.

To the Bioprocessing Research Industry Club (BRIC), Biotechnology and Biological Sciences Research Council (BBSRC) and Medimmune for providing essential funding to enable the research to be undertaken. Additional recognition goes to BRIC for the unique and benefitting opportunities, industrial insights and events offered to me throughout the studentship.

To my supervisors at my industrial collaborator MedImmune, Dr Diane Hatton and Dr Claire Harris, for your unwavering help, project guidance, industrial perspective and proof-reading throughout the project. I feel you went above and beyond your duty, for which I am very grateful. Additionally, thanks to Dr Ray Field, Dr Greg Dean and the rest of the Cell Line Development (CLD) team at MedImmune for their input in various ways; along with the flow cytometry support from Dr Guglielmo Rosignoli and Dr John Ferguson. It has been a pleasure to work with you all. Beneficially MedImmune provided access to their cell lines and plasmids, and the opportunity to use equipment and their laboratories; necessities for the progression of this project.

To Professor David James, for the opportunity to study for a PhD within your laboratory (DCJ research group) at The University of Sheffield and for providing continued supervision, guidance, knowledge, critique and enthusiasm throughout.

To the numerous members past and present of the DCJ research group. Thank-you for your troubleshooting when needed, helpful advice, knowledge sharing, encouragement and, perhaps most of all, friendship, working with you all over the last four years has been a pleasure. An extra acknowledgement goes to the other members of the “MedImmune team”, Dr Adam Brown, Dr Joe Cartwright, Yash Patel and Nic Barber - for your collaborations, assistance and resource sharing; and, along with Devika Kalsi, Katie Syddall, Joanne Longster, Darren Geoghegan and Joseph Longworth for many a lunch-time and cake breaks and making the

Acknowledgements

whole experience just all the more enjoyable. Extended thanks go to the members of the D72 office who made for a very pleasant working environment. Gratitude also goes to Dave Wengraf, for his continued and extensive technical laboratory support, you definitely helped to keep the lab running and the incubators rotating!

To those who have taken the time to review a draft of this thesis, your feedback and comments were very useful and much appreciated.

To my many friends and extended family including the new Arnall contingent. I have been very grateful for the support and encouragement throughout this process. Thank-you to those who have shared their company for an excursion out in the hills, a board-game, a weekend getaway or just a cocktail or two, it has been much enjoyed. A mention also needs to go to the feline companions, past and present, near and afar, for their companionship and providing many a comical moment.

To my Mum, Dad and sister, Emma – you are just the best, and I cannot thank-you enough for all that you do. Thank-you for always seeing the funny side, providing much needed reality and perspective, providing many a holiday opportunity and encouraging a love for the much-needed outdoors. And for trying to understand the use of hamster ovaries to make “mono-body anti-clones”!

And finally, but by no means least, to my husband Andy, who has journeyed this PhD together with his own. Thank-you for everything always. For your ceaseless love, strength, support and smile.

Contents

Abstract	i
Acknowledgements	iii
Contents	v
Table of Figures	x
Table of Tables	xiii
Acronyms and Abbreviations	xv
1. Introduction	1
1.1. Thesis Overview.....	2
2. The Production and Development of Biopharmaceuticals	4
2.1. Biopharmaceuticals.....	4
2.1.1. mAbs as Therapeutic Proteins.....	5
2.1.2. Novel Therapeutic Molecules.....	7
2.2. Manufacture of Recombinant Protein Therapeutics.....	8
2.2.1. Expression Systems for Biopharmaceutical Manufacture.....	8
2.2.2. CHO Cells.....	10
2.3. Biopharmaceutical Production.....	12
2.3.1. Expression Plasmid.....	12
2.3.2. Transfection Approach.....	15
2.3.3. Stable Versus Transient Expression.....	18
2.3.4. Cloning Process.....	21
2.3.5. Upstream Manufacturing Process.....	22
2.3.6. Downstream Processes.....	23
2.4. Cellular Level Processes in Mammalian Recombinant Protein Production.....	23
2.4.1. ER Protein Folding and Assembly.....	24
2.4.1.1. Polypeptide Translocation into the ER Lumen.....	25
2.4.1.2. Polypeptide Modification and Folding in the ER Lumen.....	26
2.4.2. ER Quality Control Mechanisms.....	29
2.4.3. The Secretory Pathway.....	32
2.4.4. Post-translational Bottleneck.....	34
2.5. Summary.....	35
3. Improvements in Therapeutic Protein Production	36

3.1. Introduction	36
3.2. A Need to Engineer the Production Host.....	36
3.3. Approaches to Enhance Production Levels.....	39
3.4. Cell Engineering Approaches to Enhance CHO Cell Factory Performance	41
3.4.1. Development of Omics-based Approaches	43
3.4.2. Repression of Genetic Components.....	43
3.4.3. Overexpression of Genetic Components.....	45
3.4.3.1. Anti-Apoptotic Engineering	45
3.4.3.2. Engineering Post-Translational Mechanisms.....	48
3.4.3.3. Engineering Components of the UPR	53
3.4.3.4. Engineering the Secretory Pathway	54
3.4.3.5. Other Upregulated Components.....	54
3.4.3.6. Multiple Component Engineering	59
3.4.3.7. Effector Gene Dosing	59
3.4.4. Synthetic Biology Paradigm	60
3.5. Conclusion	60
4. Materials and Methods.....	63
4.1. Plasmid DNA Preparation and Cloning	63
4.1.1. DNA Ligation.....	63
4.1.2. Transformation and Amplification of Plasmid DNA	63
4.1.3. Plasmid Isolation and Purification.....	64
4.1.4. Quantification of Plasmid DNA	65
4.1.5. Restriction Digestion of Plasmid DNA	65
4.1.6. Gel Electrophoresis.....	65
4.1.7. Gel Extraction	66
4.1.8. Sequence Verification	66
4.1.9. Commercial DNA Synthesis	67
4.2. CHO Cell Culture.....	67
4.2.1. CHO Cell Lines	67
4.2.2. Subculture.....	67
4.2.3. Cryopreservation and Cell Revival	68
4.3. Transient Transfection.....	68
4.3.1. Plate-based Electroporation for MEDI-CHO Cells	69
4.3.2. Deep-Well Post-Electroporation Culture.....	70
4.3.3. Lipofection	71

4.4. PrestoBlue® Cell Viability Assay	71
4.5. Flow Cytometry	73
4.5.1. Attune Acoustic Focusing Cytometer	74
4.5.2. BD LSR Fortessa Flow Cytometer	74
4.6. Quantification of Volumetric Titre	75
4.6.1. Quantification of IgG Titre	76
4.6.2. Quantification of Secreted Embryonic Alkaline Phosphatase Titre	77
4.7. Cell Culture Parameter Equations	78
4.7.1. Equations to Calculate Cell Growth.....	78
4.7.2. Equations to Calculate Productivity.....	78
4.8. Statistical Analysis.....	78
4.8.1. Coefficient of Variation.....	79
4.8.2. Design of Experiments.....	80
5. Development of a High-Throughput Transient Transfection Screening Platform.....	81
5.1. Introduction.....	81
5.2. Experimental Approach	86
5.3. Results	89
5.3.1. Plasmid Development to Enable Standardised Gene Expression	89
5.3.1.1. 'Expression-Plasmid' Construction	90
5.3.1.2. Verification of 'Expression-Plasmid' Functionality	91
5.3.2. High-throughput Electroporation Optimisation.....	94
5.3.2.1. Nucleofector Programme	94
5.3.2.2. DNA Load and Cell Number Optimisation	95
5.3.3. Comparison of Plate-Based Electroporation to Plate-Based Lipofection	105
5.3.4. Post-Transfection Plate-Based Culture	106
5.3.5. Plate-Based Parameter Measurements	109
5.3.6. Platform Variation	111
5.4. Discussion	117
5.4.1. Developed Transfection Platform.....	118
5.4.2. Comparison of Transfection Setup.....	120
5.4.3. Plate-Based Assays to Assess Post-Transfection Cell Culture Performance	122
5.4.4. Variation Within Platform.....	123
5.4.5. Developments in Shaking Plate Based Culture	124

5.5. Summary.....	126
6. Validation of Multiple Gene Transfection	127
6.1. Introduction	127
6.2. Experimental Approach.....	130
6.2.1. Fluorescent Protein Expression Plasmids	131
6.3. Results	133
6.3.1. Development of the Fluorescent Reporter Strategy.....	133
6.3.2. Fluorescent Protein Flow Cytometry Detection and Analysis.....	137
6.3.2.1. Challenges and Limitations in the Detection of Three Fluorescent Proteins	141
6.3.3. Normalisation to Adjust for Inherent Differences Between Fluorescent Reporter Proteins.....	142
6.3.4. Expression Adherence to Transfections with Equal Copies of Genes...	149
6.3.5. Alteration in the Ratio of Plasmids Transfected.....	153
6.4. Discussion.....	158
6.4.1. The Use of Fluorescent Proteins to Analyse Multiple Gene Expression	159
6.4.2. Co-transfection with Multiple Plasmids Compared to a Single Multigene Plasmid.....	161
6.4.3. Stoichiometry Attained Through Co-Transfection of Separate Plasmids	164
6.5. Summary.....	165
7. Implementation of the Developed Platform for Effector Gene Screening	166
7.1. Introduction	166
7.2. Experimental Approach.....	169
7.3. Results	172
7.3.1. Development of Effector Gene Library	172
7.3.2. Range of mAb Expression with Platform Setup	177
7.3.3. Effect of Multiple Gene Expression	180
7.3.4. Initial Effector Gene Screen Demonstrated the Importance of Transfect Gene Dose	181
7.3.5. Combinatorial Expression of Effector Genes.....	184
7.3.5.1. Two-Level Factorial Design	184
7.3.5.2. Central Composite Response Surface Model Design.....	188
7.3.6. Rapid Auditioning of the Effector Gene Library.....	194

7.3.6.1. Use of Deep-Well Shaking Culture Mode Improved Post-Transfection Culture	194
7.3.6.2. High-Throughput Effector Component Screen	196
7.4. Discussion	206
7.4.1. Model DTE-mAb	207
7.4.2. Overexpression of XBP1s	209
7.4.3. Effector Gene Library.....	210
7.5. Summary.....	211
8. Conclusions and Further Work	213
8.1. Summary Discussion and Conclusions	213
8.2. Alternative Platform Uses.....	217
8.3. Future Work.....	220
References	224
Appendix A.....	237
Appendix B.....	240
Appendix C.....	243

Table of Figures

Figure 2.1: Outline of the Typical Biopharmaceutical Production Processes Undertaken During Recombinant Protein Manufacture.	13
Figure 2.2: An Outline of the Processes Undertaken to Assemble Proteins in the ER.	25
Figure 2.3: A Schematic Outlining Components of the Unfolded Protein Response (UPR)	31
Figure 2.4: Schematic of the Secretory Pathway in Mammalian Cells.....	34
Figure 3.1: Schematic Demonstrating the Range of Engineered Recombinant Proteins in Clinical Trials	39
Figure 3.2: Targets of CHO Cell Engineering.....	42
Figure 3.3: Engineering of Protein Folding and Assembly in the ER	49
Figure 3.4: Engineering of UPR	53
Figure 4.1: Example Fast-ELISA Assay Standard Curve	76
Figure 4.2: Example SEAP Assay Standard Curve	77
Figure 5.1: Determination of GFP Expression using Attune Flow Cytometry Detection	88
Figure 5.2: 'Expression-Plasmid' Outline Map.....	90
Figure 5.3: Example Insert Sequence Compatible with 'Expression-Plasmid'	92
Figure 5.4: GFP Expression Verifies Functionality of Constructed 'Expression-Plasmid'	93
Figure 5.5: Programme FF-158 was the Most Suitable Nucleofector Programme Tested for Rapid, High-Level Introduction of Plasmid DNA.....	95
Figure 5.6: 0.9-1.2 µg of Plasmid DNA was the Optimal pmaxGFP Plasmid Load From a Range Based on Manufacturer's Recommendations	96
Figure 5.7: Transfections with mAb Encoding Plasmid DNA had Negative Effects on Post-Transfection Culture Viability	97
Figure 5.8: Three-Dimensional Surface Graphs Derived from DNA Load and Cell Number Optimisation Using DoE-RSM	101
Figure 5.9: Multiplexing Transfected Plasmid Type to Determine Final Transfection DNA Load	104
Figure 5.10: Comparison of Electroporation and Lipofection Transfection Methods Demonstrates Electroporation is a More Suitable Transient Transfection Approach for Rapid, High Level Gene Overexpression.....	106
Figure 5.11: 96-Well Micro-Plate Static Culture Duration is Limited	107
Figure 5.12: Comparison Between Transfections Using MEDI-CHO and CHO-T2 Cells Demonstrates Very Close Comparability of Transfection Between Cell Lines	109

Figure 5.13: Positive Correlation Between Vi Cell Trypan Blue Exclusion and (A) PrestoBlue® Resazurin Assay and (B) TO-PRO®3 Viability Assessment.	111
Figure 5.14: Experimental Outline to Investigate Platform Variation.....	112
Figure 5.15: Box Plots Representing Growth Rates from Three Separate Whole-Plate Transfection Runs with Each Transfection Seeded into Three Separate Culture Plates.....	114
Figure 5.16: Correlations in Growth Rate as A Result of Seeding Multiple Culture Wells from the Same Transfection Well.....	115
Figure 5.17: Increasing the Number of Transfection Replicates Within a Transfection Run Reduces Transfection Variation (CV).....	117
Figure 6.1: Fluorescent Protein Plasmid Maps.....	133
Figure 6.2: E2-Crimson is Not a Suitable Reporter Protein for Multiplexing with TagBFP Expression	136
Figure 6.3: Excitation and Emission Spectra and Flow Cytometer Configuration Used for TagBFP, eGFP and mCherry Detection.....	137
Figure 6.4: Flow Cytometry Gating to Isolate Triple Fluorescent Events	140
Figure 6.5: Limitations in Detection from Multiplexing Fluorescent Protein Reporters	142
Figure 6.6: Fluorescent Reporter Plasmid Titration	144
Figure 6.7: Transfection Efficiency is Consistent Between Transfections and Plasmids	146
Figure 6.8: Histograms Representing the Spread of Fluorescent Protein Expression at a Cellular Level.....	150
Figure 6.9: The Adherence of Cellular Expression to the Transfected Plasmid Ratio (1:1:1).....	153
Figure 6.10: Histograms Representing Cell Fluorescent Protein Expression from Alterations in the Plasmid Ratio Transfected	156
Figure 6.11: The Conformity of Cellular Expression with Alterations in Transfected Plasmid Ratios	158
Figure 7.1: Effector Component Screening Strategy	171
Figure 7.2: Overview of Effector Genetic Components	176
Figure 7.3: Effect of Varying Transfected mAb-Plasmid Type and Copy Number	179
Figure 7.4: Alteration in eGFP-‘Expression-Plasmid’ Levels Demonstrated Effects Upon Recombinant mAb Co-Expression.....	181
Figure 7.5: Static Screen: Effects from Co-Transfected Effector Genes at Different Plasmid Doses	182
Figure 7.6: The Effect from Multiple Effector Component Co-Transfections.....	185
Figure 7.7: The Effect from Simultaneous Co-Transfections of Multiple Effector Components at Differing Levels.....	190

Figure 7.8: Three-Dimensional Surface Graphs Representing Titre and qP Responses from Co-Transfections with CypB, Largen and XBP1s.....	192
Figure 7.9: Shaking, Fed-Batch Culture Mode Using Deep, 96-Well Plates	196
Figure 7.10: Responses from Co-Transfection of ER Protein Processing Effector Genes at Different Doses.....	198
Figure 7.11: Responses from Co-Transfection of Golgi Complex Processing and Vesicular Transport Effector Genes at Different Doses.....	201
Figure 7.12 Responses from Co-Transfection of Transcription Factor Effector Genes at Different Doses.....	203
Figure 7.13: Clear Dose-Dependent Response from XBP1s (C) Overexpression	204
Figure 7.14: Overexpression of CHO and Human XBP1s Demonstrated Very Similar Responses	206
Figure 8.1: The Platform Can Be Used to Investigate the Functionality of Synthetic Parts and the Manufacturing Performance of Novel Cell Lines	219
Appendix Figure A.1: Construction of 'Expression-Plasmid'	237
Appendix Figure A.2: Construction of eGFP 'Expression-Plasmid'	238
Appendix Figure A.3: Plasmid Maps of eGFP-C3 and SEAP Plasmids.....	238
Appendix Figure A.4: Normal Plots of Residuals for DoE-RSM: Cell Number and DNA Load	239
Appendix Figure B.1: Excitation and Emission Spectra for E2-crimson.....	241
Appendix Figure B.2: Flow cytometry Analysis of Single Fluorescent Protein Samples	242
Appendix Figure C.1: Normal and Half Normal Probability Plots for DoE Two-Level Factorial Design: CypB, ERo1L α , Largen and XBPs.....	260
Appendix Figure C.2: Normal Plots of Residuals for DoE-RSM: CypB, Largen and XBPs Normality plots for: IVCD (A) Titre (B) and qP (C).....	261
Appendix Figure C.3: XBP1s Amino Acid Sequence Alignment for human (XBP1s (H)) and CHO (XBP1s (C)) Proteins	265

Table of Tables

Table 2.1: Top Ten Best-Selling Recombinant Biopharmaceuticals in 2013.....	6
Table 2.2: Different Expression Systems Utilised for Recombinant Protein Production .	9
Table 2.3: Common Transfection Approaches.....	17
Table 2.4: Resident ER Components.....	27
Table 3.1: Anti-Apoptotic Engineering Approaches.....	47
Table 3.2: ER Protein Folding and Assembly Engineering Approaches	51
Table 3.3: UPR Engineering Approaches	55
Table 3.4: Secretory Pathway Engineering Approaches	57
Table 3.5: Other Cellular Engineering Approaches.....	58
Table 4.1: Example PrestoBlue Assay Analysis.....	73
Table 4.2: Example Compensation Matrix Table for Tag-BFP, mCherry and eGFP Detection and Analysis.....	75
Table 5.1: Levels of Factors Used in Central Composite Design Used to Optimise Transfection DNA Load and Cell Number.....	99
Table 5.2: Summary of Response Model Analysis of Outputs from the Central Composite Design Used to Optimise Transfection DNA Load and Cell Number.....	100
Table 6.1: Transfection Setup Used to Investigate Multigene Transfection.....	131
Table 6.2 Example of Normalisation Undertaken to Adjust for Total Fluorescence and Convert the Dataset into Comparable Ratios.....	147
Table 6.3: The Average and Variance in Fluorescent Protein Expression at a Cellular Level	151
Table 6.4: Example of Normalisation Undertaken to Adjust for Differing Fluorescence Intensities Derived from Altering the Transfected Plasmid Ratio.....	155
Table 6.5: The Average and Variation in Cellular Fluorescent Protein Expression from Alterations in the Ratio of Plasmids Transfected.....	157
Table 7.1: Multiple Information Sources Guided Effector Component Choice.....	174
Table 7.2: Effector Gene Library.....	175
Table 7.3: Analysis Summary of Outputs from the Two-Level Factorial Design Used to Investigate Combinations of Effector Genes.....	187
Table 7.4: Levels of Factors Used in Central Composite Design Combining CypB, Largen and XBP1s Effector Genes.....	189
Table 7.5: Summary of Response Model Analysis of Outputs from the Central Composite Design.....	191
Appendix Table B.1: Nucleotide Sequences for Fluorescent Proteins.....	240

Table of Tables

Appendix Table C.1: Effector Gene Library Information: Protein Processing in the ER	243
Appendix Table C.2: Effector Gene Library Information: Golgi Complex/Vesicular Secretion Processes.....	246
Appendix Table C.3: Effector Gene Library Information: Other	248
Appendix Table C.4: Effector Gene Library Information: Transcription Factors	249
Appendix Table C.5: Effector Gene Library Nucleotide Sequences.....	250
Appendix Table C.6: Summary Matrix for DoE Two-Level Factorial Design: Co- Expression of CypB (C), ERo1L α (E), Largen (L) and XBPs (X).....	259
Appendix Table C.7: Results of Ordinary One-Way ANOVA With Dunnett's Multiple Comparisons Test	262

Acronyms and Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
AmpR	Ampicillin resistance
ANOVA	Analysis of Variance
AT-III	Antithrombin-III
ATF4	Activating transcription factor 4
ATF6	Activating transcription factor 6
ATF6c	Cleaved activating transcription factor 6
BFA	Brefeldin A
BFP	Blue fluorescent protein
BiP	Binding immunoglobulin protein
bp	Base pairs
(C)	CHO protein sequence
Calr	Calreticulin
Canx	Calnexin
CERT	Ceramide-transfer protein
CHO	Chinese hamster ovary
CHOP	Transcription factor C/EBP homologous protein
CMV	Cytomegalovirus
COPI	Coat protein complex I
COPII	Coat protein complex II
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats-associated protein-9 nuclease
CV	Coefficient of variation
DAPI	4',6 Diamidino-2-phenylindole
DHFR	Dihydrofolate reductase
DMSO	Dimethyl sulfoxide
DoE	Design of Experiments
DoE-RSM	Design of experiments response surface model
DTE	Difficult-to-express
EBNA-1	Epstein-barr virus nuclear antigen 1
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
ERES	Endoplasmic reticulum exit sites
ERGIC	Endoplasmic Reticulum to Golgi Intermediate Compartment
ERO1	Endoplasmic reticulum oxidoreductase 1
ETE	Easy-to-Express
FACS	Fluorescent activated cell sorting
FDA	Food and Drug Administration
FSC	Forward scatter
GADD34	Growth arrest and DNA damage inducible protein 34
GAP	Deactivating GTPase activating protein
GEF	Activating guanine nucleotide exchange factor
GFP	Green fluorescent protein
GMP	Good manufacturing practice
GS	Glutamine synthetase
(H)	Human protein sequence
HC	Heavy chain
HF	High fidelity

Hsp	Heat shock protein
HSP27	Heat shock protein 27
IRE1	Inositol-requiring protein 1
LC	Light chain
mAb	Monoclonal antibody
MFI	Median fluorescence intensity
miRNA	microRNA
MSX	Methionine sulfoximine
MTX	Methotrexate
NCBI	National Centre for Biotechnology Information
OFAT	One Factor at A Time
ORF	Open reading frame
PBA	4-phenylbutyric acid
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDI	Protein Disulphide Isomerase
PEI	Polyethylenimine
PERK	Protein kinase RNA-like ER kinase
PMT	Photomultiplying tube
PPlases	Prolyl isomerase
PTM	Post-translational modifications
qP	Specific productivity
RNAi	Interfering RNA
rpm	Revolutions per minute
SEAP	Secreted embryonic alkaline phosphatase
SEM	Standard Error of the Mean
SNAP23	Synaptosome-associated protein of 23kDa
SNARE	Soluble N-ethylmaleimide sensitive factor receptors
SRP	Signal recognition particle
SSC	Side scatter
SSC-A	Side scatter area
SSC-H	Side scatter height
TALEN	Transcription activator-like effector nuclease
TGN	Trans Golgi network
tPA	Tissue plasminogen activator
TPO	Thrombopoietin
UCOEs	Ubiquitously acting chromatin opening elements
UPR	Unfolded protein response
UTR	Untranslated region
UV	Ultraviolet
v/v	Volume/volume
VAMP8	Vesicle-associated membrane protein 8
VCD	Viable cell density
VPA	Valproic acid
w/v	Weight/volume
WCB	Working cell bank
XBP1	X-box Binding Protein 1
XBP1s	Spliced X-box Binding Protein 1
XIAP	X-linked inhibitor of apoptosis
YY1	Ying-yang 1
ZFN	Zinc-finger nuclease

1. Introduction

The manufacture of biopharmaceuticals (recombinant proteins, commonly medicinal drugs, used to target diseases such as cancers and autoimmune diseases) predominantly uses Chinese hamster ovary (CHO) cells as the production host (Walsh, 2014). There is a continual pressure to improve the production system, to increase yields and reduce the costs involved, ultimately aiding the number of patients that can benefit from the drugs manufactured. The industry, established in the 1980s, has for the most-part improved the yields generated, with the typical amounts improving from milligram to gram per litre quantities (Wurm, 2004). Problematic products with poor production yields (“difficult-to-express” (DTE) products) are experienced and are heightened by the progression of “next generation” and engineered molecules. It is likely that these, characteristically low-production products, have limitations in the production processes that take place within the CHO cell factory itself.

One approach to improve the yield of these products is to engineer the cell itself and alter cellular functions directly. The use of genetic engineering to overexpress cellular components can alter the dynamics taking place within the cell. This concept is well-documented for CHO cells, with a few examples also demonstrating the use of multigene engineering approaches (Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008; Nishimiya, 2014). It is likely that when engineering any complex cellular system, simultaneous modulation and co-ordination of multiple components will be of benefit, with different combinations of components favourable for different products with different cellular requirements (Johari et al., 2015; Pybus et al., 2013). It is hypothesised that cell engineering will be a useful tool to improve the production levels of DTE products, with the level of the cellular component gene overexpressed important to consider and a greater functionality derived from the overexpression of multiple components (Hansen et al., 2017; Johari et al., 2015).

There is a need to identify functional cellular components and combinations of genes with added synergy, and determine their optimal relative stoichiometry to achieve the desired objective of improved recombinant protein production. Considering the great number of components and potential combinations of interest, a tool for this screening process is desirable (Hansen et al., 2015; Johari et al., 2015).

This thesis has three aims. Firstly, to develop a platform capable of high-throughput simultaneous screening of multiple genetic components, transiently transfected in a standardised manner. Secondly, to confirm the co-transfection of multiple plasmids to individual cells at a desired stoichiometry; to ensure multigene engineering could be tested. Thirdly, to use the platform, to demonstrate its use and to find a solution to improve the production of an industrial DTE monoclonal antibody (mAb).

1.1. Thesis Overview

In chapter two there is introductory overview of the biopharmaceutical industry; using literature to describe the processes to introduce and manufacture recombinant products in the predominant CHO cell host; and the processes undertaken at the cell level to assemble and secrete recombinant proteins.

Chapter three is a focussed review of the literature of CHO cell engineering to improve recombinant protein production. Promising outcomes from previous genetic overexpression studies are highlighted. These help to emphasise the need for screening tools to aid in testing the many components that could be targeted and a need to search for product-specific solutions. This information informed the basis of the design criteria for a high throughput transient transfection platform, and also identified some of the target genetic components for testing in the platform.

In chapter four the materials and methods used to undertake the work within this thesis are outlined.

The development of the high-throughput transient transfection platform is detailed in chapter five. The design and optimisation of a multi-well, micro-scale, standardised approach, for many engineering strategies to be evaluated at once,

is described. The transfection process is optimised for high DNA delivery (high transfection efficiency), while trying to maintain a healthy culture viability. The transfection setup is interfaced with compatible and rapid cell growth and productivity assessments.

In chapter six, the utility of the platform to undertake multiple gene co-transfection is considered. Specific stoichiometry of parts is required for multigene engineering to be studied and optimised. Single cell analysis, using flow cytometry, of the expression levels of three transfected fluorescent reporter genes is studied. The main strategies that can be used to introduce multiple genes into cells, either on separate plasmids or on singular multigene plasmid carrying all genes, are also compared.

The developed transfection platform is implemented in chapter seven. A proof-of-concept demonstration of its use to rapidly audition a library of effector genes to find a solution to improve production of a DTE-mAb, is shown, with gene dose effects evident. The screening platform is enhanced by the implementation of a deep-well plate, shaking culture format and the use of Design of Experiments (DoE) which can statistically model and reduce the number of experiments required to screen, infer and optimise effective combinatorial engineering strategies.

Conclusions of the processes and investigations undertaken are further discussed and summarised in chapter eight; along with a brief mention of other possible uses of the developed platform and suggestions for future work.

2. The Production and Development of Biopharmaceuticals

This chapter provides an overview of the production processes involved in the manufacture of recombinant protein biopharmaceuticals. The most commonly utilised expression system, using CHO cells, is described and particular focus is given to the 'upstream' bioprocesses involved in recombinant biopharmaceutical production. The key cellular processes that underpin recombinant protein production and secretion are described. Together with the following chapter, this information provides the context and rationale for the subsequent work included within this thesis.

2.1. Biopharmaceuticals

Biopharmaceuticals, also referred to as biologics, are medicinal, protein or nucleic acid, pharmaceutical products. They are biological in nature, generated in engineered biological systems and used in therapeutic, preventative or *in vivo* diagnostic applications to treat or cure diseases and injuries (Rader, 2008; Walsh, 2002). A major category of biopharmaceuticals, and of main interest here, are recombinant antibody therapeutics; others include blood-related proteins, hormones, enzymes and vaccines, and target conditions such as cancers, autoimmune, inflammatory and metabolic diseases (Walsh, 2014).

The advance of recombinant protein biopharmaceuticals stemmed from developments in the ability to genetically engineer the stable insertion of foreign DNA into mammalian cells (recombinant DNA technology) (Kretzmer, 2002) and followed success of the manufacture of recombinant insulin (Humulin), a peptide hormone, in microbial cells in 1982 (Johnson, 1983). Subsequently in 1986, the first recombinant protein therapeutic – Activase (tPA (tissue plasminogen activator)), used to dissolve blood clots, produced in genetically modified CHO

cells by Genentech – gained regulatory and market approval (Kretzmer, 2002; Wurm, 2004).

By mid-2014, the United States Food and Drug Administration (FDA) had licensed 246 biopharmaceutical products (Walsh, 2014). The industry is expanding, with 54 approvals between 2010 and 2014, and many more products in the development pipeline (Walsh, 2014). During 2013, global sales of biologics were US\$140 billion, highlighting the economic and medical significance of the biopharmaceutical and recombinant protein therapeutic market. Currently, the largest class of biologics are mAbs, contributing to over half of the 2013 annual sales and making up seven out of the ten top-selling drugs in terms of value that year (**Table 2.1**). They are the fastest-growing therapeutic type and there are many more mAb-based products at all stages of development including in multiple phases of clinical trials (Aggarwal, 2011).

2.1.1. mAbs as Therapeutic Proteins

Many of the top-selling recombinant proteins are mAbs, immuno-therapies for arthritis and cancer indications. Antibodies, also known as immunoglobulins, are glycosylated proteins (glycoproteins) naturally produced as the primary component of the adaptive immune response, protecting against foreign antigens and infectious components (Lipman et al., 2005; Schroeder and Cavacini, 2010). The basic structure of a standard IgG1 antibody protein comprises of two heavy chain (HC) and two light chain (LC) polypeptides held in a 'Y' shape by disulphide and non-covalent bonds (Lipman et al., 2005; Schroeder and Cavacini, 2010). These structures have two antigen-binding regions (Fabs) which are involved in epitope (foreign molecule) recognition. They also have an Fc domain that activates an immune response against the epitope-presenting molecule by triggering various mechanisms, including marking target cells for destruction by phagocytosis or cell lysis (Lipman et al., 2005; Schroeder and Cavacini, 2010). Unlike polyclonal antibodies, that recognise multiple epitopes (Lipman et al., 2005), mAbs have monovalent specificity, enabling precise, highly-specific targeting of discrete epitope structures.

Table 2.1: Top Ten Best-Selling Recombinant Biopharmaceuticals in 2013.

Information, taken from Walsh (2014), describing the top ten recombinant products sold in 2013, their primary therapeutic indication and the host cell they are produced in.

Rank	Product	Main Therapeutic Indication	Host Expression System	Sales (\$ billions)	Company
1	Humira (adalimumab; anti-TNF)	Rheumatoid arthritis	CHO	11	AbbVie
2	Enbrel (etanercept; anti-TNF)	Rheumatoid arthritis	CHO	8.76	Amgen, Pfizer, Takeda
3	Remicade (infliximab; anti-TNF)	Crohn's disease	Murine myeloma (Sp2/0)	8.37	Janssen Biotech
4	Lantus (insulin glargine)	Diabetes mellitus	<i>E. coli</i>	7.95	Sanofi
5	Rituxan/ Mabthera (rituximab; anti-CD20)	Non-Hodgkin's lymphoma	CHO	7.91	Biogen-IDEC, Roche
6	Avastin (bevacizumab; anti-VEGF)	Metastatic colorectal cancer	CHO	6.97	Roche/Genentech
7	Herceptin (anti-HER2)	Metastatic breast cancer	Murine cell line	6.91	Roche/Genentech
8	Neulasta (pegfilgrastim)	Chemotherapy-induced neutropenia	<i>E. coli</i>	4.39	Amgen
9	Lucentis (ranibizumab; anti VEGF)	Neovascular age related macular degeneration	<i>E. coli</i>	4.27	Roche/Genentech
10	Epogen (epoietin- α)	Anemia	CHO	3.35	Amgen

The first mAb therapeutics approved were completely murine in structure, developed from advances in hybridoma technology (Kohler and Milstein, 1975). Progressive developments created 'chimeric' partly human antibodies, and subsequently fully humanized structures before advancing, along with developments in phage display and transgenic mouse technologies, to human structures derived from human sequences. These resulting therapeutics had

lower immunogenicity, improved effector functionality and increased stability through prolonged serum half-lives (Johnston, 2007).

2.1.2. Novel Therapeutic Molecules

As time has progressed the nature of therapeutic proteins gaining regulatory approval has changed. There has been an increase in unnatural, engineered proteins, for example bispecific antibodies and Fc-fusion proteins (therapeutic proteins attached to an IgG Fc domain), which are commonly referred to as 'next-generation' molecules. As the first generation of biologics come 'off patent' there has been a trend towards 'biosimilar' and 'biobetter' therapeutics gaining approval. In 2010, it was reported that 50% of new biopharmaceuticals gaining approval were 'biosimilars' (Walsh, 2010) – antibody structures that have an identical amino acid sequence to a previously approved recombinant antibody, but are manufactured using divergent bioreactor processes and/or cell hosts (Beck, 2011). These alterations have the potential to change post-translational modifications (PTMs), which can affect product quality, safety and effectiveness (Beck, 2011). 'Biobetters' recognise the same epitope target to that of a product existing regulatory approval but have an altered protein structure to provide enhanced characteristics, for example an increased half-life (Beck, 2011). Novel protein formats and engineered therapeutics have heightened incompatibility with native cellular production hosts, which lack the machinery required to sufficiently fold and secrete these synthetic, DTE molecules. Accordingly, these therapeutics typically suffer from slow production timelines and/or poor product yields. The difficulties in producing next-generation therapeutics may stem from inadequate PTMs or protein processing, which can lead to protein aggregation and degradation and/or insolubility problems (Johari et al., 2015). Accordingly, novel cell factory and process engineering solutions are required in order to enable significant improvements in production yields and development times for DTE proteins.

2.2. Manufacture of Recombinant Protein Therapeutics

2.2.1. Expression Systems for Biopharmaceutical Manufacture

Multiple diverse expression systems are utilised for biopharmaceutical manufacture, primarily bacterial and mammalian, but also yeast and insect systems. The optimal production host for any specific protein depends on the recombinant products' characteristics. Host cell performance considerations, and relative positive and negative attributes of common expression systems, are summarised **Table 2.2**. The expression host chosen for biopharmaceutical manufacture must efficiently generate sufficient levels of correctly processed product in a consistent and uniform manner (Birch and Racher, 2006). Ability to grow in suspension culture in chemically defined animal component-free media, and protein expression stability over long-term culture are important criteria for regulatory success (Birch and Racher, 2006).

Prokaryotic systems are frequently utilised for recombinant protein production, with many successful examples, but as they lack the machinery to glycosylate nascent peptides, they are primarily limited to molecules whose therapeutic function is not depended on human PTMs. As prokaryotes process proteins to a lesser extent than eukaryotic systems, the types of molecules typically generated in these host are relatively simple structures, for example hormones and interferons (Ferrer-Miralles et al., 2009).

Many therapeutics, including mAbs, require precise folding and PTMs (particularly glycosylation and disulphide bond formation) for correct biological activity, efficacy, stability and immunogenicity (Ferrer-Miralles et al., 2009; Walsh and Jefferis, 2006). Mammalian systems are capable of generating human-like proteins in terms of molecular structure, biochemical properties and producing the necessary modifications; making them the primary expression system of choice for the production of complex therapeutic proteins (Sethuraman and Stadheim, 2006; Zhu, 2012).

Table 2.2: Different Expression Systems Utilised for Recombinant Protein Production

Information describing the advantages and disadvantages of various expression systems, taken from Andersen and Krummen, 2002; Demain and Vaishnav, 2009; Dumont et al., 2016; Ferrer-Miralles et al., 2009; Gerngross, 2004; Sethuraman and Stadheim, 2006; Walsh, 2010.

Expression Systems	Example(s)	Advantages	Disadvantages
Bacterial (prokaryotic)	- <i>Escherichia coli</i>	<ul style="list-style-type: none"> - Cheap, rapid, easy to culture - Easy to genetically manipulate - Many approved examples (proteins without human PTMs) - Well characterised - High yielding 	<ul style="list-style-type: none"> - Inadequate PTM ability - Inclusion bodies formation (re-folding required) - Produce endotoxins
Mammalian (eukaryotic)	<ul style="list-style-type: none"> - CHO - HEK-293 - BHK - NS0 - HEK93 - HT-1080 - PER.C6 	<ul style="list-style-type: none"> - Capable of human-like PTMs - Many approved examples - Easy to genetically manipulate - Well characterised - Low human pathogen susceptibility - Serum-free suspension culture 	<ul style="list-style-type: none"> - Long development times - Complex culture and nutrient requirements - Expensive
Yeast (eukaryotic)	<ul style="list-style-type: none"> - <i>Pichia pastoris</i> - <i>Saccharomyces cerevisiae</i> 	<ul style="list-style-type: none"> - Capable of PTMs including glycosylation - Cheap, easy to culture - High culture density 	<ul style="list-style-type: none"> - Undesirable, high mannose glycosylation (immunogenic) - Sub-optimal yields (compared to CHO)
Insect (eukaryotic)	<ul style="list-style-type: none"> - <i>Spodoptera frugiperda</i> - <i>Drosophila melanogaster</i> 	<ul style="list-style-type: none"> - Serum-free culture - Capable of some PTMs - Proteins secreted out of cells 	<ul style="list-style-type: none"> - Time-consuming process - Slow cultivation - Inadequate PTMs
Plant (eukaryotic)	<ul style="list-style-type: none"> - Transgenic tobacco - Oilseed rape - <i>Lemna minor</i> 	<ul style="list-style-type: none"> - Low human pathogen susceptibility - Cheap cultivation costs - Capable of some PTMs - Few successful examples - Storage organs (seeds) 	<ul style="list-style-type: none"> - Immunogenic PTMs - Poor public acceptance and commercial progression
Transgenic Animal (eukaryotic)	<ul style="list-style-type: none"> - Goat - Rabbit 	<ul style="list-style-type: none"> - Secretion of product - Approved examples 	<ul style="list-style-type: none"> - High risk of human pathogen susceptibility - High costs incurred - Glycosylation - potentially immunogenic

While mammalian systems are required to produce products with human-like PTMs, they have multiple disadvantages compared to other expression hosts, including other simpler eukaryotic cell-types (e.g. slow growth, complex media requirements). However, due to extensive research and development over the past twenty years, there have been several significant improvements in protein production processes utilising mammalian cell factories. For example, improvements in media and process design have enabled successful adaptation to suspension, serum-free growth (animal-component free), extended bioreactor production runs, and significant increases in maximum cell densities (Birch and Racher, 2006; Jayapal et al., 2007; Wurm, 2004).

2.2.2. CHO Cells

Stemming from the advantages of mammalian cells there are a range of expression systems which have been manipulated for recombinant protein production. These include immortalised CHO cells, murine cell lines and baby hamster kidney cells (Birch and Racher, 2006; Wurm, 2004). CHO cell lines have become the predominant production host (particularly for complex therapeutics), making up 35.5% of all approvals by 2014 and accounting for 68% of all mammalian-based approvals (Walsh, 2014).

CHO cells were isolated from the Chinese hamster (*Cricetulus griseus*) in the form of immortalised ovarian fibroblasts. The first cell line was established by Puck and colleagues in the late 1950s (Tjio and Puck, 1958), with early work, including mutagenesis studies, resulting in various auxotrophic cell lines. There is a wealth of different CHO cell lines with prominence today, many stemming from a 'grandfather' proline auxotroph that was isolated in the 1960s (CHOK1; Jayapal et al., 2007; Puck and Kao, 1967).

The dominance of CHO cell production systems stems from several beneficial attributes. CHO cells are able to be easily genetically manipulated; can synthesise satisfactorily human-like PTMs and can correctly fold and assemble glycosylated proteins that are bioactive, and have low immunogenicity in humans (Birch and Racher, 2006; Jayapal et al., 2007; Lai et al., 2013). They are long-standing safe production hosts, with the majority of human pathogens unable to

propagate within them (Berting et al., 2010; Jayapal et al., 2007; Lai et al., 2013). Additionally, higher eukaryotes, including CHO cells, contain functional secretion pathways, enabling cells to recognise signal sequence peptides on polypeptides and secrete a recombinant product out of the cell, easing downstream purification (Barnes et al., 2000). CHO cells are able to grow resiliently and rapidly to high cell densities in large scale bioreactor cultures. Successful adaptations and evolutions allow CHO cells to grow in animal-component-free, serum-free suspension culture, and in chemically defined media – important characteristics for efficient growth as well as regulatory approval (Jayapal et al., 2007; Lai et al., 2013). The implementation of the last >20 years of intensive CHO cell research, development and optimisation have greatly improved the production titre by over 100-fold, to over 10 g L⁻¹, and established CHO as a mammalian model organism and the bioprocess industrial workhorse (Kim et al., 2012; Lai et al., 2013).

Given their other relative advantages compared to eukaryotic hosts (see **Table 2.2**) there is significant interest in engineering microbial hosts, that are able to produce human-like proteins. Developments have included introducing an *N*-glycosylation pathway from *C. jejuni* into *E. coli* and alterations in the glycan composition and improvements in the efficiency at which glycosylation reactions take place (Jaffé et al., 2014). Other developments have been with the use of immortal human cell lines (e.g. HEK293, HT1080 and PER.C6). They are able to synthesise, very desirably, natural-like proteins, with accurate structures and position of PTMs, particularly glycosylation patterning (Dumont et al., 2016). These systems show promise but the advantages of the CHO cell system currently outweigh these, therefore CHO cells will remain the preferred host for production of human-like proteins for the foreseeable future (Bandaranayake and Almo, 2014).

Many different CHO cell derivatives and cell lines have developed and evolved over the numerous years of CHO cell exploitation, resulting in great diversity in the genotypes and phenotype being used (Wurm, 2013). The majority of biopharmaceutical companies have each established and optimised their own, in-house, production CHO host cell line providing improved standardisation, predictability and reduced cell line development timescales. For regulatory, therapeutic and commercial success it is advantageous for a production cell line

to grow quickly to high densities within the bioprocessing setting and have a high specific productivity with consistent and the required product quality (Birch and Racher, 2006).

2.3. Biopharmaceutical Production

A typical manufacturing process for recombinant protein production in CHO cells is outlined in **Figure 2.1**. Firstly, the gene or genes, which encode the therapeutic protein, along with a gene encoding a selection marker, are included on an expression plasmid which is then introduced into a production CHO cell host. Culture, in media containing the selective pressure of a specific chemical inhibitor that matches the introduced selection marker, follows to isolate both cells which have successfully taken up the recombinant plasmid and establish stable expression. Slow screening processes follow, to identify and isolate an individual cell (a clone) which has highly desirable cell culture and manufacturing performance. This clonal cell line is expanded and used in large-scale, fed-batch cultures to generate large amounts of the recombinant product. Lastly, there are multiple downstream processes used to isolate, purify and concentrate the recombinant protein so that it is suitable as a medicinal product that meets regulatory standards (Birch and Racher, 2006; Jayapal et al., 2007; Wurm, 2004). Many developments and advances have been made throughout this production pipeline, some of these, along with a more in-depth description of a typical production process using CHO cells, follow.

2.3.1. Expression Plasmid

Expression plasmids (also known as expression vectors) are the principal and a well-established method of delivering the product genes of interest into the host CHO cell line. Coding sequences within the plasmids are routinely enhanced for CHO cell expression through codon optimisation and the elimination of potentially problematic sequences such as cryptic splice sites (Birch and Racher, 2006). Present on the plasmids are various elements which aid successful transcription, translation and secretion of the recombinant genes.

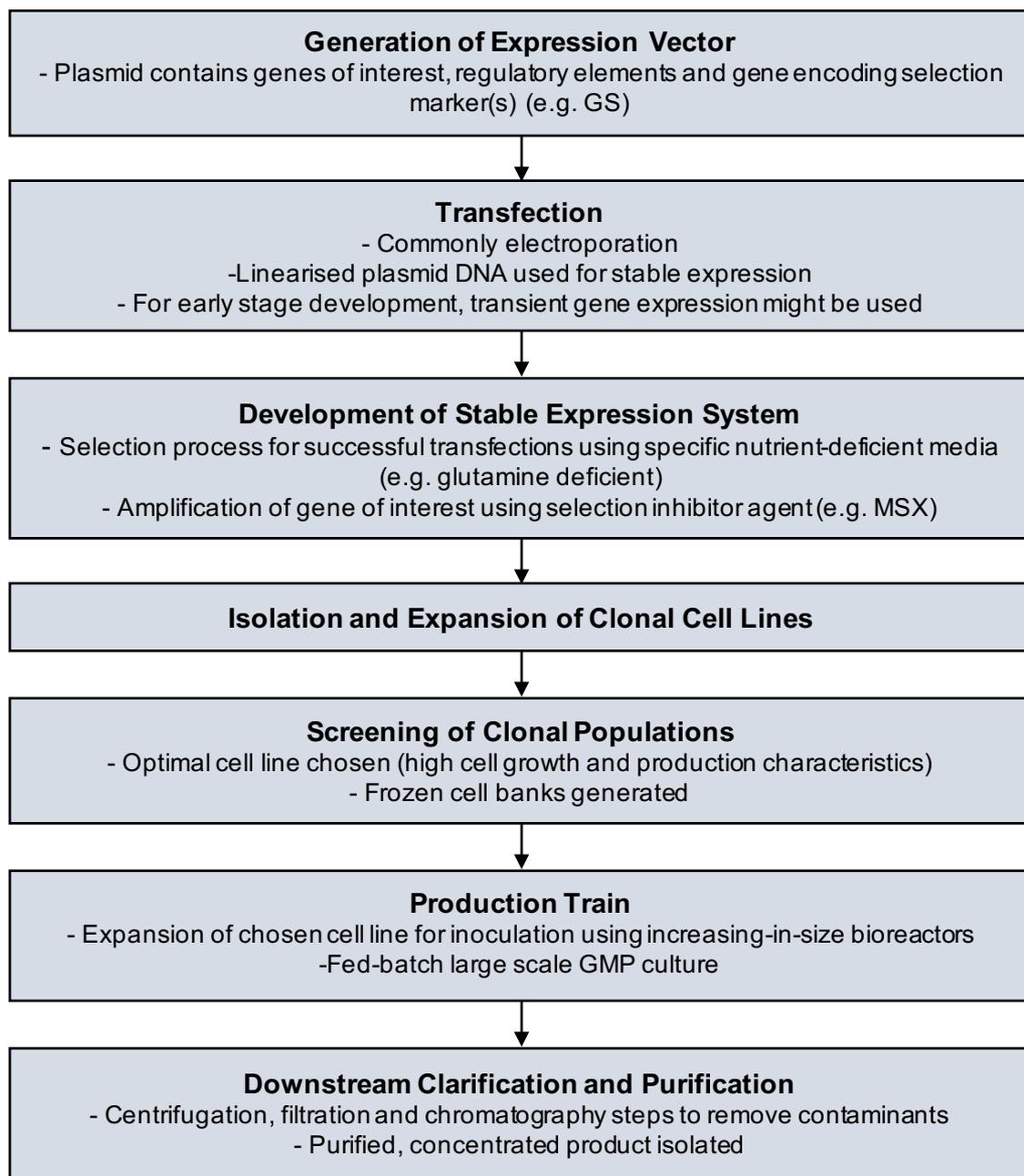


Figure 2.1: Outline of the Typical Biopharmaceutical Production Processes Undertaken During Recombinant Protein Manufacture.

Information taken from Birch and Racher, 2006; Jayapal et al., 2007; Wurm, 2004. Abbreviations used in the figure: glutamine synthetase (GS), methionine sulfoximine (MSX) and good manufacturing practice (GMP).

Traditionally, primarily due to its availability, a strong viral promoter is employed (hCMV-IE1), generating high levels of expression. Recently there have been promising developments in synthetic promoters, engineered to provide modular, predictive expression with increased consistency and stability. They are a very

promising alternative to standard cytomegalovirus (CMV) based promoters to increase control and flexibility in expression levels (Brown et al., 2014). RNA elements, such as polyA tails and Kozak sequences, are included to help regulate and improve product mRNA translation, stability and export (Rita Costa et al., 2010). Inclusion of signal peptide sequences help facilitate protein export through the secretory pathway and out of the cell into the supernatant, this makes the product easier to isolate, an aid to product harvesting.

Another important plasmid feature is the presence of a selection marker gene, often a gene encoding an enzyme which is essential when cells are grown in a specific nutrient-deficient media (Wurm, 2004). They are included to enable isolation of successful transfectants - cells where the recombinant gene has integrated into an active part of the chromosome (heterochromatin). The most routinely used selection markers are dihydrofolate reductase (DHFR) and glutamine synthetase (GS); both essential to cells grown in media lacking in glycine/hypoxanthine/thymidine or in glutamine, respectively (Birch and Racher, 2006; Lai et al., 2013). Due to the lack of important metabolite(s), only cells expressing the selection marker gene, and therefore the genetically linked recombinant product gene(s), will remain viable, proliferate and importantly be producing the recombinant product. To select for amplification of the recombinant genes and enhanced production capabilities, the selection marker gene is often linked to a weak promoter and specific inhibitors of the enzyme are included, and subsequently elevated, in the culture media. For the DHFR and GS systems, methotrexate (MTX) and methionine sulphoximine (MSX) inhibitors are used respectively (Birch and Racher, 2006; Lai et al., 2013). Recently, to further increase the effectiveness of the GS selection system, a GS knockout cell line has been engineered (Fan et al., 2012). It contains no endogenous GS genes, therefore is unable to make any essential glutamine, unlike previous CHO cell lines. This results in a more potent selection system tool, improving the generation of high yielding cell lines, with a two to three fold increase in productivity documented (Fan et al., 2012).

Targeted site gene integration, which allows the recombinant selection marker and product genes to be integrated into a highly active “hotspot” region of heterochromatin, resulting in higher and more predictable expression, is another

recent development which begins with the expression plasmid. Approaches to target integration include the recombinases (e.g. Cre/Lox), endonucleases (e.g. zinc finger nucleases) and artificial chromosome expression technology (Kim et al., 2012; Lalonde and Durocher, 2017). Specific flanking sequences or elements, for example identical sequences to those present in the host cell, need to be incorporated into the expression plasmid for targeted integration. A targeted integration approach requires knowledge of a highly expressing, stable site in the host genome to achieve success (Lalonde and Durocher, 2017). Additionally, higher expression has been achieved through the use of regulatory elements such as ubiquitously acting chromatin opening elements (UCOEs); which function to keep heterochromatin 'open' in an active transcriptional state, increasing levels of recombinant gene expression (Lalonde and Durocher, 2017).

2.3.2. Transfection Approach

Transfection is the delivery of foreign nucleic acid (in this case the DNA expression plasmid just described) into cells. Frequently transfection is used as an analytical tool to study the function of genes or their products, to increase or decrease expression of specific genes, or of most relevance here and in the bioprocessing industry, used to generate therapeutic protein (e.g. mAb) producing recombinant cell lines (Kim and Eberwine, 2010; Wurm, 2004). The fundamental objectives of the transfection approach are to reproducibly achieve a high transfection efficiency and a high level of recombinant gene expression with minimal cytotoxicity and minimal effect on cell physiology (Kim and Eberwine, 2010). The transfection process can be manipulated to result in stable (continual) gene expression, very common for large-scale recombinant therapeutic GMP production, or transient (temporary) gene expression, a method which is used with increasing frequency and enhanced capability to supply recombinant therapeutic protein for early stage development (Hu et al., 2012; Rita Costa et al., 2010).

There are a plethora of applicable transfection techniques categorised into biological, chemical and physical based technologies; with varying cost, resource requirements, efficacy, ease and licensing (Rita Costa et al., 2010). Some

advantages, disadvantages and examples within each of the three categories are summarised in **Table 2.3**.

Biological transfection approaches derive from nature, primarily being virus-mediated. They are highly efficient and easy methods of introducing DNA into a foreign host. Concerns in terms of safety from potentially immunogenic or cytotoxic effects, along with plasmid size limitations make them a rare choice for biopharmaceutical manufacture (Douglas, 2008; Kim and Eberwine, 2010).

Chemical methods are based upon a variety of positively charged chemicals which all form positively charged complexes with negatively charged nucleic acid molecules. The complexes create electrostatic interactions with a cell's negatively charged plasma membrane before entering the cell through endocytosis (Kim and Eberwine, 2010; Rita Costa et al., 2010). The main chemical-based transfection methods are based on cationic lipids (e.g. Lipofectamine[®]), cationic polymers (e.g. polyethylenimine (PEI)) and calcium phosphate (DNA-calcium phosphate co-precipitation). Overall, chemical methods require significant optimisation (e.g. ratio of DNA to chemical, to balance toxicity and efficacy), can usually be used in a regulatory-friendly serum-free media and are capable of high transfection efficiency (Rita Costa et al., 2010). The commercially available lipid-based method Lipofectamine[®] is able to yield high levels of product following optimisation, but a major downside comes with an inability to scale the method due to the significant costs (Geisse and Fux, 2009). A polymer based method PEI is a common chemical-based transfection approach and is particularly popular for transient expression systems. It is relatively low cost and, importantly, a scalable method. However, PEI is not as efficient as Lipofectamine[®] at delivering DNA to the nucleus where DNA is transcribed. Also, PEI commonly encounters poor transfection efficiency and harmful cytotoxicity, although this can be improved through optimisation, particularly of the amount and ratio of DNA and PEI polymer used. There can also be inherent variability in the PEI polymer itself (Cohen et al., 2009; Geisse, 2009; Thompson et al., 2012; Zhu, 2012)

Table 2.3: Common Transfection Approaches.

Chemical and physical approaches are most suitable for biopharmaceutical manufacture. Based on table from Kim and Eberwine (2010).

Approach	Method	Advantages	Disadvantages	Examples	References
Biological	Viral-Mediated	<ul style="list-style-type: none"> - Efficient and easy to use - Common, long-standing, well-developed method 	<ul style="list-style-type: none"> - Potential biohazard - Limited gene-length carrying capability - Potentially immunogenic and cytotoxic 	<ul style="list-style-type: none"> - Adenovirus - Alpha virus (Sindbis, Semliki Forest virus) - Vaccinia virus 	Kim and Eberwine, 2010; Wurm and Bernard, 1999
Chemical	Polymer-mediated	<ul style="list-style-type: none"> - Efficient and robust - No gene carrying length limitation - Many examples commercially available 	<ul style="list-style-type: none"> - Toxicity problems - Variable transfection efficiency obtained from changes in cell type/condition - Optimisation of DNA to chemical ratio needed - Variable costs – cationic polymers relatively inexpensive, unlike expensive cationic lipids - Calcium phosphate requires serum in the media 	<ul style="list-style-type: none"> - PEI - DEAE-dextran - Dendrimer 	Jordan and Wurm, 2004; Kim and Eberwine, 2010; Rita Costa et al., 2010
	Cationic Lipid			<ul style="list-style-type: none"> - Lipofectamine® 2000, LTX and 3000 - DOTMA - DOTAP 	
	Co-precipitation			<ul style="list-style-type: none"> - Calcium phosphate - Calcium chloride 	
Physical	Electroporation	<ul style="list-style-type: none"> - Robust, consistency achievable - Can be highly efficient and have low toxicity - Relatively simple methodology, suitable for scale-up 	<ul style="list-style-type: none"> - Specialised instruments, devices and techniques required - Often expensive 	<ul style="list-style-type: none"> - Amaxa Nucleofector™ - Maxcyte® 	Kim and Eberwine, 2010; Lin-hong et al., 2002
	Direct Injection			<ul style="list-style-type: none"> - Microneedle - Atomic force microscope tip 	
	Biolistic Particle Delivery			<ul style="list-style-type: none"> - Gene Gun 	

Physical based approaches are frequently mechanically based (e.g. direct micro-injection and particle bombardment) which can be very costly, destructive, have poor-throughput and require a lot of time-consuming skill (Kim and Eberwine, 2010; Mehier-Humbert and Guy, 2005). However, the most commonly utilised physical approach, and perhaps even transfection method, is electroporation. Electroporation involves short, intense electrical pulses which alter the membrane potential across the plasma membrane, producing temporary permeable pores allowing nucleic acid molecules to enter (Mehier-Humbert and Guy, 2005; Neumann et al., 1982; Rita Costa et al., 2010). It does require specialist, expensive equipment and can negatively affect cell viability but the method is relatively straight-forward, quick, very efficient, consistent, can transport DNA directly to the nucleus and does not require any additional 'carrier' molecules (Fratantoni et al., 2003; Kim and Eberwine, 2010). It is a commonly used approach by the biopharmaceutical industry for the development of stable cell lines.

2.3.3. Stable Versus Transient Expression

Stable transfection is where the recombinant gene(s) are continuously expressed and maintained. For successful stable transfection, the transfected DNA must be integrated into a transcriptionally active part of a cell's genome so that it is subsequently replicated and maintained during cell division (Hu et al., 2012; Rita Costa et al., 2010). The plasmid DNA is usually linearised prior to transfection to aid in successful integration (Wurm, 2004). Stable transfection is a lengthy process with integration events being a rare occurrence. The use of a genetically-linked selection marker on the plasmid, and supplementation with a suitable selective agent following transfection, allows more efficient isolation of stable transfectants that have integrated and subsequently maintained the expression plasmid. The screening of these transfectants is time- and resource-intensive, but it is capable of resulting in a high-yielding cell line with consistent product quality, when combined with a cloning process (Rita Costa et al., 2010).

Transient transfection is temporary and short-lived, with a limited period of expression. The approach is faster than stable transfection (weeks rather than months), but currently yields lower amounts of product. Unlike stable expression,

no selection or screening steps are required. Expression is extrachromosomal, with no chromosome integration, and the transfected DNA is not replicated, so is lost during culture through cell division and environmental factors (Baldi et al., 2007; Hu et al., 2012; Rita Costa et al., 2010). Traditionally, transient gene expression has been utilised extensively for small-scale research and analytical purposes. But it is increasingly being used to supply recombinant therapeutic protein for early stage development. It is a useful approach for generating material rapidly, for example for drug discovery, to explore the ability of a product to be expressed; for pre-clinical studies such as toxicology tests, or to investigate the effect of a process on product quality (Baldi et al., 2007; Rita Costa et al., 2010). Utilising transient gene expression, to generate milligram per litre amounts of product in less than 5 weeks, can reduce risks present in the drug development pipeline and potentially increase a drug's speed to market (Cain et al., 2013). The transfection efficiency, the proportion of the population taking up and expressing the introduced DNA (Hu et al., 2012), of a transient transfection process is important. The proportion of cells expressing the recombinant DNA can be improved through introducing more copies of the plasmid DNA. Caution and consideration is needed though as the process of introducing DNA can be harmful, so a balance between cell viability and transfection efficiency must be sought (Hu et al., 2012; Rita Costa et al., 2010).

There have been many developments and improvements in transient transfection processes, which have been reflected in the increasing yields obtained from these processes. Many of the substantial developments stem from PEI-based transfection. PEI has become a dominant approach, well-proven to quickly and cost-effectively produce large amounts of product. It is a relatively cheap and, importantly, scalable method (Daramola et al., 2014; Zhu, 2012). Transient transfection improvements incorporate addition of components to the culture media – including dimethyl sulfoxide (DMSO), lithium acetate, sodium chloride and valproic acid (VPA), along with an alteration to a mild hypothermic culture temperature (Wulhfard et al., 2010; Ye et al., 2009; Zhang et al., 2010). Altering the transfection parameters, including increasing the density of cells being transfected and reducing the amount of cytotoxic PEI being used, improve culture viability and, along with the use of a mild hypothermic condition, have

substantially improved productivity capabilities from transient transfections (Rajendra et al., 2011). Further advances have also been made through genetically engineering the host cell. Successful examples include overexpression of CELO adenoviral Gam 1, increasing transcription activation (Hacker et al., 2005); overexpression of Bcl-x_L, reducing cell death (apoptosis) (Majors et al., 2008b); and overexpression of X-box binding protein 1 (XBP1) and endoplasmic reticulum oxidoreductase 1 α (ERO1-L α), increasing secretion capacity, resulting in high yields of 875 mg L⁻¹ (Cain et al., 2013). The CHO cell host has been engineered with viral elements to result in enhanced and prolonged retention and replication of transfected plasmid DNA (Daramola et al., 2014; Durocher and Loignon, 2011; Kunaparaju et al., 2005). Viral elements, for example Epstein Barr-virus nuclear antigen-1 (EBNA-1) have been linked with nuclear import of plasmid DNA, which increases the amount of plasmid capable of being transcribed, along with enhancing transcription itself (Geisse, 2009). One of the most successful examples is by Daramola *et al* (2014). The CHO cell host was engineered to express both EBNA-1 and GS genetic elements, transfected plasmids contained a viral OriP DNA element and, along with a scalable, optimised transient transfection methodology, and a mild hypothermic condition, created high yields of up to 2 g L⁻¹ (Daramola et al., 2014).

Electroporation commonly is the transfection method of choice for recombinant protein manufacture from stable cell lines; it successfully and efficiently introduces DNA reproducibly to small volumes of culture (millilitres) which can then be scaled up through a selection and amplification process (Zhu, 2012). For greater transient expression scope and success, there is a need for rapid, relatively large-scale, high efficiency transient production methods. Chemical based methods, such as PEI, approach to meeting these criteria despite their drawbacks of poorer efficiency and batch to batch inconsistency. However scalable, more-efficient electroporation based transient transfection is under development. MaxCyte technology, utilising a continuous flow electroporation approach, is able to transfect a range of cell hosts and culture volumes (up to 1 L containing 2x10⁷ cells) and with optimisation can rapidly and reproducibly generate beneficial titres of >1 g L⁻¹ (Fratantoni et al., 2003; Steger et al., 2015). For recombinant protein production development and pipelines, flow

electroporation is a clear alternative transfection method to PEI as a CHO transient expression approach.

2.3.4. Cloning Process

Transfection generates a varied population of cells, for example, in terms of their growth, productivity and stability, features which would be unpredictable and create product quality inconsistencies if used for recombinant protein manufacture (Priola et al., 2016). Primarily from a regulatory, but also from a manufacturing perspective, a cloning process is required following the stable transfection and selection process, to minimise heterogeneity. Isolating cell lines from single cells and assessing their characteristics allows for development of a favourable and homogenous cell line. A high specific productivity (the ability for a cell to produce and secrete product) is of fundamental importance, but subsequently clones are assessed for other important manufacturing characteristics. These include: robust high growth rates; successful adaption to serum-free, suspension culture; and stable, consistent production and product quality over a defined number of generations (Li et al., 2010). Achieving a desirable and appropriate combination of these attributes is rare, but screening clones in a swift and thorough manner helps to increase the likelihood, and reduce the timescales to success (Agrawal et al., 2013).

In the past, the cloning process involved limiting dilution, where single cells were diluted down into separate wells of microtitre plates, which were all assessed for titre production (Agrawal et al., 2013; Priola et al., 2016). This approach was slow, costly and labour-intensive and good clones could be missed due to insufficient screening throughput. Developments in increased throughput and automation have improved screening processes. Common approaches now include, for example, fluorescence activated cell sorting (FACS) – where the levels of recombinant product are linked to a fluorescent based signal; millions of cells can be assessed and sorted based upon their productivity in minutes (Agrawal et al., 2013; Kim et al., 2012; Priola et al., 2016). Other automated and robotic systems, e.g. ClonePix and CellSelector, increase throughput and reduce timescales and labour intensity dramatically; thousands of clones can be narrowed down faster and undergo more in-depth screening for suitability for large scale recombinant

protein manufacture and stability. After an optimal manufacturing, clonal cell line has been isolated, master and working cell banks are generated, ready for good manufacturing practice (GMP) production.

2.3.5. Upstream Manufacturing Process

The actual manufacturing process to generate the required, often large, amounts of recombinant product begins with a biomass expansion process of the frozen, clonal working cell bank (WCB). A seed train, comprised of rocker bag or stirred tank bioreactors, increasing in size in a step-wise manner, is used to generate sufficient biomass to inoculate the large production bioreactors (up to 20,000 L in size) (Birch and Racher, 2006; Hu et al., 2012; Wurm, 2004). Disposable systems, e.g. plastic culture bags, are increasing in use during this process; with advantages stemming from a reduced need for cleaning and sterilising processes, which lowers costs and timescales (Butler and Meneses-Acosta, 2012; Hu et al., 2012). Process consistency and robustness can be influenced by the many parameters controlled within the bioreactor manufacturing process, including pH, temperature and dissolved oxygen levels (Birch and Racher, 2006). Fed-batch mode, where bioreactors are supplemented with concentrated feeds at particular points during culture, is the most commonly utilised culture approach for recombinant protein manufacture. Nutrients that would otherwise become exhausted can be added which brings benefits including an increase in the maximum viable cell density achieved and an extension to culture viability and the duration of production (Birch and Racher, 2006).

Significant improvements in the production of recombinant protein manufacture and the processes have been made. In 2004 increases in yield of 100-fold over a 20-year duration were reported, with titres of 50 mg L⁻¹ in 1986 improved to 4.7 g L⁻¹ in 2004. This was attributed to both significant increases in cell densities and production durations and increased specific productivities (from <10 to ~90 pg cell⁻¹ day⁻¹) (Wurm, 2004). The enhancements in titre, through improvements in specific productivity, cell growth and culture viability, are primarily from successive progressions in media, feeding and bioprocess developments, some of these are highlighted subsequently in **Section 3.3**.

2.3.6. Downstream Processes

In order to isolate and purify the recombinant therapeutic protein from the cell culture process, several downstream processing steps are undertaken (Shukla and Thömmes, 2010). They are not of importance here, but a brief description completes the manufacturing process. It is of crucial importance, for the success and approval of a product, that it is free from impurities and contaminants. In general, the cell culture volume is harvested by centrifugation, and filtration used to remove the majority of cells and cellular debris. A high specificity protein A affinity chromatography step follows, this removes impurities such as host cell proteins and DNA, and results in 98% purity (Birch and Racher, 2006; Shukla and Thömmes, 2010). The product is eluted using a low pH buffer and further polishing steps of anion and cation exchange further reduce levels of impurities (aggregates, host cell proteins, DNA). Clearance of viral material is achieved using viral inactivation and size-based filtration. Lastly, an ultrafiltration or diafiltration step concentrates the final drug product in an appropriate formulation buffer (Birch and Racher, 2006; Shukla and Thömmes, 2010).

2.4. Cellular Level Processes in Mammalian Recombinant Protein Production

The introduced foreign gene sequences undergo a plethora of complex, sequential cellular processes in order for the host cell to produce a modified and fully folded recombinant protein, which is secreted out of the cell. In mammalian cells, less complex proteins are able to fold fairly simply within the cell cytoplasm; more complex proteins, like recombinant proteins such as mAbs, made up of multiple polypeptide chains with extensive PTMs, require the distinct folding environment and processing components provided within the endoplasmic reticulum (ER) cellular compartment (Braakman and Bulleid, 2011; Lodish et al., 2008).

The coding sequence of the transfected recombinant gene is first transcribed, converting the DNA sequence into a single stranded RNA sequence, then processed into an mRNA transcript which relocates to the cytoplasm. Translation of this mRNA transcript, by ribosomal machinery, converts the genetic sequence

into a linear polypeptide chain of amino acids. Proteins destined to be secreted from the cell, contain a short signal sequence at their N-terminus. This sequence is recognised by a signal recognition particle (SRP), which directs the polypeptide to the ER for processing, via binding to the SRP receptor on the ER membrane. The growing polypeptide chain is co-translationally translocated into the ER, enabling protein folding to occur within the ER component, in-parallel with translation, as opposed to in the undesirable cytoplasm. Once within the ER, polypeptide chain(s) are able to mature, with enzymatic modifications, and assemble into their correct three-dimensional conformation; with quality control systems present to reduce the likelihood of incorrectly folded proteins being secreted from the cell. The mature proteins are concentrated at the ER membrane and transported in vesicles to the Golgi complex. Proteins are processed within the ER and the Golgi complex with PTMs, including glycosylation modifications, essential processing for correct protein function, folding and half-life. The secretory pathway finishes with secretory vesicles transferring the recombinant product through the plasma membrane and out of the cell (Braakman and Bulleid, 2011; Lodish et al., 2008).

2.4.1. ER Protein Folding and Assembly

The main cellular processes undertaken in the ER, to fold and process polypeptides are summarised in **Figure 2.2**. Further details are described in the following sections.

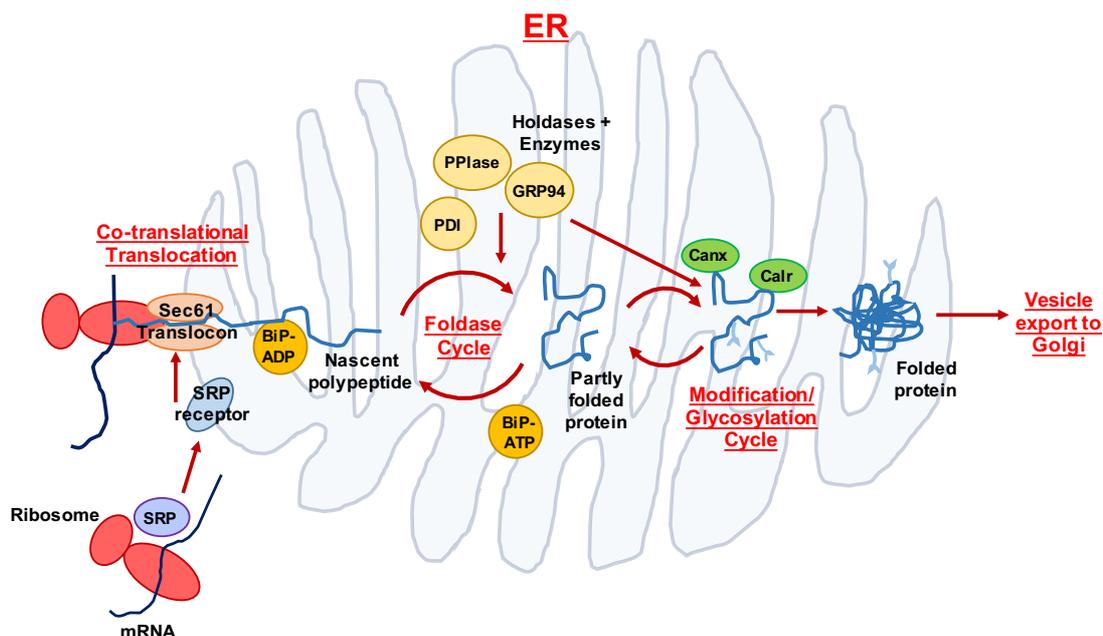


Figure 2.2: An Outline of the Processes Undertaken to Assemble Proteins in the ER

A nascent polypeptide that has begun to be translated is recognised by a signal recognition particle (SRP) component in the cytoplasm, targeting the complex for co-translational translocation into the ER. When in the ER lumen, it associates with ER chaperone binding immunoglobulin protein (BiP) as it continues to be translated, and begins to be modified and folded. Other resident ER components, many of them enzymes such as protein disulphide isomerases (PDIs) and peptidyl-prolyl *cis-trans* isomerases (PPIases), play important roles in correctly folding and modifying the developing protein. When partly folded, the polypeptide chain is usually glycosylated with the addition of specific glycans, which are essential for specific protein function and activity. Only correctly folded proteins are exported via vesicles and progress along the secretory pathway to the Golgi complex. The figure adapted from Nyathi et al., 2013 and Schröder, 2008.

2.4.1.1. Polypeptide Translocation into the ER Lumen

The progression of a secretory protein through the ER begins with its transport into the lumen of the ER compartment. This is undertaken in parallel while the nascent polypeptide is still being translated, referred to as co-translation translocation (Nishimiya, 2014; Nyathi et al., 2013). The newly-formed hydrophobic signal sequence directs the polypeptide-ribosome complex to the ER, through binding to an SRP, a six-protein complex; temporarily pausing translation while the translocation at the ER membrane is established. The SRP - nascent polypeptide - ribosome complex binds to the SRP receptor on the ER membrane, which then associates with Sec61, a translocon port, spanning the

ER membrane (Chakrabarti et al., 2011; Nyathi et al., 2013). The SRP and receptor separate and the polypeptide chain continues to be translated and enters and extends through the translocon into the lumen of the ER. Interaction with ER components heat shock proteins (Hsp), Hsp70 and Hsp40 (Binding immunoglobulin protein (BiP) and Sec63 respectively), play a part in this translocation and translation resuming process (Alder et al., 2005; Lang et al., 2012; Nyathi et al., 2013). One of the first stages of protein processing in the ER lumen is the removal and degradation of the signal sequence. Upon translocation and elongation completion, the ribosome disassociates, the translocon closes off and further protein modification and folding within the ER follows.

2.4.1.2. Polypeptide Modification and Folding in the ER Lumen

Once within the lumen of the ER, polypeptide chains undergo several PTM processes and fold and assemble into the specific three-dimensional conformation of the protein. For a typical mAb, the post-translational processes include several enzymatic steps, including disulphide bond formation, prolyl isomerisation and glycosylation (Braakman and Bulleid, 2011). In **Table 2.4** which follows, some of the many components and enzymes residing in the ER are summarised. Their primary functions are associated with correctly folding and processing polypeptide chains present in the ER, either directly through association with the polypeptide itself, or indirectly through association with other ER-residing components.

Table 2.4: Resident ER Components

Some of the key resident ER components which function in protein folding, modification and assembly. The information is taken from Braakman and Bulleid, 2011; Ellgaard and Helenius, 2003; Schröder, 2008.

ER Functional Component(s)	Protein Family	Function(s)
Chaperones and Co-Chaperones		
BiP	Hsp70	<ul style="list-style-type: none"> - Broad specificity chaperone - ATPase - Protein folding: interacts with hydrophobic regions on polypeptides – stabilising folding intermediates, preventing aggregation - Targets proteins for ER associated degradation (ERAD) - Gates translocon apparatus for polypeptide entry into the ER - Holdase function – e.g. holds mAb HC until it is displaced by LC binding - Unfolded protein response (UPR) regulation
Grp94	Hsp90	<ul style="list-style-type: none"> - Chaperone - Holdase and foldase activity - Necessary for particular protein secretion and ERAD
ERDJ1 - 6	Hsp40	<ul style="list-style-type: none"> - Co-chaperones - Work in association with other chaperones (particularly BiP) - Associated with protein folding and ERAD
Grp170	Hsp110	<ul style="list-style-type: none"> - Nucleotide exchange factor for BiP
Peptidyl Prolyl <i>cis/trans</i> Isomerases (PPlases)		
CypB	Cyclophilin B	<ul style="list-style-type: none"> - Catalyse <i>cis/trans</i> isomerisation of peptidyl-prolyl bonds
FKBP 2, 7, 9, 10, 11, 14	FKBP	<ul style="list-style-type: none"> - Catalyse <i>cis/trans</i> isomerisation of peptidyl-prolyl bonds - Lots of redundancy so precise functions unknown
Oxidoreductases		
PDI, ERp57, ERP72, ERp46 (plus many others)	PDI	<ul style="list-style-type: none"> - Catalyse disulphide bond formation - Chaperone/co-chaperone function: stabilises intermediates of multimeric complexes - Contribute to ERAD and protecting against ER stress
ERO1α, ERO1β	Sulfhydryl oxidase	<ul style="list-style-type: none"> - Oxidase function: re-oxidise PDI - Contribute to regulating redox conditions within the ER
Glycan-binding Proteins		
Calreticulin, calnexin	Calnexin family	<ul style="list-style-type: none"> - Lectin chaperones (binds to carbohydrates) - Bind to monoglucosylated N-linked oligosaccharides and assist in their folding and ER retention - Holdase function
UGGT	Glucosyl-transferase	<ul style="list-style-type: none"> - Plays a role in correct N-linked glycosylation – undertakes re-glucosylation of incorrectly folded proteins
ERMan	Mannosidase	<ul style="list-style-type: none"> - Identifies misfolded/struggling proteins, removes them from glycosylation cycle, through mannose sugar hydrolysis, and targets them for ERAD

Polypeptide chains begin in the ER in a reduced form. The formation of covalent disulphide bonds, between cysteine residues on the polypeptide chain, contribute to stabilising protein structures, preventing them from re-forming into their undesirable reduced states. They also play an important role in the assembly of multimeric proteins through the formation of disulphide bonds between separate polypeptide chains. The oxidation process of disulphide bond formation, one of the key processes in protein folding, is undertaken by a family of protein disulphide isomerases (PDIs; oxidoreductases) present within the ER lumen (Braakman and Bulleid, 2011; Ellgaard and Ruddock, 2005). In conjunction, the re-oxidation process, to restore the enzymatic capabilities of the PDI enzymes, is achieved with other ER resident components - ER oxidoreductins (e.g. ERO1). This important re-oxidation process is aided by the oxidising environment of the ER lumen and the presence of glutathione (Appenzeller-Herzog et al., 2010; Braakman and Bulleid, 2011; Frand and Kaiser, 1998; Pollard et al., 1998).

Prolyl isomerisation, where a peptide bond of proline residue with any other amino acid transitions fairly equally between the *trans* and *cis* form of the prolyl bond, is a naturally slow process. Specific prolyl bond orientation, and not an equilibrium mixture of the two, is needed for proper protein folding and functionality. Without enzymatic acceleration, the process would be rate-limiting and problematic. There are several peptidyl-prolyl *cis/trans* isomerases (PPIases) in the ER lumen which catalyse the critical process of conversion of *cis-trans* prolyl peptide bonds (Braakman and Bulleid, 2011; Schmidpeter and Schmid, 2015; Wedemeyer et al., 2002).

The covalent attachment of oligosaccharide molecules (also known as sugars or glycans) to residues on the polypeptide chain, is known as glycosylation (Pandhal and Wright, 2010). The pattern of glycosylation is of great importance for correct protein folding, secretion, function, activity, immunogenicity and half-life, and of relevance here due to the prevalence of glycosylated recombinant therapeutic proteins in development (Sethuraman and Stadheim, 2006; Varki, 1993). Within the ER, proteins primarily undergo *N*-linked glycosylation - additions of large hydrophilic sugar moieties to a nitrogen on the side chain of an asparagine residue (Aebi, 2013; Braakman and Bulleid, 2011; Pandhal and Wright, 2010). The addition and processing of glycans takes place in a sequential manner within

the ER lumen, with some anchoring of components to the ER membrane. For most human proteins, there is a core oligosaccharide moiety initially added (Glc₃Man₉GluNAc₂) to an Asn-X-Ser/Thr element on the polypeptide (Chakrabarti et al., 2011; Sethuraman and Stadheim, 2006). It is then processed through trimming by glycoside hydrolases (I and II), which remove individual glucose molecules from the oligosaccharide; before a mannosidase removes the terminal mannose molecule which targets the protein for removal from the ER. Further processing subsequently takes place in the Golgi complex (Pandhal and Wright, 2010; Schröder, 2008). After the removal of the first two glucose sugars, two ER lectin chaperones, calreticulin and calnexin (Calr and Canx), attach to the monoglucosylated glycoprotein. Their roles are associated holdase and foldase functions, preventing aggregation and degradation. They are released upon removal of the third glucose molecule, by glycoside hydrolase II. This trimming process plays a valuable role in ER quality control, in the tracking of correctly folded proteins and designating problematic molecules for degradation. Glucose molecules can be enzymatically added and removed, in co-ordination with Calr and Canx binding and release, in a cyclic manner, to ensure protein folding accuracy and quality (Ellgaard and Frickel, 2003; Schröder, 2008). Once in the correct conformation the protein begins transfer to the Golgi complex for further processing; alternatively, it is diverted from Calr and Canx cycle and targeted for degradation (Ellgaard and Helenius, 2003).

2.4.2. ER Quality Control Mechanisms

The processes being undertaken within the ER are essential for correct protein folding and subsequent functioning, but they are ultimately fundamental for cell survival too as misfolded protein accumulation can result in apoptotic pathway induction. It is important, to prevent downstream problems, for cells to only release and export correctly conformed proteins to the Golgi complex. The ER lumen encompasses quality control mechanisms to prevent or protect against mis- or un-folded proteins remaining in the folding and assembly process (Ellgaard and Helenius, 2003).

One part of the protection mechanism is ER-associated degradation (ERAD) which functions to remove proteins failing to fold sufficiently, be it slowly or

incorrectly, from the ER. Misfolded or aggregated proteins are removed from the ER and returned to the cytoplasm, where they are ubiquitinated (modified with ubiquitin molecules - markers for degradation), stripped of any glycans, before being broken down completely by the proteasome (Schröder, 2008). The process is advantageous in clearing the ER and keeping ER machinery available, preventing problematic proteins from otherwise stalling and clogging up ER processes. ER components such as BiP, PDI and EDEM1-3 (ER degradation enhancing a mannosidase like proteins 1-3) all play a role in detecting misfolded or aggregated proteins and directing them into the ERAD process (Schröder, 2008).

If the demand on the ER or the frequency of problematic proteins increases the ER compartment can become overloaded and stressed. If the folding requirement exceeds the folding capabilities of the ER, this leads to unfolded and incorrectly folded proteins accumulating. A process called the unfolded protein response (UPR) is then triggered; the purpose of which is to restore homeostasis to the ER, and when this is unachievable, to activate programmed cell death (apoptosis) (Hetz, 2012; Nishimiya, 2014; Walter and Ron, 2012). There are three signalling cascades that can be activated during an ER perturbation causing the UPR (outlined in **Figure 2.3**). The ER resident chaperone BiP plays a pivotal role in the activation of the signalling cascades. Under normal conditions BiP is bound to the three membrane-bound signal transducer components of the UPR – activating transcription factor 6 (ATF6), inositol-requiring protein 1 (IRE1) and protein kinase RNA-like ER kinase (PERK), repressing their activity (Hetz, 2012; Walter and Ron, 2012). When the concentration of unfolded proteins increases in the ER lumen, BiP switches from being bound to the UPR signal transducer components to preferentially interacting with hydrophobic regions on nascent unfolded chains, reducing their aggregation potential and premature transportation to the Golgi complex (Bertolotti et al., 2000; Chakrabarti et al., 2011; Shen et al., 2002). The release of BiP from the UPR signal transducer components results in their parallel activation, and commencement of their downstream signalling cascades.

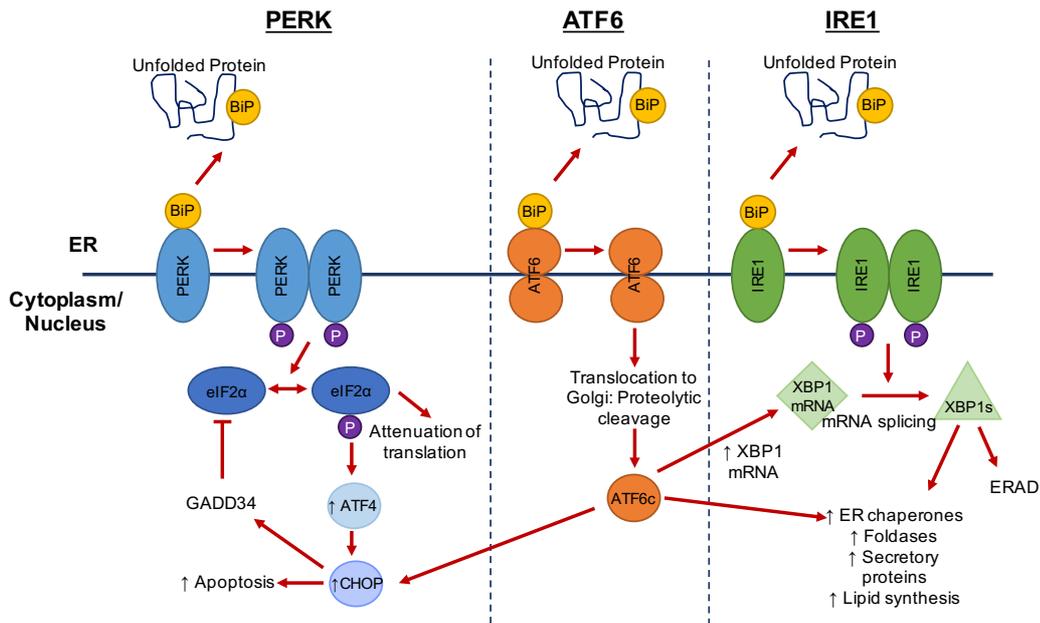


Figure 2.3: A Schematic Outlining Components of the Unfolded Protein Response (UPR)

Under normal cellular conditions, three separate UPR transactivators, PERK, ATF6 and IRE1 are bound to ER chaperone BiP. When unfolded proteins accumulate, the ER becomes stressed. The UPR is activated through the dissociation of BiP to the three transactivators as BiP preferentially associates with the unfolded proteins. PERK dimerises, phosphorylates eIF2 α , and leads to a reduction in protein translation, reducing the ER burden. It can also induce cell apoptosis through increased ATF4 translation, a necessary requirement if the balance in the ER cannot be restored. Upon BiP release ATF6 moves to, then undergoes proteolytic cleavage in, the Golgi complex. This cleaved form (ATF6c), when in the nucleus, regulates genes to increase the ER capabilities and reduce ER stress. Effects include increasing XBP1 translation, chaperone expression and the size of the ER. IRE1 dimerises and phosphorylates upon activation, this can then cause the splicing of XBP1 mRNA. This activates this global regulator of ER associated components, increasing the capacity and function of the compartment working towards restoring normal function and balance within the ER. The figure is adapted from Dinnis and James, 2005; Hussain et al., 2014 and Nishimiya, 2014.

The PERK signalling pathway primarily serves to reduce the protein folding load the ER receives by attenuating protein translation. PERK is a transmembrane kinase which oligomerises upon activation (BiP release) and phosphorylates eIF2 α (ubiquitination translation initiation factor) inactivating eIF2 which leads to a global inhibition of mRNA translation (Harding et al., 1999; Walter and Ron, 2012). Conversely, eIF2 inhibition increases translation levels of a transcription factor, ATF4. ATF4 induces expression of various genes involved in pro-survival

functions such as protein folding and balancing redox. It also induces CHOP (transcription factor C/EBP homologous protein) which regulates genetic components involved with apoptosis, cell death signalling pathways (Nishimiya, 2014; Oyadomari and Mori, 2004; Walter and Ron, 2012).

There are two forms of ATF6 in the cell, ATF6 α and ATF6 β , with ATF6 α the form associated with UPR signal transduction (Yamamoto et al., 2007). It is a transcription factor and, upon activation in response to ER stress, is transported to the Golgi complex where it undergoes proteolytic processing, resulting in its cleaved form: ATF6c. This 50kDa cleaved cytosolic domain advances to the nucleus to activate a plethora of genes to boost the capabilities of the ER, which help to meet the increased demand and ER stress (Chakrabarti et al., 2011). It upregulates the expression of ER chaperones genes including BiP and PDI as well as ERAD and XBP1 genes (Hetz, 2012; Walter and Ron, 2012; Yoshida et al., 2001).

The third UPR transducer, the IRE1 signalling pathway, works to increase the capacity and functional capabilities of the ER in a global way. Upon activation IRE1 dimerises, initiating its endoribonuclease activity, which splices the mRNA encoding the transcription factor XBP1. It cleaves the mRNA to remove an intron which, upon translation, produces the active form of the protein (XBP1s; Calfon et al., 2002; Chakrabarti et al., 2011; Walter and Ron, 2012). XBP1s proceeds to upregulate and switch on many components of the ER, including chaperone and ERAD genes, working to restore the capacity and normal workings of the ER and target the removal of unfolded proteins to degradation pathways (Chakrabarti et al., 2011; Lee et al., 2003).

2.4.3. The Secretory Pathway

Once a protein has been correctly processed, folded and assembled in the ER, it progresses through the secretory pathway. The secretory pathway involves the movement of proteins through a linear series of lipid bilayer membrane-bound compartments facilitated by membrane-bound transport vesicles. The series of membrane-bound entities undergo fusing and budding processes transporting

the protein along the secretory pathway and ultimately out of the cell (Lodish et al., 2008).

The steps incorporated into protein secretion in mammalian cells are outlined in **Figure 2.4**. Firstly, protein cargo leaves the ER at ER exit sites (ERES) through the formation of anterograde vesicle carriers, coated in coat protein complex II (COPII), which bind and concentrate the exiting proteins. The vesicles bud off from the donor ER membrane before fusing to the acceptor membrane on the ER-to-Golgi intermediate compartment (ERGIC). At the ERGIC, cargo is sorted – with transport proteins either being retained in the ERGIC or recycled back to the ER. The compartment produces large, anterograde transporting vesicles, containing coat protein complex I (COPI) coated vesicle carriers, which transfer cargo protein to the Golgi complex. COPI and COPII proteins are coating machinery and make up the transportation carrier vesicles. For successful membrane trafficking, the coat complexes work in synergy with Sar/ARF GTPases, deactivating GTPase activating proteins (GAPs) and activating guanine nucleotide exchange factors (GEFs). The Golgi complex is comprised of several stacked membranes (cisternae). The compartments contain various components of glycosylation machinery and function to modify and complete the glycosylation patterns linked to the cargo proteins. A *cis*-Golgi complex, positioned first closest to the ER, contains early modification machineries, such as mannosidase-II. Concluding glycosylation machinery, such as galactosyl-transferase, is present in the subsequent *trans*-Golgi complex, which faces the plasma membrane. Proteins are then transferred to the *trans*-Golgi network (TGN) to be lastly packaged and sorted. Some proteins are secreted by exocytosis, directly out of the cell across the plasma membrane, others are transferred to the endosomal system and others are stored in secretory granules (Lodish et al., 2008; Szul and Sztul, 2011).

Much of the secretory process, of trafficking between membranes, is regulated by members of the 'Rab small GTPase protein family' (Fukuda, 2008). Vesicle transportation and trafficking, between membranes, is mediated and regulated by two protein families: the tether protein family and the soluble N-ethylmaleimide sensitive factor receptors (SNARE) protein family. Before a vesicle fuses to an acceptor membrane, tethering factors create a bridging link attaching the donor

and acceptor membranes (Hong and Lev, 2014). Membrane fusion is then mediated by members of the SNARE protein family; v-SNAREs and t-SNAREs which are bound to the vesicle and target membranes, respectively, interact to continue the protein cargo transportation process (Jahn and Scheller, 2006; Szul and Sztul, 2011).

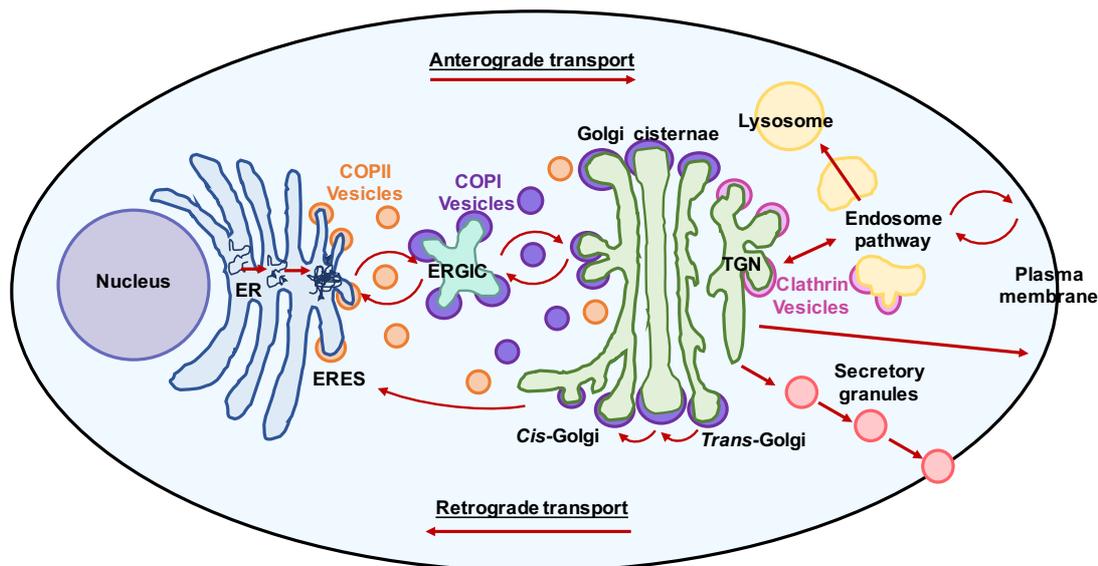


Figure 2.4: Schematic of the Secretory Pathway in Mammalian Cells

Folded proteins, destined for secretion out of the cell, undergo anterograde transport through a series of membrane-bound compartments. Once correctly folded, proteins leave the ER at ERES in COPII-coated vesicles. They undergo transportation; firstly, to ERGIC, for sorting, then onto the Golgi complex, comprised of several membrane cisternae, where further protein glycosylation modifications are undertaken. Proteins are then sorted in the TGN and delivered to the plasma membrane, to secretory granules or to endosomes. Components of the secretory pathway are returned and recycled back towards the ER, in COPI-coated vesicles by retrograde transportation. The figure adapted from Szul and Sztul, 2011.

2.4.4. Post-translational Bottleneck

It has been articulated in the literature that there are often bottlenecks downstream of translation, resulting in post-translational operations being rate-limiting (Dinnis and James, 2005; Le Fourn et al., 2014; Nishimiya, 2014; Schröder, 2008) An increase in transcription, through applying the current engineering strategies of amplification of gene copy number, use of strong promoters along with other plasmid and transfection approaches, does not

necessarily result in increased or sufficient increase in productivity levels (Le Fourn et al., 2014). It is thought that, in many cases, there is little to no correlation between the amount of recombinant protein secreted and the level of transcription undertaken (Nishimiya, 2014). This often leads to the implication that rate-limiting steps are likely to be downstream of transcription and translation. Cells may be unable to undertake sufficient processing of recombinant proteins, which consequently increases cellular and ER stress and toxicity levels and results in lower than optimal product titres.

2.5. Summary

A broad overview of the processes, platforms and developments related to the production of biopharmaceuticals has been described. Details of upstream process were of focus, including the dominance of the CHO cell host and methods used to introduce and express recombinant protein genes. These details were provided to help summarise the bioindustrial setting and its influences, along with contextualising the work undertaken within this thesis. The processes that take place within cells to synthesis, fold, process and secrete recombinant proteins were also outlined. A knowledge of these process provides context and understanding for the next chapter, where a specific and targeted investigation into CHO cell engineering is explored. The information was relevant to help provide design criteria for the development of a CHO cell engineering platform and information of previously engineered cellular pathways and could be potential useful targets. The chapter provides further context for the direction of the work undertake within the thesis.

3. Improvements in Therapeutic Protein Production

This chapter provides an insight into some of the methods and approaches utilised to improve the CHO cell production factory and the yield that it generates. It is focussed on the approach of overexpressing components of the cell, particularly those involved in protein folding and assembly, with the objective to increase the capability of the CHO cell as a production host. This information, together with the previous chapter, provides some context and rationale to the subsequent work included within this thesis.

3.1. Introduction

In the previous chapter the manufacture of recombinant therapeutic proteins using CHO cell factories was described. This provided a broad insight into the processes that take place, both in the biopharmaceutical production pipeline and at a cellular level, within the CHO cell itself. The aim of this chapter is a focussed investigation into the previous work that has been undertaken to improve this CHO cell production process and to increase the yields generated; continually of interest from both an economic and a therapeutic perspective. There is targeted attention to previous engineering examples demonstrating improvements made to the CHO cell factory and its production capabilities. The intention when reviewing the literature was to identify design requirements relevant for the development of a CHO cell engineering platform technology, in addition to highlighting genetic components that could be incorporated into the process.

3.2. A Need to Engineer the Production Host

It is inherently expensive to produce recombinant therapeutic proteins, therefore various engineering approaches have been implemented throughout upstream

process development to increase the concentration of recombinant protein produced per bioreactor run (Wurm, 2004). In the last 25 years, since the first recombinant therapeutic proteins successfully gained FDA-approval, the volumetric yields obtained from biopharmaceutical production have increased over 100-fold (Wurm, 2004). Average productivities have improved from milligram per litre to tens of grams per litre (Wurm, 2004). There is still drive for further improvement, with yields from CHO-based systems still below what is achievable with microbial-based fermentation.

There is ongoing pressure on the biopharmaceutical industry to improve the recombinant protein production process. One of the leading drivers for improvements is an increased demand for therapeutic drugs (Kelley, 2009). Some of the reasons for this increasing demand include the trend of increasing life expectancy, along with lifestyle changes, resulting in a higher prevalence of cancer-related diseases (Jemal et al., 2011; Mariotto et al., 2011; Siegel et al., 2016). Targeted therapies for cancer are the most common recombinant mAb drug products being developed and manufactured. In order for health services and governments to continue to try and meet this growing demand for treatments, there is a real need for a reduction in the cost of products, and for industries to supply more cost-effective treatments that are realistic for use in large populations, including those in the developing world. Improvements in the production process will contribute to a reduction in cost, and as a result, more people will have access to, and benefit from recombinant therapeutic products.

There is an expanding number of therapeutics in the development and approval process (Butler and Meneses-Acosta, 2012), for example, between 2003 and 2006 the number of drug candidates in clinical trials increased from 75 to 400 mAb products alone (Coco-Martin and Harmsen, 2008). The type of recombinant products being tested in these trials are of increasing complexity; with an increased prominence of engineered antibody, bispecific and fusion products, exemplified in **Figure 3.1** (Kontermann and Brinkmann, 2015; Nelson, 2010; Walsh, 2010; Walsh, 2014). Many of these have increased effectiveness, with increased functionality, targeting more than one antigen, and are smaller in size bringing them in closer proximity to their target. Their increased engineered structures can be much more challenging to express and manufacture, often

referred to as DTE proteins. Titres produced can be unpredictable, and often lower than standard gram per litre production titres achieved by standard mAbs (Johari et al., 2015; Pybus et al., 2013). These products are likely to have heightened associated process and production costs, along with prolonged bioprocessing development. There is a need to develop manufacturing strategies for these candidates to reduce their burden on, or loss from, the development pipeline. The increased presence of biosimilar and biobetter candidates in drug development (highlighted in **Section 2.1.2**; Beck, 2011; Walsh, 2010) also demonstrate the developments and trends in drug manufacture. For these types of product to be competitors of the original, the production process needs to generate drug product with robust and/or improved product quality, at a reduced cost and timeframe, to decrease the cost of goods and therefore market price. Taken as a whole, there is a continuing incentive for drug development to be improved, and drug manufactured in a more cost and time effective way (Birch and Racher, 2006).

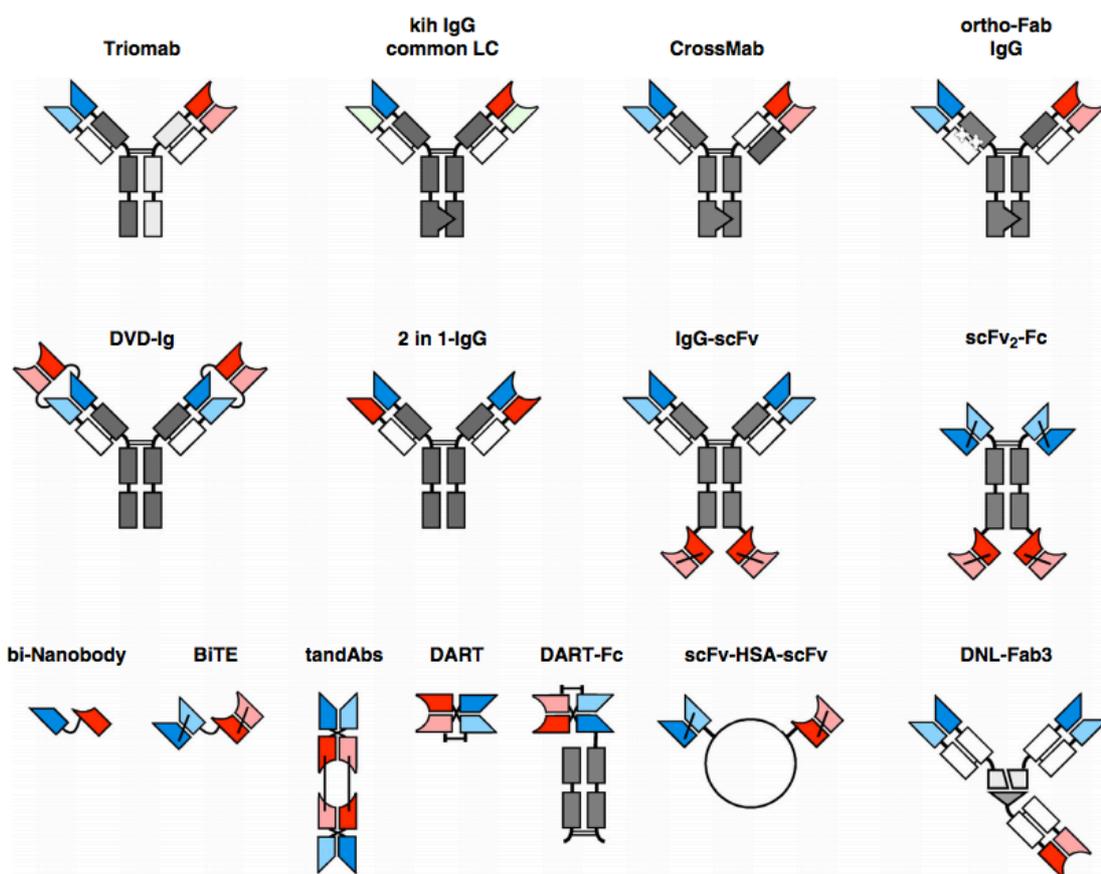


Figure 3.1: Schematic Demonstrating the Range of Engineered Recombinant Proteins in Clinical Trials

There are increasing numbers of engineered, more complex recombinant products within development. This schematic (taken directly from Kontermann and Brinkmann, 2015) displays some of the range of structures currently in biopharmaceutical development, clinical trials or have been successfully approved drugs for clinical use. Some are bispecific or multi-valent derivatives of mAb structures (top two lines), others are fusion proteins (bottom line).

3.3. Approaches to Enhance Production Levels

Improvements in recombinant protein production, in terms of volumetric yields, cost, efficiency, and timelines, stem from multiple optimisation steps in the manufacturing process. This has involved improving the cell line itself, and the culture conditions, and by increasing the size of the bioreactor used for the production (Walsh, 2010). The upstream strategies employed to achieve the improvement in recombinant protein yields stem from process optimisation and expression-based technology developments (Birch and Racher, 2006; Wurm, 2004). Some are expanded on subsequently, but include fed-batch feeding

strategies, extended culture durations, use of biphasic cell culture, and a better understanding of the nutritional requirements of cells (Birch and Racher, 2006; De Jesus and Wurm, 2011; Xie and Wang, 2006). In addition, improvements stem from introducing an ability to amplify selection marker genes through the use of essential enzymes which indirectly increases the copy numbers of the recombinant gene of interest, and therefore transcription. Knockout cell lines have been generated to increase the stringency of selection systems. There have been improvements in the identification process of high producing clones and in targeting gene integration into a transcriptionally active section of heterochromatin. (Birch and Racher, 2006; Fan et al., 2012; Wurm, 2004). Many improvements have involved methods, such as directed evolution and mutagenesis, which are usually difficult to foresee and predict, require significant time, and are often unreliable and hard to control (Hu et al., 2012; Jayapal et al., 2007). The confidential and proprietary nature of advances made by industry, have hampered the full-disclosure of all improvements in the field (De Jesus and Wurm, 2011; Wurm, 2004).

One of the biggest contributions has been from changes to media design and feeding regimes (De Jesus and Wurm, 2011). Fed-batch cultures are widely used with calculated, optimised and concentrated feeds, boosting and prolonging cell growth and therefore productivity (Whitford, 2006). For example, in an 18 day period, using a fed-batch optimised culture mode, cell densities of 20×10^6 cells mL^{-1} and yields of $10\text{-}13 \text{ g L}^{-1}$ were possible (Huang et al., 2010). The implementation of disposable equipment, including single-use bioreactors, is a more recent improvement. Single-use equipment reduces turnaround times and the need for cleaning; a combination of which can lower capital costs by 40% (Shukla and Gottschalk, 2013). A biphasic culture approach is routinely implemented to increase titre; the approach is based upon knowledge that recombinant protein production is elevated when cell growth is slowed (Dinnis and James, 2005). During biphasic growth, cells increase in number during an initial growth phase, then growth is slowed when viability is high, to increase productivity. The recombinant product accumulates as the cell is using energy to make the product, rather than replicating. Methods to slow cell growth include altering osmolality to halt cell division, and a reduction in culture temperature

(Dinnis and James, 2005). For example, a common implementation of two phase culture begins with an initial 37°C period of growth, which generates high levels of biomass. A shift to a hypothermic culture temperature (32°C), reduces cell growth but increases protein synthesis, through increased mRNA levels and stability, and therefore the yields generated (Fox et al., 2005; Kou et al., 2011; Masterton and Smales, 2014). However, it must be noted that the use of a reduced temperature can have the potential to negatively affect product quality (Sou et al., 2015).

Small molecule effectors, low molecular weight compounds, also referred to as chemical chaperones, have shown to be another means of enhancing recombinant protein production. They can be easily added to the culture and provide protein chaperone-like functionality, in aiding and enhancing protein folding within the ER, and some have been shown to reduce the levels of misfolded protein (Cortez and Sim, 2014; Rajan et al., 2011). Chemical chaperones have potential therapeutic effects in themselves for neurodegenerative diseases associated with protein folding problems, such as Alzheimer's (Cortez and Sim, 2014). Their method of action is not completely understood, which is not ideal for therapeutic use or in manufacture, but have been beneficially linked to reducing protein aggregation, improving protein trafficking and secretion, and stabilising misfolded proteins (Johari et al., 2015; Perlmutter, 2002; Roth et al., 2012). Promising compounds include glycerol, proline, DMSO, 4-phenylbutyric acid (PBA), VPA, sodium butyrate, and betaine. Some of these compounds are osmolytes, altering osmotic pressures, enhancing protein stabilisation and folding; others have histone deacetylase activity, reducing transcriptional gene silencing. All have been shown to boost recombinant protein production in CHO cells, and reduce unwanted protein aggregation (Hwang et al., 2011; Johari et al., 2015; Roth et al., 2012; Sung and Lee, 2005; Wulhfard et al., 2010).

3.4. Cell Engineering Approaches to Enhance CHO Cell Factory Performance

Another prevalent approach utilised, to enhance recombinant protein production, is genetic engineering of the CHO cell factory machinery. Approaches have

commonly involved the upregulation or overexpression of particular genes; but, with the advancement of technology and knowledge, approaches such as Clustered Regularly Interspaced Short Palindromic Repeats-associated protein-9 nuclease (CRISPR/Cas9) mean genetic downregulation and suppressive approaches are also being utilised (Fischer et al., 2015; Lim et al., 2010). The main aim of engineering the cell host is to improve its bioprocessing capabilities; to increase the specific productivity (qP) of the cell and/or to increase the growth capabilities or duration of culture, which will indirectly increase yield (Mohan et al., 2008). There are many features of the cellular machinery that have been targeted and altered (**Figure 3.2**), including the key bioprocessing steps of protein folding, modification, assembly in the ER, transport and secretion through the Golgi complex, as well as cell proliferation, stress and death pathways (Kim et al., 2012; Mohan et al., 2008).

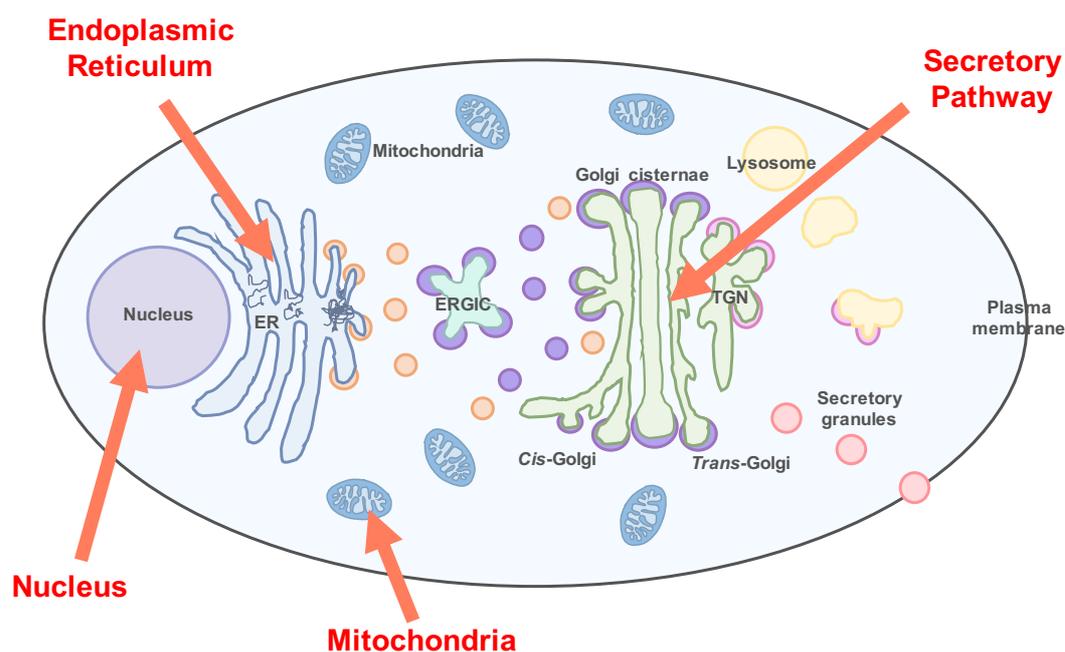


Figure 3.2: Targets of CHO Cell Engineering

The majority of cell engineering strategies have been targeted within the ER, where the key processes of recombinant protein folding and assembly take place. Fewer studies have targeted the Golgi complex, in processes related to recombinant protein secretion out of the cell. Other major targets have been aimed at prolonging cell culture longevity including anti-apoptotic pathways in the mitochondria. Genes which function in the nucleus to alter gene transcription and react to cell stress have been other common targets.

3.4.1. Development of Omics-based Approaches

There has been a surge in the progression of omics-based technologies and information that have accompanied and advanced cellular engineering of production systems. Omics technologies, including genomics, transcriptomics and proteomics, involve large scale studies of molecules, such as DNA, RNA and proteins, respectively, over a time course or between different conditions (Lewis et al., 2016). Recent CHO-based omics developments, stem from improvements in ease, ability, cost, and speed of sequencing technology, have included sequencing of the first CHO genomes (Brinkrolf et al., 2013; Lewis et al., 2013; Xu et al., 2011) and transcriptomic-based studies (Becker et al., 2011; Hernández Bort et al., 2011; Rupp et al., 2014). Despite all the successes, one of the major limitations remains an insufficient ability to process and interpret the large amounts of data generated (Lewis et al., 2016). Omics-sourced information can contribute to characterising particular CHO phenotypes, finding genes differentially expressed at distinctive time points in culture (Kildegaard et al., 2013; Nishimiya, 2014). Omics has the potential to improve numerous aspects of CHO cell factories and recombinant protein production. Particularly by contributing to engineering of the host to improve manufacturing performance; and in assisting with identifying and understanding potential cellular targets to engineer, which may improve production capabilities (Datta et al., 2013; Hansen et al., 2017; Lewis et al., 2016).

3.4.2. Repression of Genetic Components

Despite not being the focus here, there is a plethora of approaches involved in down-regulating CHO cell components. The principle of genetic down-regulation being that, if unfavourable detrimental genes can be repressed, cellular processes can be favourably altered, resulting in improved bioprocessing and productivity (Fischer et al., 2015; Krämer et al., 2010). There are various methods developed that can result in genes being silenced, or even stably deleted.

Firstly, there are approaches which function to knock out genes. Traditionally this was achieved using random mutagenesis, with chemical or radiation treatment. These are very unspecific approaches, but have been successful, for example,

in knocking-out the DHFR gene, generating DHFR negative cell lines that are the precursors of many cell lines in use today (Fischer et al., 2015). The use of nucleases, which cut and alter genomic DNA, are a more precise, highly specific approach. Methods include zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas9. ZFNs have been used to beneficially repress pro-apoptotic genes; including BAX and BAK, resulting in a five-fold increase in mAb production (Cost et al., 2010). They have also been used to improve the functionality of a recombinant therapeutic through altering the glycan profile of the product. A knock out of the FUT8 gene, which catalyses the addition of fucose sugars onto the Fc region of mAbs, results in mAbs without a core fucose residue, this enhances their antibody-dependent cellular cytotoxicity (ADCC) and anti-tumour activity (Yamane-Ohnuki et al., 2004). Developments in CRISPR/Cas9 technology are improving the ease, cost, speed, and quality of results from nuclease-based genetic knockdown (Lee et al., 2015). This tool has been used to create multiple genetic knockdowns in a cell line, which has repressed expression of FUT8, BAX and BAK simultaneously, without introducing off target effects (Grav et al., 2015).

A second knockdown approach utilises interfering RNA (RNAi) to cause gene silencing (Wu, 2009). Short, double-stranded RNA, known as siRNA or shRNA, which is complementary to the target mRNA has been used to beneficially alter apoptosis, glycosylation, metabolism and cellular production (Fischer et al., 2015; Wu, 2009). For example, cells with reduced levels of apoptosis, were generated through siRNA targeted to various caspases and BAX/BAK (Kim and Lee, 2002; Lim et al., 2006; Sung et al., 2007).

The third approach to genetic repression utilises microRNA (miRNA). These are naturally occurring, non-coding RNA, which function to regulate large, or multiple cellular pathways, through inhibiting translation, similar to transcription factors (Barron et al., 2011b; Hackl et al., 2012). Developments in miRNA have stemmed from the increase in knowledge from genomic and other global omics studies, but there is a need for further phenotypic characterisation of miRNAs for continued success. The use of miRNAs has shown to increase cell growth and specific productivity. For example, components of the miR-30 family, have been linked to enhanced protein production (Fischer et al., 2014). miR-7, identified through

studies of hypothermic culture, caused an increase in qP, despite negatively affecting cell growth (Barron et al., 2011a). Repression, silencing or knock-out of genes, is a promising and developing alternative to overexpression genetic engineering to improve host bioprocessing capabilities and qualities, although increased knowledge and understanding of targets is needed to help with selection and to prevent unwanted off-target effects.

3.4.3. Overexpression of Genetic Components

Overexpression of genetic components has been a common approach, over the last decade, to improve recombinant protein production and host cell manufacturing performance. This approach primarily harnesses recombinant DNA technology; ectopic transgenes, referred to as effector genes herein, are introduced and expressed in the host cell, often alongside the recombinant gene product. There are many published examples of the upregulation of different components in the literature, with many of these highlighted in **Table 3.1 to Table 3.5**. Many studies involve overexpression of components in key cellular processes, for example anti-apoptotic, UPR, folding and assembly pathways. There is a focus on ER-related components, which directly or closely interact with the recombinant protein. In recombinant protein production, particularly TGE of DTE proteins, where recombinant gene copies are high, there can be an imbalance between protein synthesis, assembly and secretion processes, these have the potential for improvement. There are often post-translational bottlenecks and mismatched rates of transcription, translation, folding and secretion (Davies et al., 2011; O'Callaghan et al., 2010; Pybus et al., 2013). Engineering the ER remains an obvious target for genetic expression based approaches.

3.4.3.1. Anti-Apoptotic Engineering

Reducing apoptosis, a genetically-regulated type of programmed cell death, has been a common engineering target. The rationale behind targeting apoptosis, is that by reducing the level of cell death, cultures might beneficially maintain or improve their viability, extend their culture longevity, increase their growth and/or increase the maximum VCD reached (Kim et al., 2012; Mohan et al., 2008). Overexpression of anti-apoptotic genes has been the primary target, with

examples shown in **Table 3.1**. In particular, members of the Bcl-2 family of proteins which function to inhibit apoptosis have been targeted (Chiang and Sisk, 2005; Kim and Lee, 2000; Majors et al., 2008b; Majors et al., 2009). Upregulation of various Bcl-2 components has been successful in approximately doubling the titres of the recombinant proteins being expressed.

Table 3.1: Anti-Apoptotic Engineering Approaches

Genetic engineering examples targeting anti-apoptotic components. SGE and TGE refer to stable gene expression and transient gene expression, respectively. Abbreviations: Heat shock protein 27 (HSP27). In addition to specific references within the table the following reviews were also sources of information: Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008 and Nishimiya, 2014.

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
Bcl-2	2-3-fold increase	Antibody	CHO	SGE	SGE	<ul style="list-style-type: none"> - Cells cultured in with sodium butyrate (chemical enhancer, may have improved production) - Extended culture longevity (overcame sodium butyrate growth inhibition) 	Kim and Lee, 2000
Bcl-xL	1.9-fold increase	Antibody (mAb)	CHO	SGE	SGE	<ul style="list-style-type: none"> - Increased cell viability - Further enhancement with sodium butyrate treatment 	Chiang and Sisk, 2005
	1.7-3.7-fold increase	Fusion protein	CHO	SGE	TGE	<ul style="list-style-type: none"> - Prolonged culture viability 	Majors et al., 2008
HSP27	2.3-fold increase	Antibody (mAb)	CHO	SGE	SGE	<ul style="list-style-type: none"> - 2.2-fold higher peak cell density and prolonged culture viability 	Tan et al., 2015
Mcl-1	1.2-1.35-fold increase	Antibody (mAb)	CHO	SGE	SGE	<ul style="list-style-type: none"> - Bcl-2 family protein member - Prolonged culture viability 	Majors et al., 2009

3.4.3.2. Engineering Post-Translational Mechanisms

Targeting components of the post-translational machinery, cellular targets that function to modify, fold, and assemble proteins within the ER, is another common target for genetic engineering. The main aim behind the efforts in engineering this part of the cell, is to try and reduce post-translational bottlenecks. As mentioned previously, it has been found that the titre does not always correlate with the number of copies of the recombinant gene present, the levels of its mRNA or the amount of intracellular recombinant protein (Fann et al., 1999; Ku et al., 2007; Mohan et al., 2008). This leads to the idea that there are production bottlenecks in translational, post-translational, and/or secretory cellular pathways, these therefore have the potential to be engineered, reduced or even overcome. Cells that have varying productivity outputs can have the same level of mRNA, indicating that transcription is usually not the rate limiting step in production (Hussain et al., 2014; Nishimiya, 2014). In addition to this, studies of proteins, including global proteomic studies, within recombinant producing cell lines, have elucidated that there is a positive relationship between cell productivity output and the relative amounts of ER chaperones and foldases present (Dinnis et al., 2006; Dorner et al., 1989; Smales et al., 2004). There is also evidence of a similar correlation on a transcriptome level (Doolan et al., 2008). High levels of expression are associated with upregulation of components of the ER pathway, for example proteins like BiP, PDI and Grp94. Protein studies of recombinant producing hosts have also revealed that ER chaperone levels increase during periods of ER stress and in late stage culture, where recombinant protein productivity is highest (Dorner et al., 1989; Nishimiya, 2014; Pascoe et al., 2007). High abundance of ER components underpins high production characteristics, and are therefore attractive engineering targets.

Many post-transcriptional targets involved in protein folding and assembly pathways in the ER have been upregulated in recombinant producing cell lines with various effects. Previous targets are exemplified in **Figure 3.3** and in **Table 3.2**. The hypothesis is that increasing the folding and assembly machinery within a cell would increase its production capabilities (Borth et al., 2005).

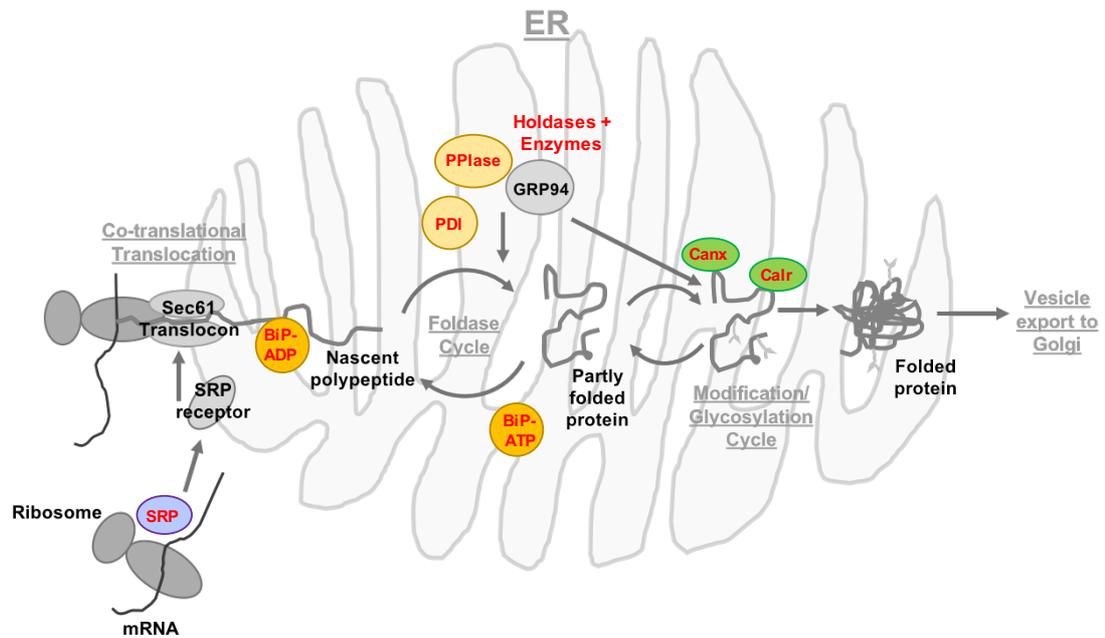


Figure 3.3: Engineering of Protein Folding and Assembly in the ER

ER components, functioning in relation to protein folding (e.g. PDI, PPlase) and processing into and through the ER (e.g. SRP and BiP) that have been previously overexpressed are highlighted. Further information regarding overexpression can be found in **Table 3.2**.

Numerous studies have focussed on overexpression of PDI, an enzyme catalysing disulphide bond formation in forming proteins in the ER. These studies provide a good example of how varied, inconsistent, and in some ways contradictory the effects of cell engineering approaches can be. Upregulation of PDI has been shown to produce negative, positive and no effect on recombinant protein production (Davis et al., 2000; Mohan et al., 2007). Each study was performed under different circumstances; for example, different product genes, cellular hosts, and expression platforms. All these changes will alter the bottlenecks and requirements of the production process and therefore the responses generated (Hansen et al., 2017; Mohan et al., 2008). It is suggested that upregulation of ER proteins like PDI would be more beneficial when its requirement was higher, for example in the production of DTE proteins or multimeric proteins, where there is a higher number of disulphide bonds to be formed (Pybus et al., 2013). The variable outcomes are seen in other engineering examples too, for example upregulation of BiP, has been shown to both increase and decrease titre in different examples (Borth et al., 2005; Brown et al., 2011;

Johari et al., 2015; Pybus et al., 2013). Due to their potential benefit, and major role within recombinant protein folding and assembly in the ER, these components still remain potential targets to boost production. There are many other ER components that hold promise; upregulations of some of them has caused beneficial effects on titre, such as the chaperones Canx and Calr (Chung et al., 2004) and ER enzyme ERO1L (Mohan and Lee, 2010). However, there are many components that remain untested or have not been published in the literature.

Table 3.2: ER Protein Folding and Assembly Engineering Approaches

Genetic engineering examples targeting protein folding and assembly components in the ER. In addition to specific references within the table the following reviews were also sources of information: Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008 and Nishimiya, 2014.

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
BiP	1.34-fold decrease	Antibody (mAb)	CHO	SGE	SGE		Borth et al., 2005
	1.5-fold decrease	Human coagulation factor VIII	BHK	TGE	SGE		Brown et al., 2011
	No effect	Porcine coagulation factor VIII	BHK	TGE	SGE		Brown et al., 2011
	Up to 1.5-fold increase	Antibody (DTE mAb)	CHO	TGE	TGE	- Increased qP and reduced cell growth - No effect on an ETE mAb	Pybus et al., 2013
	Up to 1.3-fold increase	Fc-fusion protein	CHO	TGE	TGE	- Titre increase from increased qP - Dose response effect	Johari et al., 2015
Calr + Canx	1.9-fold increase	Thrombopoietin	CHO	SGE	SGE	- No effect on cell growth - Inducible expression system	Chung et al., 2004
CypB	1.4-fold increase	Antibody (mAbs)	CHO	TGE	TGE	- Improved cell growth with no effect on qP - No effect on an ETE mAb	Pybus et al., 2013
	Up to 1.3-fold increase	Fc-fusion protein	CHO	TGE	TGE	- Increase in titre stem from increased growth - Dose response effect	Johari et al., 2015
ERO1L	1.37-fold increase	Antibody	CHO	TGE	SGE	- No or negative effect for SGE	Mohan and Lee, 2010

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
ERp57	1.7-2.1-fold increase	Thrombopoietin	CHO	SGE	SGE	- Member of PDI protein family - Inducible expression system	Hwang et al., 2003
ERp72	No effect	Antibody (mAb)	CHO	TGE	SGE	- Member of PDI protein family	Hayes et al., 2010
PDI + BiP	1.35-fold decrease	Antibody (mAb)	CHO	SGE	SGE		Borth et al., 2005
PDI + ERO1L	1.55-fold increase	Antibody	CHO	TGE	SGE	- No or negative effect for SGE	Mohan and Lee, 2010
PDI	No effect	IL-15	CHO	SGE	SGE		Davis et al., 2000
	Decrease	TNFR:Fc	CHO	SGE	SGE		Davis et al., 2000
	1.37-fold increase	Antibody (mAb)	CHO	SGE	SGE		Borth et al., 2005
	1.27-fold increase	Antibody	CHO	SGE	SGE	- Inducible expression system	Mohan et al., 2007
	No effect	Thrombopoietin	CHO	SGE	SGE	- Inducible expression system	Mohan et al., 2007
	No effect	Antibody (mAb)	CHO	TGE	TGE		Hayes et al., 2010
	No effect	Multiple mAbs	CHO	TGE	TGE	- No effect on an ETE mAb	Pybus et al., 2013
	2-fold increase	Antibody	COS-1	TGE	TGE		Nishimiya et al., 2013
	Up to 1.2-fold increase	Fc-fusion protein	CHO	TGE	TGE	- Increased aggregate formation	Johari et al., 2015
	No effect	Antibody (mAb)	CHO	TGE	SGE	- Pancreatic PDI (PDlp)	Hayes et al., 2010

3.4.3.3. Engineering Components of the UPR

Targeting components of the UPR has been another established approach used to enhance recombinant protein production, exemplified in **Figure 3.4** and **Table 3.3**. Manipulating proteins involved in the UPR is a more global approach, and could have improved benefit over more targeted approaches (Lim et al., 2010). Upregulation of signal transducers that activate UPR pathways can increase the expression of an extensive array of ER genes, in a meaningful stoichiometric fashion; and is described as a more promising engineering approach (Lee et al., 2003; Mohan et al., 2008).

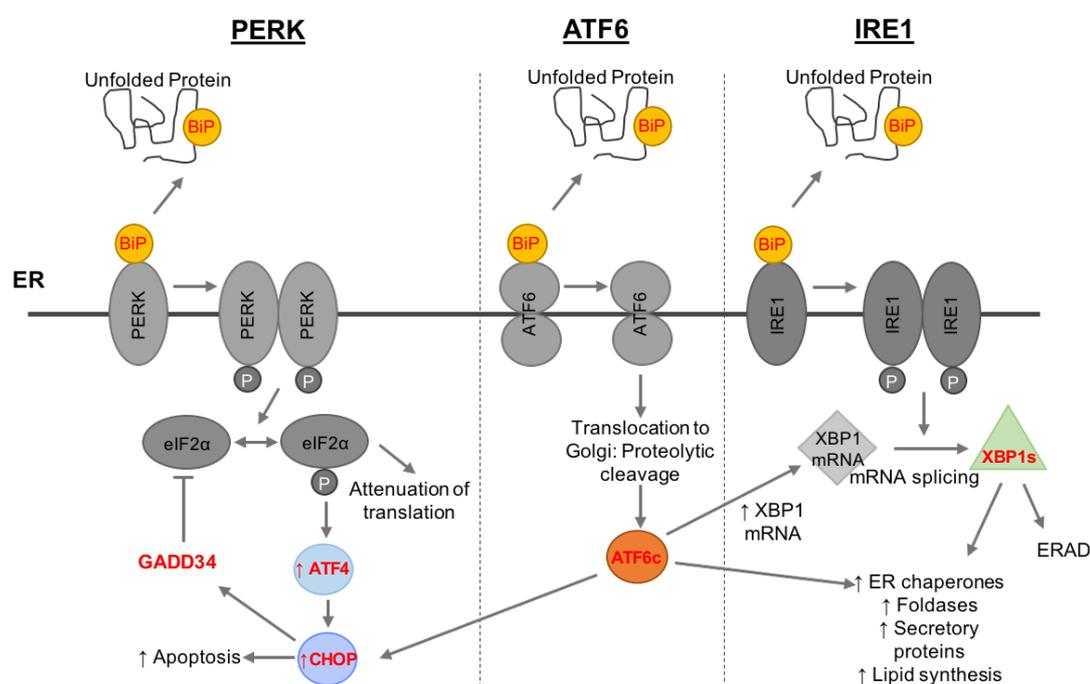


Figure 3.4: Engineering of UPR

Components with roles in the UPR have been previous engineering targets. Particularly transactivators, ATF6c and XBP1s, which both activate many downstream cellular components and pathways. Further information regarding overexpression can be found in **Table 3.3**.

The active form of XBP1 (XBP1s) has been overexpressed in several studies, predominantly with advantageous consequences (Becker et al., 2010; Codamo et al., 2011; Ku et al., 2007; Pybus et al., 2013; Tigges and Fussenegger, 2006). When activated, XBP1s, a transcription factor, initiates expression of a wide range ER components, for example BiP and PDI (Lee et al., 2003), and physically increases the size and capacity of the ER compartment (Tigges and

Fussenegger, 2006). However, overexpression of XBP1s does not always lead to increases in titre. Its effect is more apparent in TGE systems and for the production of DTE proteins; where the host secretory capacity is often increased beyond that of SGE systems, where the number of recombinant gene copies being processed are often lower (Ku et al., 2007; Pybus et al., 2013). Other engineered components of the UPR include the transcription factors ATF4, ATF6 and CHOP (Haredy et al., 2013; Nishimiya, 2014; Ohya et al., 2008; Pybus et al., 2013). These too affect the global status of the ER, affecting translation and ER capacity. It should be noted, engineering UPR components can have the potential to cause negative effects, in the form of initiating apoptosis, a major UPR downstream pathway (Becker et al., 2010; Pybus et al., 2013).

3.4.3.4. Engineering the Secretory Pathway

Components of the secretory pathway have been less frequently investigated as engineering targets, with the effect from overexpression of only a handful of components being described (**Table 3.4**). This leaves considerable scope for improvements through secretory pathway engineering. Published examples primarily involve engineering vesicle to membrane fusion and trafficking, at points between the ER, Golgi complex, and plasma membrane (Peng et al., 2011; Peng and Fussenegger, 2009). Examples have demonstrated that successful improvements in titre are achievable through secretory pathway engineering, and are potentially further enhanced when more than one component is overexpressed at once (Peng and Fussenegger, 2009).

3.4.3.5. Other Upregulated Components

There are many other parts of the cellular machinery that can be engineered, some previous examples are shown **Table 3.5**. For example components of the cell cycle, metabolism, and glycosylation pathways (Fischer et al., 2015; Mohan et al., 2008). Engineering transcription factors and large-scale regulators, such as YY1 and mTOR, have a wide scope in positively affecting the host's bioprocessing capabilities, through their effect on many downstream components (Dreesen and Fussenegger, 2011; Tastanova et al., 2016).

Table 3.3: UPR Engineering Approaches

Genetic engineering examples targeting UPR components. Abbreviations: Growth arrest and DNA damage inducible protein 34 (GADD34), x-linked inhibitor of apoptosis (XIAP), and Secreted Embryonic Alkaline Phosphatase (SEAP). In addition to specific references within the table the following reviews were also sources of information: Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008 and Nishimiya, 2014.

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
ATF4	2-fold increase	Antithrombin-III	CHO	SGE	SGE		Ohya et al., 2008
	1.8-2.5-fold increase	Antibody (mAb)	CHO	SGE	SGE		Haredy et al., 2013
	Up to 1.4-fold increase	Human Proteins (hα1AT, hC1INH)	CHO	TGE	TGE		Hansen et al., 2015
ATF6c	Up to 1.5-fold increase	Antibody (DTE mAb)	CHO	TGE	TGE	- Increased qP and reduced cell growth - No effect on an ETE mAb	Pybus et al., 2013
	Up to 1.1-fold increase	Fc-fusion protein	CHO	TGE	TGE	- Increases qP, represses growth	Johari et al., 2015
CHOP	2-fold increase	Antibody	COS-1	TGE	TGE	- Overexpression of CHOP with other ER effectors further enhanced effects.	Nishimiya et al., 2013
GADD34	1.4-fold increase	Antithrombin-III	CHO	SGE	SGE	- Transcription factor (increases dephosphorylation of eIF2α)	Omasa et al., 2008
XBP1s	6-fold increase	SEAP	CHO	TGE	SGE	- Expanded ER	Tigges and Fussenegger, 2006
	4-fold increase	SAMY	CHO	TGE	SGE	- Expanded ER	Tigges and Fussenegger, 2006

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
XBP1s	No effect	Antibody (mAb), EPO, IFN γ	CHO	TGE	SGE		Ku et al., 2007
	2.5-fold increase	EPO	CHO	TGE	TGE	- No effect for SGE	Ku et al., 2007
	2-fold increase	EPO	NS0	TGE	TGE	- No effect for SGE	Ku et al., 2007
	1.4-fold increase	Antibody (mAb)	CHO	SGE	SGE		Becker et al., 2008
	No effect	Human Factor VIII	CHO	TGE	TGE		Campos-Da-Paz et al., 2008
	Increase	Human Factor VIII	HepG2	TGE	TGE		Campos-Da-Paz et al., 2008
	No effect	Antithrombin-III	CHO	SGE	SGE		Ohya et al., 2008
	1.4-fold increase	Antibody (mAb)	CHO	SGE	SGE	- Elevated cell death	Becker et al., 2010
	1.37-fold increase	Antibody (mAbs)	CHO	TGE	TGE	- Hypothermic (32°C) culture used	Codamo et al., 2011
	Up to 1.6-fold increase	Antibody (mAb)	CHO	TGE	TGE	- Increased qP and reduced cell growth - No effect on an ETE mAb	Pybus et al., 2013
	No effect	tPA	CHO	SGE	SGE		Rahimpour et al., 2013
	Up to 1.4-fold increase	Human Proteins (α 1AT, hC1INH)	CHO	TGE	TGE		Hansen et al., 2015
	Up to 1.1-fold increase	Fc-fusion protein	CHO	TGE	TGE	- Stimulates qP, represses growth	Johari et al., 2015
XBP1s + ERO1Lα	5.3-6.2-fold increase	Antibody (mAb)	CHO	SGE	TGE	- Product quality unaffected	Cain et al., 2013
XBP1s + XIAP	2-fold increase	Antibody (mAb)	CHO	SGE	SGE	- Reduced cell death	Becker et al., 2010

Table 3.4: Secretory Pathway Engineering Approaches

Genetic engineering examples targeting components of the secretory pathways. Abbreviations: Ceramide-transfer protein (CERT), Synaptosome-associated protein of 23kDa (SNAP23) and Vesicle-associated membrane protein 8 (VAMP8). In addition to specific references within the table the following reviews were also sources of information: Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008 and Nishimiya, 2014.

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
CERT	1.26-fold increase	Antibody (mAb)	CHO	SGE	SGE	- Role in transport of proteins from ER to Golgi complex to plasma membrane	Florin et al., 2009
CERT S132A	1.35-fold increase	tPA	CHO	SGE	SGE	- Mutated version of CERT (phosphorylation resistant)	Rahimpour et al., 2013
Sly1 + Munc18c + XBP1	Up to 20-fold increase	Antibody (mAb)	CHO	SGE	SGE	- Sly1 and Munc18c proteins regulate membrane fusion, part of exocytosis - Increases from SGE of Sly, Sly+Munc and Sly+XBP1 also reported - Increased through increased qP	Peng and Fussenegger, 2009
SNAP-23	3-fold increase	Antibody (mAb)	CHO	SGE	SGE	- A member of the SNARE protein family - Increase also noted for TGE and other products	Peng et al., 2011
VAMP8	3-fold increase	Antibody (mAb)	CHO	SGE	SGE	- A member of the SNARE protein family - Increase also noted for TGE and other products	Peng et al., 2011
30Kc6	3.8-fold increase	Antibody (mAb)	CHO	SGE	SGE	- Anti-apoptotic protein isolated from silkworm haemolymph - Hyperosmotic medium used, contributed to enhanced effect	Wang et al., 2012

Table 3.5: Other Cellular Engineering Approaches

Genetic engineering examples targeting a variety of miscellaneous cellular components. Abbreviations: Ying-yang 1 (YY1). In addition to specific references within the table the following reviews were also sources of information: Fischer et al., 2015; Hansen et al., 2017; Hussain et al., 2014; Kim et al., 2012; Lim et al., 2010; Mohan et al., 2008 and Nishimiya, 2014.

Target	Effect on Titre	Product	Host	Effector Gene Expression	Recombinant Product Expression	Key Facts	Reference
E2F-1	No effect	Antibody (mAb)	CHO	SGE	SGE	- Cell cycle transcription factor - Increased maximum cell density by 20%	Majors et al., 2008a
mTOR	4-fold increase	Antibody (mAb)	CHO	TGE	SGE	- Stimulated cell growth, viability and qP	Dreesen and Fussenegger, 2011
SRP14	Increase	Antibody (mAb)	CHO	SGE	SGE	- DTE and ETE mAbs tested - Enhanced when SRP14 was overexpressed with other translocation components	Le Fourn et al., 2014
YY1	Up to 6-fold increase	Antibody (mAb)	CHO, Human cell lines	TGE	SGE	- Transcription factor - Species specific interactions – CHO YY1 effective in CHO cells, human YY1 effective in human cells - Various results from TGE and SGE expression with different products	Tastanova et al., 2016
ZFP-TF (LK52)	10-fold increase	Antibody (mAb)	CHO	TGE	SGE	- An artificial zinc finger protein transcription factor	Kwon et al., 2006

3.4.3.6. Multiple Component Engineering

Most genetic engineering examples described to-date involve the upregulation of a singular genetic component. There are few positive examples describing alterations of multiple genetic components, through co-overexpression of combinations of genes together (some of which are shown in **Table 3.2**, **Table 3.3** and **Table 3.4**). For example, work undertaken by Cain et al, involving stable upregulation of XBP1s and ERO1L α , significantly improved productivity, with an approximate 6-fold increase in volumetric titre of a transiently produced recombinant product (Cain et al., 2013). Co-expression of multiple components have the potential to have additive, or even synergistic effects. Secretory pathway components, Sly1, Munc18c and XBP1s, were shown to have beneficial effects when expressed alone, but demonstrated an additive effect on secretion by upregulation of all three components in concert (Peng and Fussenegger, 2009). Genetic overexpression can also successfully be combined with small molecule chemical chaperones or process alterations too, to synergistically improve DTE protein titres (Johari et al., 2015). These examples help to highlight that it is unlikely that upregulation of a singular component would produce the optimal response.

3.4.3.7. Effector Gene Dosing

The amount of effector gene being expressed within a host production system directly affects the response it elicits (Hansen et al., 2017), a concept that has been highlighted in some previously reported studies (Davis et al., 2000; Johari et al., 2015; Tastanova et al., 2016). This may mean the results, trends and effects reported may have been different if the dose of effector gene used had been altered or titrated during an investigation (Hansen et al., 2017). Differing effects are seen, for example, in a study by Davis and colleagues, overexpression of PDI at a low level was found to have no effect on fusion protein expression, but at higher levels it became detrimental and caused a reduction in titre (Davis et al., 2000). There is likely to be an optimal effector gene dose, as found for YY1 overexpression (Tastanova et al., 2016), and a titration should be incorporated into all studies. Gene titration could be achieved using synthetic promoters to alter gene transcription level or, more simply, by altering the level of gene

transfected, with filler, non-coding DNA used to maintain amount of DNA transfected (Brown and James, 2016; Estes et al., 2015; Hansen et al., 2017; Tastanova et al., 2016).

3.4.4. Synthetic Biology Paradigm

These approaches to engineering the CHO cell host machinery, altering and exploiting the normal cellular biology using genetic components, to improve its manufacturing capabilities can be identified and associated with the growing field of synthetic biology. Synthetic biology, is a developing area that combines biology and engineering to alter biological-based systems to undertake functions which do not exist *per se* in nature (Serrano, 2007). The field applies design, development and manufacture of novel parts (characterised genetic components), devices and systems (assembly of combinations of components into pathways) along with remodelling current biological set-ups already existing in nature, in order to generate useful functions and products (Benner and Sismour, 2005; Ellis et al., 2011; Kitney and Freemont, 2012; Serrano, 2007). There is a great level of unpredictability, inconsistency, variability, and complexity built into this synthetic biology paradigm (Kwok, 2010). To try and predict success in genetically engineered systems, the outcome of expressing novel components in host cells needs to be somewhat known prior to implementation. How parts combine together within pathways and systems, the effect of expression from multiple components simultaneously, and a reliable prediction of optimal relative ratios of multiple parts all need to be taken into consideration (Andrianantoandro et al., 2006).

3.5. Conclusion

There are many examples of engineering strategies undertaken to modify the CHO cell host, with just some of these highlighted within this chapter. As shown towards the end of the previous chapter (**Section 2.4**) there is a wealth of cellular components which function in relation to the process of producing and secreting a recombinant product. Engineering investigations, to date, have perhaps only scratched the surface in terms of CHO cell engineering. The advancements in omics level data, and increased knowledge and understanding of the CHO cell

host can be exploited to further engineer the CHO cell factory for recombinant protein production.

The outcomes from the literature demonstrate several relevant points that should be considered when undertaking CHO cell engineering and developing a CHO cell engineering platform, these include:

- i. Product specific solutions are required. Different products have resulted in different effects from particular engineering strategies. For example, as demonstrated in a study with a set of eight mAbs, a range of cell engineering solutions were demonstrated to have different effects (Pybus et al., 2013). This is likely to become a greater requirement with the increase in the number of engineered product variants, the CHO cell host may well need to be engineered to accommodate these developments. “Synthetic products” will likely benefit from “designer synthetic cell factories”.
- ii. The level of the engineering strategy applied is important. Very few studies have integrated different doses of genetic engineering strategies, but it has been demonstrated to be a relevant engineering criterion in studies by Tastanova and Johari (Johari et al., 2015; Tastanova et al., 2016). The level of gene expression will have a direct effect on its response.
- iii. The added benefit of multiplexing engineering strategies. When engineering a multi-component cellular system, where processes are taking place simultaneously and in association, it would be of benefit for many parts of the cell to be engineered. Combining strategies is likely to have significantly additional potential benefit (Cain et al., 2013; Johari et al., 2015; Peng and Fussenegger, 2009) but introduce added complexity and number of engineering strategies to test.

These requirements of CHO cell engineering can begin to be combined and addressed through the implementation of a screening tool – to test and compare many strategies at once, increasing the throughput and capability to test for an optimal solution (Hansen et al., 2017; Johari et al., 2015; Pybus et al., 2013). To date, the majority of published engineering strategies have been made with fairly low throughput and lengthy duration SGE systems. A high-throughput TGE screening platform has been highlighted as a useful tool to implement,

overcoming some of the current limitations in CHO cell engineering, and useful in finding solutions more rapidly. A screening system concept is well established in biopharmaceutical stable cell line generation, to identify optimal producing clonal cell lines (**Section 2.3.4**). This paradigm, along with that derived from the area of synthetic biology, can be combined to try and harness the potential of CHO cell engineering. A cell engineering screening process will be particularly useful for DTE engineered products, for finding optimal solutions potentially beyond that capable from screening for natural heterogeneity alone. The development of a high-throughput transient transfection platform, aimed at CHO cell engineering, is described in **Chapter 5**.

4. Materials and Methods

This chapter describes the technical approaches, methods, and materials used to generate the results that are described and explained in the subsequent chapters.

4.1. Plasmid DNA Preparation and Cloning

4.1.1. DNA Ligation

All ligations were performed using the NEB Quick Ligation™ Kit (New England Biolabs, Hitchin, UK), following the manufacturer's instructions. Insert and backbone vector fragments were combined in molar ratios of 6:1 and 3:1 of insert to backbone DNA with quick ligase and reaction buffer. The mixture was incubated for 5 minutes for the DNA fragments to join. A negative control, of just backbone vector without insert, was used to measure the amount of background colonies formed from self-ligation of the backbone. The ligation mixture was either stored at -20 °C or immediately used in a transformation reaction, outlined below.

4.1.2. Transformation and Amplification of Plasmid DNA

Library efficiency DH5α competent *E. coli* cells (Invitrogen, Thermo Fisher Scientific, Loughborough, UK) were used for the amplification of plasmid DNA, following the manufacturer's protocol. For transformation, competent cells were thawed on ice in 50 µL aliquots, mixed with 1 ng of plasmid DNA and incubated on ice for 30 minutes. The mixture was heat shocked in a water bath at 42°C for 45 seconds, then immediately incubated on ice for two minutes. Cells were diluted 1 in 20 with LB broth (Fisher Scientific, Thermo Fisher Scientific) and incubated at 37°C, 350 rpm shaking for one hour. 100 µL of cells were spread

onto LB agar plates (Fisher Scientific, Thermo Fisher Scientific), supplemented with the appropriate antibiotic, in all cases herein, either 100 µg/mL ampicillin or 50 µg mL⁻¹ kanamycin (Sigma-Aldrich, Dorset, UK), depending on the plasmid being amplified. Plates were inverted and incubated at 37°C for ~16 hours for colonies to form.

Individual colonies were picked to inoculate 5 mL starter cultures of LB broth supplemented with antibiotic and were incubated at 37°C, at 200 rpm for 8 hours. 100 µL of starter culture was then used to inoculate 100 mL of antibiotic-supplemented LB broth and the culture incubated at 37°C, 200 rpm for ~16 hours. After incubation, cells were harvested by centrifugation at 6000 x g for 15 minutes. The plasmid DNA was either extracted immediately from the cell pellet or pellets were stored at -20°C for subsequent extraction.

For long-term storage of plasmids, frozen glycerol stocks of transformed cells were generated. Bacteria were grown overnight in LB broth supplemented with antibiotic, and 500 µL of the culture mixed with 200 µL of 100% glycerol (Fisher Scientific, Thermo Fisher Scientific) and stored in screw-cap CryoTube™ vials (Sigma-Aldrich) at -80°C. To recover bacteria, a sterile loop was used to streak frozen cells onto an agar plate containing the appropriate antibiotic and incubated overnight at 37°C.

4.1.3. Plasmid Isolation and Purification

A mini- or a maxi-prep kit (Qiagen, Manchester, UK), was used to extract and purify plasmid DNA from *E. coli* cells, following the manufacturer's protocol. Alkaline lysis was used to break open the bacterial cells, the majority of cellular debris was precipitated then removed by centrifugation or filtration. The cleared lysate was added to an anion exchange resin tip or spin column which bound and isolated plasmid DNA, using low-salt and pH conditions. This was washed with medium salt buffer to further remove impurities, such as RNA and protein; eluted using a high-salt buffer and precipitated using isopropanol. Purified plasmid DNA was resuspended into either nuclease-free H₂O (Qiagen) or elution buffer (buffer EB) (Qiagen) and stored at -20°C.

4.1.4. Quantification of Plasmid DNA

A NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific) was used to determine the concentration and purity of plasmid DNA. Plasmid DNA was measured in triplicate and the average concentration used. To ensure the plasmid DNA was of high quality, DNA was only used that had a 260:280 nm absorbance ratio of 1.8-1.9 and a 260:230 nm ratio of 2-2.2, DNA with ratios outside of these ranges could indicate the presence of contaminants, such as RNA or leftover reagents from the purification process. DNA was diluted in nuclease-free H₂O to the concentration required for CHO cell transfections. The resulting concentration of the diluted DNA was confirmed using the NanoDrop™.

4.1.5. Restriction Digestion of Plasmid DNA

Restriction enzymes cut at sequence specific sites, and can be used to confirm the presence of an insert in a vector – a diagnostic digest. For diagnostic digests 1 µg of plasmid DNA was digested using 1 µL of NEB High Fidelity (HF®) restriction enzyme (New England Biolabs), 5 µL of CutSmart™ buffer (New England Biolabs) in a total reaction volume of 50 µL with nuclease-free H₂O. The reaction was incubated at 37°C for 15 to 60 minutes before being inactivated by the addition of 10 µL 6X purple gel loading dye (New England Biolabs). Double digestions were prepared using the same method, but with 1 µL of each HF® restriction enzyme included. Digests used for plasmid cloning were scaled up so 3 µg of plasmid DNA was digested to ensure sufficient DNA was obtained, the amount of restriction enzyme was increased accordingly, with 1 µL of restriction enzyme per 1 µg of DNA.

4.1.6. Gel Electrophoresis

Gel electrophoresis separates DNA based on size, and enables the visualisation and isolation of DNA fragments. 0.8% (w/v) agarose gels were made by dissolving 0.8 g of agarose (Fisher Scientific, Thermo Fisher Scientific) in 100 mL 1X TAE buffer (Sigma-Aldrich), 10 µL of 10000X SYBR safe dye (Invitrogen, Thermo Fisher Scientific) was added to allow the DNA to be visualised under blue or UV light. DNA was combined with 6X loading dye before being loaded onto a

gel; approximately 100-2000 ng of DNA in volumes between 20-60 μL , depending on the nature of the experiment, were loaded. Hyperladder 1 (Bioline, London, UK) was included on all agarose gels as a molecular weight marker. To separate DNA fragments, agarose gels were run at a constant 90 V for at least 40 minutes, and were imaged using a UV light box or UV transilluminator.

4.1.7. Gel Extraction

Fragments generated from plasmid digestion, were separated using gel electrophoresis, and the required fragment cut out from the agarose gel using a scalpel. To isolate purified DNA, from the gel slice, a Qiagen gel extraction kit (Qiagen) was used, following the manufacturer's instructions. The gel slice was first dissolved then added to a spin column, which, under high salt conditions, binds nucleic acids. Impurities (e.g. agarose and salts) were removed by washing the column with ethanol-containing buffer before the DNA was eluted in nuclease-free H_2O . A NanoDrop™ spectrophotometer was used to quantify the resultant DNA concentration.

4.1.8. Sequence Verification

DNA sequencing, using specific primers positioned around the site of the insert, verified plasmid sequences and ligation reactions. Primers were identified using Vector NTI software (Thermo Fisher Scientific) from MedImmune's in house primer library (MedImmune) or designed using SnapGene software (SnapGene, Chicago, UK) and synthesised by Thermo Fisher Scientific. For sequencing, plasmid DNA was diluted to 100 ng μL^{-1} and primers to 10 μM , both in nuclease-free H_2O . Where possible, both forward and reverse primers covering the region of interest were used. DNA sequencing was performed by the University of Sheffield Core Sequencing Facility (Core Genomic Facility, University of Sheffield, Sheffield, UK) or by MedImmune's in-house sequencing service (MedImmune). To check the success of insertions into the 'expression-plasmid' the forward primer TCTCACCGTCCTTGACACGAAG and the reverse primer TTGTCCAAACTCATCAATGTATCTTAGGCC were used.

4.1.9. Commercial DNA Synthesis

DNA sequences for effector genes and fluorescent proteins were synthesised by GeneArt™ (Thermo Fisher Scientific). The gene sequences for the fluorescent proteins were sourced from Evrogen (Evrogen, Cambridge Bioscience) for TagBFP and from Clontech (Clontech, Takara Bio Europe, Saint-Germain-en-Laye, France) for eGFP and mCherry. The gene sequences for the effector genes were obtained from online resources (e.g. National Centre for Biotechnology Information (NCBI) database). All sequences were optimised for *Cricetulus griseus* (Chinese hamster) expression by GeneArt™. A Kozak sequence upstream and restriction enzyme sites, flanking the coding sequence, were added and the coding sequences checked so no additional sites of the enzymes used for cloning were present. All synthesised DNA was sequence-verified by GeneArt to ensure its identity.

4.2. CHO Cell Culture

4.2.1. CHO Cell Lines

MedImmune kindly provided two of their proprietary cell lines for this work (MedImmune, Cambridge, UK). A CHOK1 derived, parental host: MEDI-CHO, a suspension-adapted CHO cell line, not producing any recombinant product, and a derivative of the MEDI-CHO cell line, CHO-T2. This cell line stably expresses EBNA-1 and GS and was developed for transient transfections (Daramola et al., 2014).

4.2.2. Subculture

Cells were routinely grown at 37°C, in 5% CO₂ (v/v), shaking at 140 rpm in an Infors HT Multitron Pro incubator (Infors HT, Surrey, UK). MEDI-CHO and CHO-T2 cells were passaged every 3-4 days to maintain exponential growth, into fresh pre-warmed culture media: CD-CHO media (Gibco™, Invitrogen, Thermo Fisher Scientific) supplemented with 6 mM L-glutamine (Gibco™, Invitrogen, Thermo Fisher Scientific) for MEDI-CHO cells, and CD-CHO media supplemented with 25 µM MSX (Sigma-Aldrich) and 100 µg mL⁻¹ hygromycin-B (Invitrogen, Thermo

Fisher) for CHO-T2 cells. Cells were seeded at 2×10^5 cells mL^{-1} in vented, non-baffled Erlenmeyer flasks (Corning[®], Amsterdam, The Netherlands). Cell density and culture viability were assessed using a Vi-Cell XR automated cell counter and viability analyser (Beckman Coulter, High Wycombe, UK).

4.2.3. Cryopreservation and Cell Revival

To reduce the need for long-term culture, which introduces the potential of genetic drift, master and working cell banks were generated and stored long-term in liquid nitrogen (-195.79°C). Cell banks were generated from cells in mid-exponential phase (day four after passaging), with four passages between the master and working cell banks. Cells were centrifuged at 200 RCF and resuspended into cold cryopreservation media containing 92.5% CD-CHO and 7.5% dimethyl sulfoxide (DMSO) (Sigma-Aldrich). 1×10^7 cells per CyroTube[™] vials (Sigma-Aldrich) were slowly chilled to -80°C , over 24 hours, in a controlled manner using a Mr Frosty container (Thermo Fisher Scientific). CyroTube[™] vials were placed in liquid nitrogen tanks for prolonged storage.

To recover cells from long-term storage, vials were thawed in a 37°C water bath, cells were transferred to a 50 ml falcon tube before 42.5 mL of pre-warmed CD-CHO slowly added. Cells were then centrifuged at $200 \times g$ for 5 minutes; supernatant removed and pellets resuspended in 10 mL of culture media and cell density measured. Cells were seeded at 3×10^5 cells mL^{-1} and passaged 3 days later. Experiments were only performed on cultures between passages 4 and 16.

4.3. Transient Transfection

For transfection, MEDI-CHO or CHO-T2 cells were taken from early exponential phase; day three of culture, where viable cell density (VCD) was approximately 2×10^6 cells mL^{-1} and where viability was $>95\%$. MSX and hygromycin-B selection were removed from CHO-T2 cultures two passages prior to transfection to prevent negative effects on transfection efficiency. The following protocol describes the electroporation and static post-electroporation culture for MEDI-CHO cells, the alterations for CHO-T2 post-electroporation culture are then

described, but the electroporation process itself is identical. A protocol for Lipofection-based transfection, used for MEDI-CHO cells, is also included.

4.3.1. Plate-based Electroporation for MEDI-CHO Cells

Electroporation (Nucleofection) was performed using the Amaxa™ Nucleofector™ system (Lonza, Basel, Switzerland), with the Nucleofector™ 96-well Shuttle™ add-on device (Lonza), the protocol being based on the manufacturer's instructions (Lonza, 2009). Prior to MEDI-CHO cell transfection, 500 µL of culture media (CD-CHO, supplemented with 6 mM L-glutamine) was transferred to each well of a flat-bottomed, sterile 24-well plate and 100 µL the outer edge wells (to reduce an evaporation effect) of 96-well microtitre culture plates (Nunc™, Thermo Fisher Scientific). These were transferred to a static humidified incubator to equilibrate to 37°C and 5% CO₂.

Prior to transfection MEDI-CHO cells were resuspended in nucleofection solution; SG Cell Line 96-well Nucleofector™ Solution pre-mixed with Supplement (Lonza), which was combined at a 4.5:1 (solution: supplement) on the day of use. A sterile round-bottom 96-well plate was used to mix plasmid DNA with the nucleofection solution mix. Firstly, 7.5 µL of the nucleofection solution mix was added to separate wells in the round bottom plate, corresponding to the number and layout of wells being transfected. Plasmid DNA, suspended in nuclease-free H₂O, was then added to the nucleofection solution. Each well contained 1000 ng of DNA in total (unless otherwise specified). nuclease-free H₂O was used to make up the total volume in each well to 10 µL. To prepare the cells for transfection, cells were centrifuged at 130 x g for 5 minutes. The supernatant was completely removed, with a vacuum pump or careful pipetting. The cell pellet was resuspended in nucleofector solution at a concentration of 115.3×10^6 cells mL⁻¹ so there were 2.33×10^6 cells per 15 µL for each transfection. At this stage, a sample of the resuspended cells was used to check cell density and viability. To each well of the 96 well plate containing plasmid DNA, 15 µL of cell suspension was added and mixed. 20 µL of each DNA-cell mix was transferred to the 96-well Nucleocuvette™ plate (Lonza). To reduce electroporation error, bubbles were carefully removed using fine sterile needles (Microlance™, BD BioSciences, Oxford, UK). The plate was electroporated within the Nucleofector™ 96-well

Shuttle™ add-on device using programme FF-158 (unless otherwise specified). Immediately afterwards 80 µL of pre-warmed culture medium was added to all transfected wells. The viability from three randomly selected wells were assessed to provide an average post-transfection cell viability and VCD measurement. Cells were transferred to the 24-well microtitre plates, already containing culture media, to obtain a seeding density of approximately 0.35×10^6 cells mL⁻¹, with pre-warmed CD-CHO added so that the total volume per well was 600 µl. In most experiments, cells from these 24-well plates were used to seed 96-well plates, 100 µL of each culture was transferred into separate wells in the 96-well plate (excluding outer edge wells). Use of 96-well plates increased throughput and allowed the use of plate-based downstream titre and viability measurements. All plates were cultured at 37°C, 5% CO₂, 80% humidification in a static incubator for 72 hours, unless otherwise stated. Multiple 96-well plates were seeded in parallel for use in different assays (due to the low sample volume), including one plate used immediately for PrestoBlue® analysis (see **Section 4.4**).

4.3.2. Deep-Well Post-Electroporation Culture

For transfection of CHO-T2 cells, the previous electroporation protocol was followed, but with the following changes to allow the cells to be cultured for longer in deep well plates post transfection.

Prior to transfection, 525 µL of CD-CHO was transferred to each well of a sterile 96 deep-well (square, V-bottomed) plate (Geiner Bio-one, Stonehouse, UK). ‘Sandwich cover’ lids (Duetz, Enzyscreen B.V, Heemstede, Netherlands) were used to cover and seal the deep-well plates, reducing evaporation, cross-contamination, and to improve gas transfer. The plates were transferred to a shaking incubator (320 rpm, 25 mm throw) to equilibrate to 37°C, 85% humidification and 5% CO₂. A clamp system (Duetz, Enzyscreen B.V) was used to hold the deep well plates and lids in place; enabling high-speed, vigorous, orbital shaking, essential for efficient mixing and cell culture performance. The electroporation protocol was followed in the same way as MEDI-CHO cells. After electroporation and cell counting, the volume of cells required to seed at 0.4×10^6 cells mL⁻¹ was transferred to wells in the deep-well plates, then pre-warmed CD-CHO was added so that the total volume was 575 µl per well. Prior to incubation

100 μL from each well was transferred into a flat-bottomed 96-well plate (Nunc) for PrestoBlue[®] analysis (**Section 4.4**). The deep-well plates containing 475 μL of cell culture were then incubated for 5 days, using the same culture conditions as above, unless otherwise stated. At 72 hours culture media within the deep-well cultures was supplemented. To maintain a constant volume within each well, plates were centrifuged at 300 x g for five minutes then 10%, based on the initial culture volume, was removed (47.5 μL). It was replaced with the same volume (47.5 μL) of a mixture (1:1 ratio) of CHO CD EfficientFeed[™] A and B liquid feed supplements (Gibco[™], Thermo Fisher Scientific), with gentle pipetting to resuspend cells.

4.3.3. Lipofection

Chemical mediated transfection was performed using Lipofectamine[®] LTX with PLUS[™] reagent (Invitrogen, Thermo Fisher, Scientific), following the manufacturers protocols (Invitrogen, 2011; Invitrogen, 2013). Prior to transfection, MEDI-CHO cells were seeded at 4×10^5 cells mL^{-1} in 500 μL culture media in shallow 24-well culture plates (Nunc, Thermo Fisher Scientific), then incubated in a static incubator (37°C, 5% CO_2 , 80% humidification) for an hour. For every transfection, the ratio of plasmid DNA to Lipofectamine[®] used was 1:3 (w/v). Plasmid DNA (1 μg) was combined with PLUS[™] reagent (1 μL) and OptiMEM[™] I Reduced Serum Medium (Gibco[™], Thermo Fisher Scientific), making the volume 25 μL . Separately 3 μL Lipofectamine[®] LTX reagent was combined with 22 μL of OptiMEM[™]. The separate mixtures were combined, gently mixed and incubated at room temperature for five minutes for lipid (Lipofectamine[®])-DNA complexes to form; before the 50 μL volume was added and gently mixed to cells within the 24-well plate. Cultures were placed in a static incubator for 24-72 hours before cell viability and protein expression were assessed.

4.4. PrestoBlue[®] Cell Viability Assay

For high-throughput analysis of biomass, a plate-based PrestoBlue[®] assay was used, which utilises a resazurin-based solution and requires very small sample volumes. When added to live cells, the permeable PrestoBlue[®] viability reagent

(Invitrogen, Thermo Fisher Scientific) is converted to a fluorescent red colour by the reducing environment in a cell, this colour change is detectable on a fluorescence plate reader.

Immediately prior to use PrestoBlue mix was prepared at a ratio of 1:1 with PrestoBlue® and CD-CHO culture media, both at room temperature. Samples were either cultured in, or transferred to just before measuring, clear, flat-bottomed 96-well plates (Nunc, Thermo Fisher Scientific) at a volume of 100 µL per well. To account for background fluorescence three media-only samples were included on each assay plate. To each sample 22.2 µL of PrestoBlue® mix was added and immediately mixed using 700 rpm orbital shaking for 20 seconds. Plates were then incubated at 37°C for 35 minutes, then shaken again at 700 rpm for 20 seconds and the fluorescence measured using 540 nm excitation and 590 nm emission filters on a PHERAstar plus micro-plate reader (BMG Labtech, Aylesbury, UK). Prior to fluorescence measurements, gains were adjusted so a highly viable (and therefore fluorescent) sample was at 80% of the maximum signal, ensuring all samples were within the detectable range. When necessary, a cell concentration standard curve was included on assay plates; samples were treated the same as above but measured in-parallel using the trypan blue exclusion method on a Vi-Cell XR viability analyser to determine VCD (cells mL⁻¹). Data was analysed using PHERAstar analysis software (BMG Labtech), the average background media-only fluorescence on each assay plate, was calculated, to standardise values between assays. Normalised fluorescence was calculated by the following Equation (4.1) (an example is shown in **Table 4.1**):

$$\text{Normalised fluorescence} = \frac{\text{Sample fluorescence}}{\text{Media only fluorescence}} \quad (4.1)$$

Table 4.1: Example PrestoBlue Assay Analysis

All raw fluorescence values (top table) are divided by the average media only background fluorescence (in red) to give a relative ratio to the background fluorescence (bottom table). This makes the values comparable across assays and between time points.

Raw values	126610	125793	122927	112614	123603	109609	101876	107568	103070	116907	106317	91422
	133558	147280	161348	114314	131385	118606	115845	110890	114185	95678	114442	125662
	130005	121056	104831	116982	129798	105033	118698	108229	107360	108392	98077	89806
	129747	128861	100506	185084	93428	108281	123016	111707	81313	108340	99434	91028
	114476	121917	108455	102426	115354	100990	108216	109540	89446	103376	94365	85002
	76837	116396	97401	114334	111413	107710	119856	106437	85376	102566	92526	96259
	133684	127368	96703	96016	93925	90740	112274	106403	102978	87281	96945	92976
	118684	125811	94809	110314	93222	95936	53670	53388	53018			
Average background fluorescence:												
53359												
Ratio to background	2.373	2.357	2.304	2.111	2.316	2.054	1.909	2.016	1.932	2.191	1.992	1.713
	2.503	2.760	3.024	2.142	2.462	2.223	2.171	2.078	2.140	1.793	2.145	2.355
	2.436	2.269	1.965	2.192	2.433	1.968	2.225	2.028	2.012	2.031	1.838	1.683
	2.432	2.415	1.884	3.469	1.751	2.029	2.305	2.094	1.524	2.030	1.864	1.706
	2.145	2.285	2.033	1.920	2.162	1.893	2.028	2.053	1.676	1.937	1.769	1.593
	1.440	2.181	1.825	2.143	2.088	2.019	2.246	1.995	1.600	1.922	1.734	1.804
	2.505	2.387	1.812	1.799	1.760	1.701	2.104	1.994	1.930	1.636	1.817	1.742
	2.224	2.358	1.777	2.067	1.747	1.798	1.006	1.001	0.994			

4.5. Flow Cytometry

To measure cellular expression of fluorescent proteins, for example the transfection efficiency from green fluorescent protein (GFP) transfections, flow cytometry analysis was used. Living cells were analysed directly in their culture media. Different flow cytometers were used depending on the nature of the experiment. For all, the photomultiplier tube (PMT) settings were adjusted for optimal detection of CHO cells, based on their size, shape and autofluorescence. Cells were positioned based on size (using forward scatter (FSC)) and granularity (side scatter (SSC)) in the centre of the plot, with viable cells then gated to exclude dead cells or debris in further analysis. For most analyses, 10,000 viable cells (events) were measured for further fluorescence analysis. The peak of a negative, mock or non-transfected population of cells was positioned at 10^2 fluorescence units, unless otherwise stated, and used to determine background autofluorescence levels.

4.5.1. Attune Acoustic Focusing Cytometer

The Attune[®] Acoustic Focusing Cytometer (Applied Biosystems, Life Technologies, Thermo Fisher Scientific) was used to measure GFP expression. The optimised voltages (mV) used were 1100, 2800, 1000 2000, for FSC, SSC, BL1 and RL1, respectively. For large numbers of samples or those with low volumes (<250 μ L), an Attune[™] NxT Autosampler (Life Technologies, Thermo Fisher Scientific) was used; this analysed samples directly from a 96 multi-well plate enabling rapid, high-throughput analysis.

Intracellular GFP was excited using a 488 nm laser; with emitting fluorescence measured using a 530/30 bandpass filter (BL1). If viability was measured TO-PRO[®]-3 Iodide (Life Technologies, Thermo Fisher Scientific) stain was added. The stock solution (1 mM) was diluted 1:500 using PBS to a working concentration of 0.002 mM. When adding to a 96 well plate, the working concentration was further diluted one in four in PBS and 8 μ L added and mixed to each well immediately prior to flow cytometric analysis. TO-PRO[®]-3 Iodide was excited using the 633 nm laser; with emitting fluorescence measured using a 660/20 bandpass filter (RL1). The TO-PRO[®]-3 Iodide stain was stored in the dark and used until it became colourless. All data was analysed using Attune Cytometric Software (version 2) (Life Technologies, Thermo Fisher Scientific).

4.5.2. BD LSR Fortessa Flow Cytometer

For triple fluorescent protein transfection experiments, fluorescent expression was measured after 24 hours. To ensure sufficient cells for analysis after 24 hours, cells were seeded post-transfection at approximately 1×10^6 cells mL^{-1} in 96-well culture plates.

The BD LSR Fortessa[™] flow cytometer (BD BioSciences) was used to measure the co-expression of multiple fluorescent protein expression, as it is suitable for multicolour applications. The lasers (colours) and bandpass filter sets (channels) used were: 405 nm (violet), 488 nm (blue) and 532 nm (green) lasers; 450/50 (BV421), 525/50 (FITC), 610/20 (PE-Texas Red), for TagBFP, eGFP and mCherry, respectively.

To exclude dead cells from analysis, a DAPI (4',6-Diamidino-2-Phenylindole) (BioLegend®, San Diego, USA) DNA stain, which is not taken up by living cells, was added to cells immediately prior to flow cytometric analysis. Stock solution of DAPI (5 µg mL⁻¹) was diluted 1 in 500 in PBS before 100 µL was added to each well. To measure DAPI staining a 355 nm (UV) laser and a 450/40 (DAPI) bandpass filter was used for excitation and emission, respectively. Cells that were positive for DAPI staining were excluded from further analysis.

The optimised voltages (mV) used for CHO cells were 20, 336, 148, 143, 190, 203, for FSC, SSC, UV, BV421, FITC and PE Texas red respectively. Samples were measured using a high-throughput autosampler analysing samples directly from a 96-well plate, with 200 µL of cells and DAPI-PBS solution per well. Data collected using BD FACSDiva™ software (BD BioSciences) was analysed using FlowJo® Software (FlowJo LLC, USA). Minimal amounts of compensation were applied to correct for small amounts of spectral overlap. In **Table 4.2** there is an example compensation matrix applied, which shows how much spill over from each fluorophore is detected in the other channels.

Table 4.2: Example Compensation Matrix Table for Tag-BFP, mCherry and eGFP Detection and Analysis

		Target/Detector			
		BV421	FITC	PE-Texas-Red	DAPI
Emission Source	BV421	100	0.3	0	9.3
	FITC	0.5	100	0	0
	PE-Texas-Red	0.1	0	100	0
	DAPI	0	0	0	100

4.6. Quantification of Volumetric Titre

Supernatant samples were isolated using centrifugation at 1000 x g for 5 minutes. The supernatants were transferred into a new tube, and if not used immediately, stored at -20°C for short term storage.

4.6.1. Quantification of IgG Titre

Recombinant IgG levels were measured using a FastELISA[®] Human IgG Quantification Kit (RD-Biotech, 2BScientific, Oxford, UK), following the manufacturer's instructions. Briefly, all samples were diluted at least 1 in 2 using the supplied diluent before being transferred to the pre-coated 96-well plate, followed by incubation with HRP (horseradish peroxidase) conjugate. Non-bound HRP was removed through washing steps, then the HRP substrate was added and incubated before addition of the stop solution to end the reaction. End point absorbance was measured at 450 nm and 620 nm using a PowerWave HT Microplate Spectrophotometer plate reader (BioTek, Swindon, UK) and KC4 software (version 3.1) (BioTek). On each plate, provided standards of known concentration, were included. Cubic spline standard curves, using the protein standards, were generated on GraphPad Prism software (version 7) (GraphPad Software, Inc., San Diego, USA) and used to interpolate unknown recombinant protein concentrations. An example is shown in **Figure 4.1**.

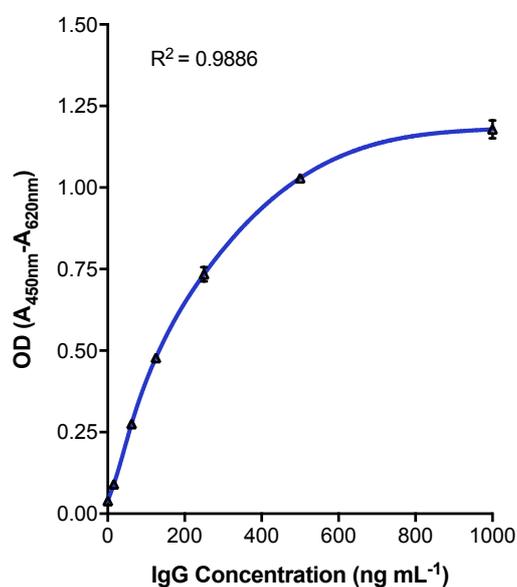


Figure 4.1: Example Fast-ELISA Assay Standard Curve

Standards of known protein concentration were included on each FastELISA[®] assay and their absorbance at 450 nm used to generate cubic spline standard curves; allowing concentrations of unknown IgG samples to be inferred. Where possible, standards were included in at least duplicate. Data shown from duplicate standards, with error bars showing standard deviation.

4.6.2. Quantification of Secreted Embryonic Alkaline Phosphatase Titre

Recombinant Secreted Embryonic Alkaline Phosphatase (SEAP) levels were measured using a Colorimetric SensoLyte® pNPP SEAP Gene Assay Kit (AnaSpec, Cambridge Bioscience, Cambridge, UK), following the manufacturer's instructions. This kit converts the enzymatic properties of SEAP proportionally into a measurable colour change. Briefly, all samples were diluted at least 1 in 2 using culture media, before 50 μL of each sample was transferred to a flat-bottomed 96-well plate (StarLab, Milton Keynes, UK), followed by the addition of 50 μL of pNPP substrate to each well. Kinetic absorbance readings, at 405 nm, were measured every 5 minutes for a total of 60 minutes, using a PowerWave HT Microplate Spectrophotometer plate reader (BioTek) and KC4 software (version 3.1) (BioTek). This ensured readings were within the linear range of detection. On each plate, a serial dilution of known SEAP concentrations was included. A standard curve using these known protein concentrations was generated on GraphPad Prism software (version 7) (GraphPad Software Inc.) and used to interpolate unknown SEAP concentrations. An example is shown in **Figure 4.2**.

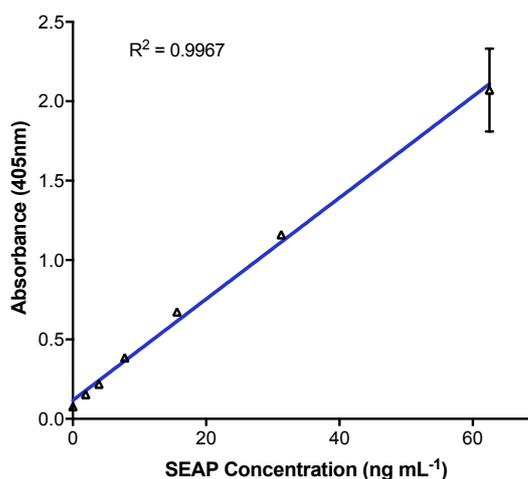


Figure 4.2: Example SEAP Assay Standard Curve

Standards of known concentration were included in each SEAP assay and their absorbance at 405 nm used to generate a standard curve; allowing concentrations of unknown SEAP samples to be inferred. Standards were included in duplicate. Data shown for duplicate standards, with error bars showing standard deviation.

4.7. Cell Culture Parameter Equations

4.7.1. Equations to Calculate Cell Growth

Cell specific growth rate, μ (day^{-1}), was calculated using the following equation:

$$\mu = \frac{\ln(N_2/N_1)}{\Delta t} \quad (4.2)$$

Integral of viable cell density (IVCD) was calculated using the following equation:

$$IVCD = \left(\frac{N_1 + N_2}{2} \times \Delta t \right) \quad (4.3)$$

For both equations N_1 and N_2 are the concentration of viable cells (cells mL^{-1}) at the first and second time points, respectively, and t represents time, in days.

4.7.2. Equations to Calculate Productivity

Specific Productivity, qP ($\text{pg cell}^{-1} \text{ day}^{-1}$) was calculated using the following equation:

$$qP = \left(\frac{T_2 - T_1}{(N_1 + N_2)/2} \right) t \quad (4.4)$$

Where N_1 and N_2 are the concentration of viable cells (cells mL^{-1}) at two different time points (first and second time points, respectively), t represents the change in time, in days, and T_1 and T_2 are the amounts of recombinant product (titre) in mg L^{-1} at the first and second time points, respectively.

4.8. Statistical Analysis

All calculations for mean, median, standard deviation, standard error of the mean (SEM), 95% confidence interval and coefficient of variation were performed using Microsoft® Excel®, Graphpad Prism software or R. Box plots and Pearson's correlation in **Section 5.3.6** were generated within R. In **Section 7.3.6.2**, for wide-scale statistical comparisons, a one-way analysis of variance (ANOVA) with a multiple comparison Dunnett's test was performed within Graphpad Prism. This was used to statistically compare each transfection-sample mean to a single control-transfection mean.

4.8.1. Coefficient of Variation

Coefficient of variation (CV), to provide an indication of relative variation, was calculated using the following equation:

$$CV (\%) = \left(\frac{\text{Standard Deviation}}{\text{Mean}} \right) \times 100 \quad (4.5)$$

In **Section 5.3.6** investigations into platform variation were made; including an investigation into how relative variation (CV) altered by changing the level of replication within a transfection run and to decide a suitable number of transfection replicates to use. Growth rates were calculated, using Equation (4.2), for each well of nine separate post-transfection culture plates, resulting in approximately 864 growth rates in total. Growth rates were randomly combined to represent different numbers of transfection replicates (singular and duplicates through to up to six replicate wells), with each post-transfection culture plate treated independently. For each resulting combined grouping of replicates an average growth rate was calculated and from these average growth rates a mean and standard deviation was calculated for each number of replicates (one to six). Relative variation (CV) was then calculated, using Equation (4.5), for each replicate number and from these an average CV estimated. Combining wells into groups of replicates was a random process, so the process, and calculation of subsequent statistics, was repeated a hundred times to account for any error within the randomising replicates process. An average CV and level of error overall was calculated for each number of replicates (**Figure 5.17**). When repeating the random process for singular transfections (replicate of one) the mean, standard deviation and resulting CV remain constant as no wells were combined together.

In **Section 6.3.4** and **Section 6.3.5** investigations into fluorescent protein expression variation were made. For each fluorescent protein (colour), the mean relative fluorescence of the population was calculated from the normalised fluorescence values for each cell. The standard deviation, in relative fluorescence, for each fluorescent protein was also calculated from the normalised fluorescence values for each cell, to provide an estimate of the level of variation between cells within the population. These attributes were used to

calculate the spread of relative fluorescence levels for each fluorescent protein for the whole cell population, i.e. the CV (using Equation (4.5). Total cellular fluorescence was accounted for when investigating the variation in expression; when calculating all averages and variance weighted statistics, established from the total cellular fluorescence of each cell, were calculated within R.

4.8.2. Design of Experiments

To simultaneously investigate multiple parameters within an experiment, DoE Design Expert 10.0.6.0 software (Stat-Ease, Minneapolis, USA) was used. The software enabled mathematical modelling of output variables within two-level factorial designs (used to identify both single and combinatorial parameters) and response surface models (used to investigate combinatorial effects and to predict optimal solutions between multiple parameters). The software performed statistical analysis on the generated mathematical models using ANOVA to identify any significant experimental parameters (factors). In addition, predictive models of output variables, in terms of any experimental factors which were identified as having a significant impact, were produced. Further information on the use of DoE can be found in **Section 5.3.2.2**, **Section 7.2** and **Section 7.3.5**.

5. Development of a High-Throughput Transient Transfection Screening Platform

The development of a high-throughput transient transfection setup, aimed to simultaneously screen the functionality of multiple effector genes, is described in this chapter. To facilitate standardised and equivalent transfection and expression of effector genes within the screen, an ‘Expression-Plasmid’ was designed and constructed. The parameters of electroporation programme, cell number and amount of plasmid DNA to use within the micro-scale transfection platform were optimised. This resulted in high, immediate gene expression for a large proportion of the population, with minimal cytotoxicity. Incorporation of 96-well plate-based formats allowed for increased throughput, and the same format used throughout the platform. Variation, inherent within the micro-scale setup, can be reduced through the use of triplicate transfections and inter-transfection normalisation controls. In summary, the resulting platform enables hundreds of simultaneous, micro-scale, high-efficiency, standardised, transient co-transfections of CHO cell cultures, each with multiple plasmids.

Acknowledgments

The deep-well plate-based shaking culture setup was developed by Dr Joseph Cartwright (University of Sheffield). The R script to analyse variation within the transfection platform was undertaken in R by Dr Joseph Cartwright.

5.1. Introduction

There is a wealth of evidence in the literature, as highlighted in **Chapter 3**, demonstrating that cell engineering strategies have much potential benefit to improve the production capabilities of CHO cell factories. The overexpression of

genetic components is a well-established approach. To-date, the majority of published examples describe strategies involving the alteration of one or two different components. The approaches are low throughput, which is noteworthy considering the wealth of potential targets that include approximately 24,000 genes in the *C. griseus* genome (Lewis et al., 2013). Amongst these studies there are contradictory and variable findings, with different studies, demonstrating varied effects from the overexpression of the same effector component on the production of different recombinant proteins. An example of this irregularity has been recently highlighted in a review by Hansen *et al.*, highlighting differences from several studies of PDI overexpression (Hansen et al., 2017). It demonstrates differing outcomes upon production of different recombinant products and therefore specific engineering solutions are likely to be needed for different recombinant products. A recent study, with several mAb products showed that even related molecules of the same subclass can have product-specific differences. The different genetic engineering strategies tested had varied efficacy upon recombinant protein production for the selection of related mAbs (Pybus et al., 2013). The cellular constraints arising in cells producing and secreting recombinant products varies between molecules and cell hosts. Bespoke, product-specific solutions will likely to be more optimal and of more benefit.

In most studies reported, singular engineering solutions were implemented; there are few examples of simultaneous overexpression of multiple components beneficially generating synergistic effects (Cain et al., 2013; Le Fourn et al., 2014; Mohan and Lee, 2010; Peng and Fussenegger, 2009). An engineering solution integrating multiple components, although more complex, is anticipated to be of more potential benefit, in terms of enhancing productivity, and potentially able to alleviate more than one cellular constraint (Hansen et al., 2017; Pybus et al., 2013; Xiao et al., 2014). When altering multiple components simultaneously, their proportions relative to each other is important and should be part of investigations to try and find an optimal and effective balance between components (Hansen et al., 2017; Xiao et al., 2014). These types of solutions appear to have the most potential benefit for DTE proteins, where there are likely to be more cellular constraints to overcome (Johari et al., 2015).

Compared to most studies, there was an increase in throughput in the range of engineering approaches simultaneously tested in a recent study by Johari *et al* (Johari et al., 2015). Most studies report engineering a few components at most, whereas in this work twelve different engineering solutions, including five different genetic effectors, were tested in-parallel in a TGE approach. Despite the success of the study in greatly improving the productivity of a DTE fusion protein, it was emphasised that progression of high-throughput approaches would be beneficial in finding advantageous effects to improve CHO cell factories. To harness and progress from the findings described in the literature, it is clear that an approach utilising a platform rapidly assessing the impact of genetic components and combinations of components on a particular recombinant product would be advantageous in designing and engineering superior CHO cell factories (Johari et al., 2015; Pybus et al., 2013). To enable multi-component synthetic DNA solutions, two fundamental design criteria need to be addressed; selection of the specific and minimal combination of components that operate in synergy to achieve the desired objective and determination of their optimal relative functional stoichiometry.

There is a clear requirement to screen for individual genetic component and multi-component functionality. There are many potential beneficial components, which could be overexpressed at different levels and in different combinations, and are likely to be context and product specific. With this in mind, the work in this chapter describes the development of a high-throughput transient transfection platform. An approach of this nature was recently demonstrated by Hansen and colleagues (Hansen et al., 2015). They describe the development and potential of a 96-well plate-based, chemical transfection platform and exemplify its use to investigate genetic engineering approaches on the effect of producing two recombinant non-mAb products of human origin. This recent example helps to demonstrate the relevance, flexibility and potential gain derived from high-throughput screening approaches. One of the primary purposes for the high-throughput screen being developed and described here is to screen for the functionality of a library of effector genes for their potential improvement on the production of a DTE-mAb of industrial interest. As combinatorial effects are likely to be of additional benefit,

the platform needs be a tool that can assess and optimise combinations of genetic components simultaneously.

There are several design criteria to incorporate into the high-throughput platform being developed. Particular attributes are desirable to enable successful use of the transfection setup in model-based, multiple-component engineering with the aim to improve the production capability of the CHO cell factory. The approach developed needs to enable simultaneous investigation of both individual and combinations of components. A summary of the desirable platform attributes is outlined below:

- i. Foremost, the platform needs to incorporate simultaneous comparisons of many genetic parts in-parallel. This will facilitate screening the effect and functionality of many numerous parts at once, a valuable trait to increase the scope of parts that can be assessed.
- ii. To enable functionality testing and comparisons between many genetic parts, the platform should harness synthetic biology principles of 'plug and play' and modularity, to allow flexibility and relative comparisons in the parts tested. Different parts should be interchangeable and expressed in a consistent and standardised way.
- iii. To expand and increase platform throughput, the setup needs to harness multi-well plate-based approaches. The use of compatible 96-well plate-based approaches throughout would increase the number of simultaneous comparisons that can be assessed. Micro-scale approaches of this nature, will incorporate low volume transfections and cultures have the potential of increasing inherent variation within the setup, this will need to be taken into account and reduced if possible.
- iv. The transfection setup needs to incorporate rapid assessments of the effect from gene overexpression, enabling increased platform throughput. To help facilitate this, the transfection setup should robustly and reliably generate high, immediate, titratable expression of the introduced genes to the whole population (i.e. high transfection efficiency). This will increase the reliability and reduce the variability of multiple gene introduction and ensure effective gene expression. This scenario needs to be balanced with

not causing too high toxicity from the transfection, which could introduce unwanted off-targets effects.

- v. The method needs to include approaches to introduce co-ordinated expression of multiple components, to systematically investigate potential combinatorial effects and enable potential optimisation of the ratio of beneficial combinations.

In this chapter, the process undertaken to develop and optimise the high-throughput transient expression platform to create a synthetic design space for CHO cell engineering is outlined. As highlighted in **Section 2.3.2** different transfection methods to introduce DNA into cells are available, here a 96-well plate-based electroporation setup was used. A low-volume, micro-scale transfection approach increased the throughput enabling a rapid screening platform. Parallel transfections like this have the benefit of reducing the associated resource, labour and time intensity. The parameters in the setup (e.g. electroporation programme, and cell number and amount of DNA transfected) were altered to achieve the desired design criteria and optimal outputs to enable an effective screening capability. A DoE methodology was partly incorporated into the optimisation process, for a systematic approach to optimising more than one factor at a time. An approach like this has been successfully used for transfection optimisation previously, particularly in relation to optimising PEI based transfection methodologies (Daramola et al., 2014; Mozley et al., 2014; Thompson et al., 2012). To implement synthetic biology concepts, a universal 'Expression-Plasmid' was generated and constructed. This allows for standardised, consistent and precise levels of expression and for interchangeability and alteration in the level of parts being tested. An empty-'Expression-Plasmid' can be utilised as 'filler' plasmid, to keep DNA loads constant between transfections. The product gene of interest can also be interchanged for a different product gene within a screen. Plate-based approaches were developed throughout the platform – for the transfection process, cell culture and output response assays. Multi-well micro-plate based static culture methods permitted high-throughput, low volume culture; although VCDs and culture durations were limited, due to the static, batch culture-mode with restricted surface area for cell growth. Implementing a more optimal culture

method, utilising deep-well plates and a shaking culture mode overcomes these limitations. The assays to measure post-transfection culture performance and effector functionality needed to be suitable and compatible with the low volume and cell number involved. Partly due to the inherent low volumes throughout the platform, multiple, identical transfections were undertaken to investigate the inherent platform variability. This helped resolve how many replicates were required to reduce some of the variation within the platform.

5.2. Experimental Approach

A multi-well micro-plate based transfection setup was considered most desirable to facilitate high-throughput in the number of side-by side transfections that could be performed. There would be reduction in resource requirements (for example, less plasmid DNA per transfection) and in labour and time intensity required (for example, increased in-parallel setup, reduced handling times). However, limitations associated with small volumes, low cell numbers and many samples to measure simultaneously would be present.

The transfection method used was electroporation. This uses an electric field to alter the permeability of cell membranes, enabling DNA to be introduced into cells. It is an industrially relevant technique, commonly used for stable cell line generation (MedImmune, personal communication). A Nucleofection electroporation approach (Lonza), with a 96-well shuttle device™ was used to deliver DNA directly to the nuclei of cells. Up to 96 independent transfections (each in a separate well) could be undertaken in-parallel (with the electroporation process taking under five minutes). Electroporation has several benefits for the high-throughput platform. Due to the nature of the method (Nucleofection) – DNA enters cells nuclei directly meaning there is immediate capability for gene expression. This is a very desirable trait for rapid assessments of gene expression. The approach has been demonstrated to introduce DNA with less variation than for a common alternative method, lipofection (Davies et al., 2013; Pichler et al., 2011), advantageous for micro-scale high-throughput transfection. Additionally, electroporation is more efficient at delivering DNA and results in higher levels of transient protein production (Davies et al., 2013; Tabar et al., 2015). Additionally, DNA delivery has been shown to be uniform throughout the

whole transfected population (Pichler et al., 2011). Electroporation appears to rapidly, efficiently and consistently deliver genes into cells.

To optimise Nucleofection for the MEDI-CHO cell line and the design criteria of interest here (rapid, high efficiency transfection), the expression of GFP was primarily used. This is a commonly used reporter for gene expression, easily measurable due to its fluorescent nature using flow cytometry. An example of flow cytometry GFP analysis is shown in **Figure 5.1**. The proportion of the population receiving the transfected DNA, the transfection efficiency, can be easily inferred from measuring GFP expression. For many initial experiments, a small pmaxGFP plasmid (Lonza; ~3.5kb; Lonza Amaxa, Nucleofector technology flyer), provided with Nucleofector kits was used. With subsequent platform development, other plasmids, which encoded for recombinant mAb genes, were introduced into cells. These plasmids were larger and were found to cause more substantial and undesirable effects on culture viability (**Figure 5.6**). Further development, in relation to the number and cells and amount of DNA being transfected, was undertaken to improve the output from transfections with more appropriate plasmids of interest (**Section 5.3.2.2**).

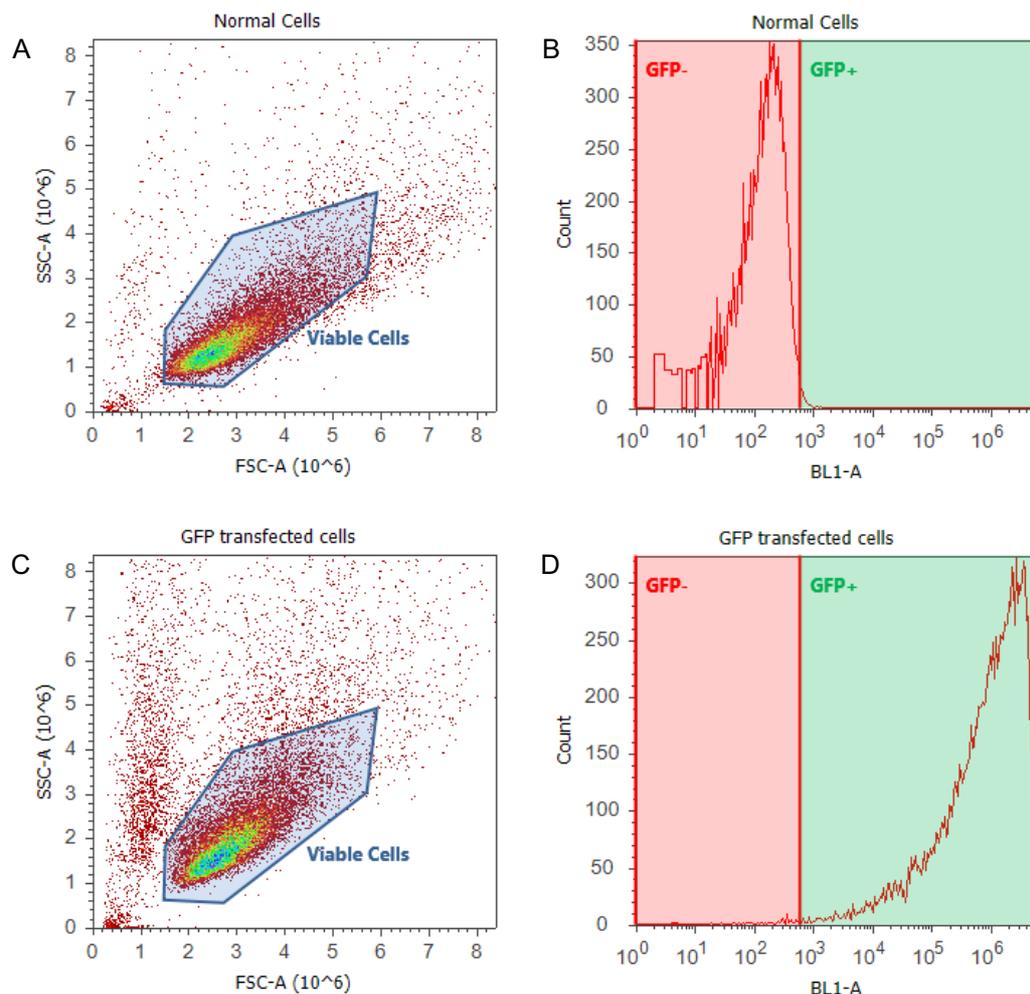


Figure 5.1: Determination of GFP Expression using Attune Flow Cytometry Detection

(A, C) A density plot, with events positioned based on their size (FSC-A) and side scatter (SSC-A); with a 'Viable Cells' gate positioned around viable cells, removing debris and dead cells from further analysis. (B, D) Analysis of the 'Viable Cells' population using a daughter histogram plot, representing the frequency (count) of relative fluorescence the events are emitting (BL1-A). A bi-marker gate using a population not expressing GFP is positioned so that 99% of this negative population falls in GFP- gate (as seen in (B)). A GFP positive population can then be quantitatively inferred using a GFP+ gate (as seen in (D)). Statistics are exported, providing percentages and median fluorescence intensity (MFI) of events within specific gates.

5.3. Results

5.3.1. Plasmid Development to Enable Standardised Gene Expression

The transfection platform needs to incorporate expression of numerous different genetic components, both on their own and in combinations with each other. The amount of DNA needs to be maintained between transfections to keep a consistent transfection efficiency and enable comparison of data from different transfections. A plasmid, designated throughout as 'Expression-Plasmid', was developed for standardised expression of synthetic genetic components.

The structure of the 'Expression-Plasmid' is outlined in **Figure 5.2**; with various design criteria incorporated. AgeI and PstI (SbfI) restriction sites, separated by 16 base pairs (bp) of spacer DNA, were included for directional insertion of different effector gene open reading frames (ORFs). The cloning site was designed to be compatible with a separate multigene plasmid system (under development in the DCJ laboratory at The University of Sheffield). Any synthetic effector gene parts were compatible and could be straightforwardly cloned in both systems. Unique restriction sites (EcoRI, HindIII, KpnI and NheI) were included to provide downstream flexibility, for example to clone in a mammalian selection marker enabling the plasmid to be used for SGE. This is not included at this stage as the system being developed is primarily aimed towards TGE. The additional plasmid size and expression of another gene would respectively generate additional negative transfection effects and biosynthetic burden within cells.

The components on the plasmid (for example promoter and untranslated regions (UTRs)), controlling expression of the inserted gene, are typical of those used for recombinant protein expression in CHO cells. Using universal parts on all plasmids being transfected ensures consistent, unbiased expression of different genes. For example, using the same CMV promoter throughout expression of all recombinant genes reduces promoter interference and bias, something that would be prevalent if different promoters were used (Brown and James, 2016).

An OriP DNA element was included, this enhances gene transcription and plasmid retention post-transfection; and the plasmid was compatible with longer

duration TGE in MedImmune's transient host cell line (Daramola et al., 2014). The OriP DNA element was flanked by SapI restriction sites, for straight-forward modification to remove the OriP from the plasmid if desired. Sequences encoding ampicillin resistance (AmpR) and bacterial origin of replication (pUC) were included to enable replication and amplification of the plasmid within bacteria.

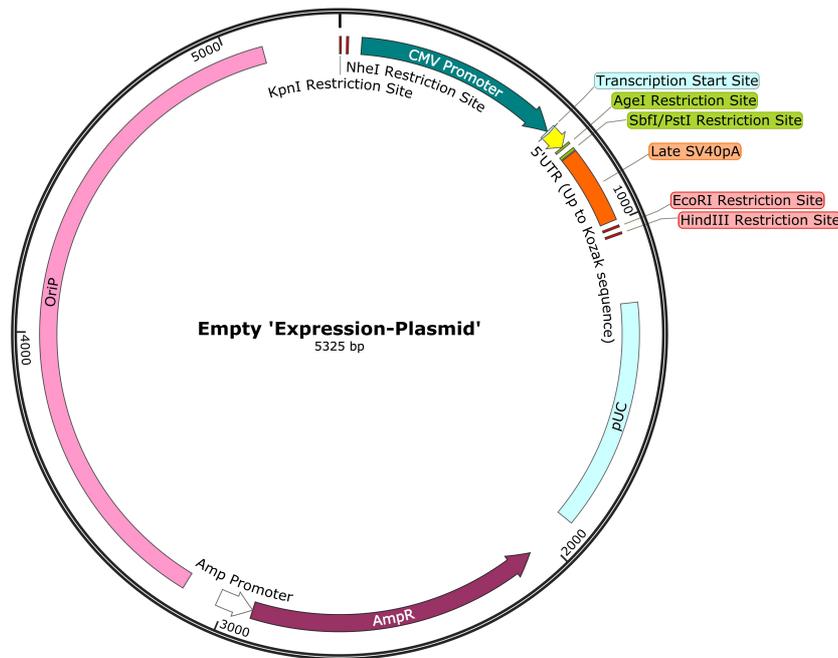


Figure 5.2: 'Expression-Plasmid' Outline Map

The developed plasmid ('Expression-Plasmid') enables standardised expression of genetic components. The plasmid contains bacterial replication and amplification components (AmpR and pUC), mammalian expression components (CMV promoter, 5' and 3' UTR sequences), enhanced transient expression sequence (OriP) and unique restriction sites for cloning purposes.

5.3.1.1. 'Expression-Plasmid' Construction

The 'Expression-Plasmid' was derived from a 'pre-existing' plasmid (referred to as Plasmid-M), kindly provided by MedImmune, which contained both a GS-cassette and an OriP DNA element. To incorporate the required design features, a section from Plasmid-M was replaced with a synthesised fragment containing the desired components. The useful region on Plasmid-M, contained sequences for ampicillin resistance (AmpR), bacterial origin of replication (pUC) and OriP were flanked by HindIII and KpnI restriction sites. A double digest with KpnI and HindIII, enabled the desired 4.2kb fragment to be isolated (**Appendix Figure**

A.1). The region to insert between these sequences was designed, based upon sequences kindly provided by MedImmune, with added alterations to incorporate desirable unique cloning sites and design features, and then synthesised by GeneArt. The insert fragment contained KpnI and NheI restriction sites followed by the sequence for the human CMV promoter, a transcription start site, the 5'UTR up until the Kozak sequence, AgeI and PstI restriction sites, the 3'UTR containing an SV40 polyA tail and finally EcoRI and HindIII restriction sites. The designed ~1kb section of sequence was isolated by a double digest with KpnI and HindIII (**Appendix Figure A.1A**) After ligation, a diagnostic digest checked for presence of the insert (**Appendix Figure A.1B**) before the plasmid was fully sequenced to confirm its identity.

5.3.1.2. Verification of 'Expression-Plasmid' Functionality

To verify the 'Expression-Plasmid' developed was functional and able to express genetic components; a GFP reporter gene was generated, cloned into the newly constructed plasmid and its expression verified using flow cytometry analysis. All inserts for the 'Expression-Plasmid' were designed and synthesised in the same way, with the designed eGFP insert shown as an example in **Figure 5.3A**. Upstream of the 'ATG' start site of the ORF, the Kozak sequence, required for effective mammalian mRNA translation initiation, was included. The nucleotide sequence encoding the protein of interest was altered using the GeneArt gene portal service (Thermo Fisher Scientific); codon usage of the nucleotide sequence was optimised for *C. griseus* expression and the sequence designed with particular restriction sites included or eliminated, to enable straightforward cloning downstream. The entire sequence was made devoid of a collection of restriction sites (AgeI, BamHI, EcoRI, HindIII, KpnI, NheI, NotI, PstI and SbfI), of interest for the in-parallel development of a multigene plasmid construct. AgeI and PstI/SbfI restriction sites were positioned and preserved at the periphery of the insert sequence, at the 5' and 3' end (immediately after the stop codon) respectively; for the insert to be directionally integrated into the expression-plasmid in a straightforward, efficient manner.

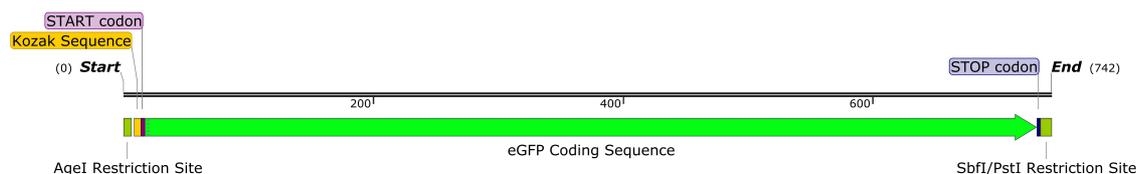


Figure 5.3: Example Insert Sequence Compatible with ‘Expression-Plasmid’

Outline illustrating the layout of an ‘insert’ sequence compatible for integration into the ‘Expression-Plasmid’. The eGFP gene sequence, downstream of a Kozak sequence, is flanked by restriction sites of AgeI and SbfI/PstI for cloning the region into the ‘Expression-Plasmid’.

To incorporate inserts into the ‘Expression-Plasmid’, both the ‘Expression-Plasmid’ and the insert-plasmid were separately digested in one-step double digests with AgeI and SbfI. For eGFP digestion an additional AgeI site was also present in the ampicillin resistance sequence on the GeneArt plasmid, but the fragment sizes resulting from AgeI and SbfI digestion (739, 1424, 1849 bp) allowed for adequate size separation of the fragments by agarose gel electrophoresis and successful isolation of the GFP insert fragment (**Appendix Figure A.2**). After ligation, diagnostic digests confirmed the inclusion of the ~700bp GFP insert with identity confirmed through sequence verification. PstI and SbfI restriction enzymes recognise similar motifs (CTGCAG versus CCTGCAGG respectively). PstI was found to be a significantly cheaper enzyme per unit and a suitable alternative. Therefore, it was used for all subsequent cloning reactions to generate inserts for the ‘Expression-Plasmid’.

To verify ‘Expression-Plasmid’ functionality, transfections with the constructed eGFP ‘Expression-Plasmid’ were compared to those with a commercial plasmid encoding the same amino acid sequence (eGFP-C3). GFP production was measured 24 hours after transfection with the two different plasmids and GFP levels were similar between the transfections (**Figure 5.4**). The analysis of GFP expression demonstrated successful functionality of the generated plasmid and a high transfection efficiency from transfections with the ‘Expression-Plasmid’ were attained. The ‘Expression-Plasmid’ without any insert, referred herein as empty-‘Expression-Plasmid’, was used to equalise DNA load to ensure all transfections were undertaken using a standardised, comparative concentration of plasmid DNA (an equalised total DNA load). This concept is important to try

and keep the dynamics of transfection consistent. The use of ‘filler’ DNA help to maintain a similar, optimised amount, and ratio, of cells and DNA between transfections, resulting in comparable transfection efficiencies, levels of gene expression and post-transfection recovery (Estes et al., 2015; Kichler et al., 2005; Rajendra et al., 2012). Cells are affected differently, in terms of both cytotoxic effects and gene expression, to different amounts of transfected DNA (Figure 5.6). Standardised transfection conditions with the use of ‘filler’ DNA should enable gene dosage to be investigated comparatively (Hansen et al., 2017), without a change in the transfection process altering the efficiency or cytotoxicity of introducing different amounts of DNA.

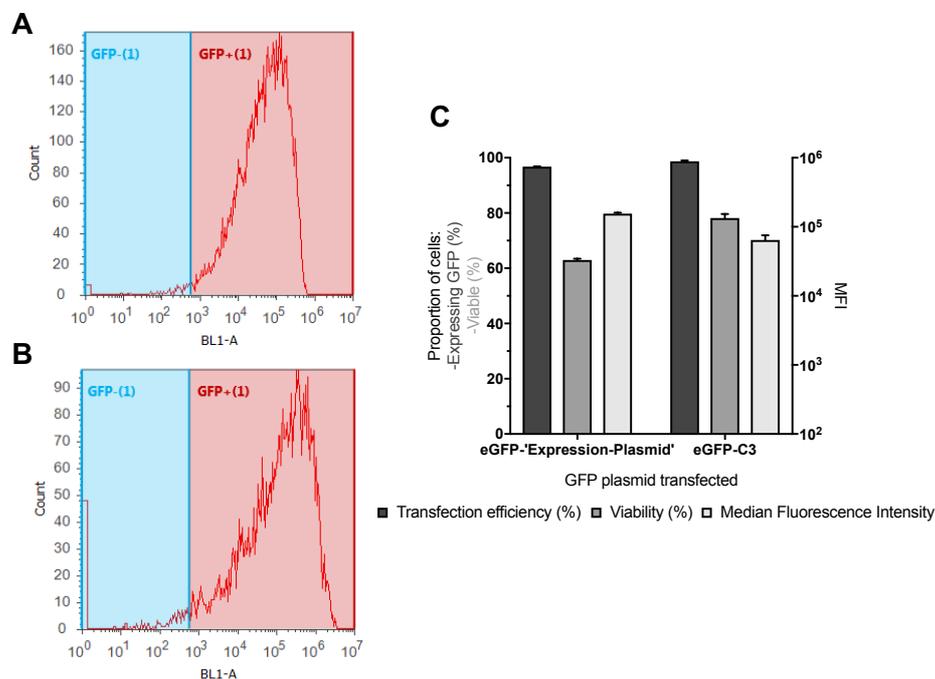


Figure 5.4: GFP Expression Verifies Functionality of Constructed ‘Expression-Plasmid’

Comparison in expression, after 24 hours, of equal copies of eGFP-C3 plasmid (A) or the ‘Expression-Plasmid’ encoding GFP (B), adjusted due to differences in plasmid size (as eGFP-C3 was 4727bp and eGFP-‘Expression-Plasmid’ was 6038bp) using empty-‘Expression-Plasmid’, with total DNA load (1µg). The ‘Expression-Plasmid’ functions correctly, as seen through very similar and high GFP expression compared to the commercial plasmid (A,B), indicating successful plasmid design. Similar levels of expression and cytotoxicity were observed, (C), but a slightly greater negative effect on viability is seen with the expression-plasmid but a higher median fluorescence was achieved. The data (C) represent the means and standard error of the mean (SEM) of three technical replicates.

5.3.2. High-throughput Electroporation Optimisation

Optimisation of the Nucleofection setup was undertaken with the aim to balance several transfection outputs and achieve a robust transient expression platform. A high transfection efficiency was very desirable to ensure the majority of the population was expressing the genes of interest, like in clonal SGE, and being affected by the expressed genes functionality. A high level of immediate gene expression was desirable to increase the likelihood of a measurable and functional effect. This would be particularly beneficial to increase the potential effect from gene overexpression and for multiple plasmid co-transfections. The electroporation method was likely to be, to a certain degree, cytotoxic to cells. Any toxicity needed to be minimised, so healthy populations were investigated and off-target effects from the transfection process were minimised. The transfection process needed to be reproducible and have low amounts of variation, important characteristics for any transfection process. To balance these outputs, the main transfection parameters that could be altered were the Nucleofection programme, the number of cells transfected and the amount of plasmid DNA in the transfection mix.

5.3.2.1. Nucleofector Programme

A range of Nucleofector programmes, recommended for suspension-adapted CHO cell lines (Lonza, personal communication) were compared to find a suitable programme for electroporation of the MEDI-CHO cells. Testing the different programmes was required as their differences, in terms of electroporation parameters (pulse length and strength), were not disclosed and were likely to vary for different CHO cell lines. This initial comparison used the manufacturer's recommended cell number and DNA load, of 1×10^6 cells with 0.4 μg of pmaxGFP plasmid (initial experiments undertaken prior to 'Expression-Plasmid' development). The most suitable transfection programme would be one which reproducibly introduced high levels of DNA expression to the entire cell population (i.e. a median fluorescence intensity (MFI) and high transfection efficiency, respectively) while balancing and minimising the level of negative cytotoxicity introduced from the transfection process (i.e. maintain a high cell viability). As can be seen in **Figure 5.5**, all the transfection programmes tested

produced high transfection efficiency (over 74%) and resulted in healthy cultures (over 77% viability). An inverse relationship was observed between transfection efficiency and culture viability, with greater transfection efficiency associated with increased cell death. Programme FF-158, was deemed to be the most suitable transfection programme; yielding consistent results, an average transfection efficiency of 86%, culture viability of 90% and over a 50-fold increase in MFI compared to an untransfected population. Desirably, programme FF-158 was capable of rapidly introducing a lot of plasmid DNA to almost the entire cell population, without too much cytotoxicity; it was therefore selected to be used for all transfections.

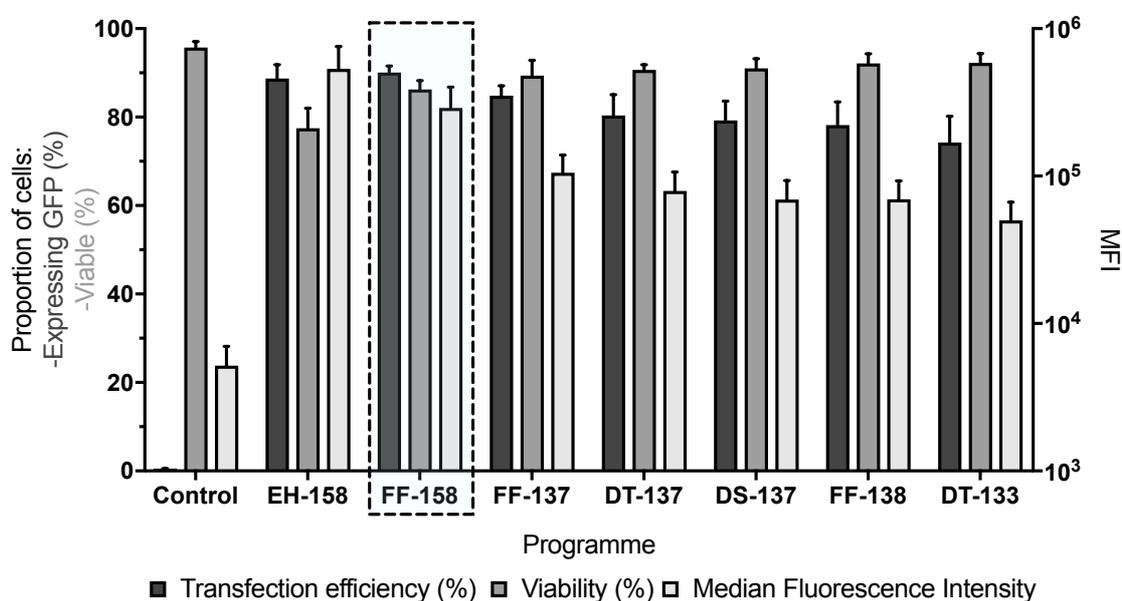


Figure 5.5: Programme FF-158 was the Most Suitable Nucleofector Programme Tested for Rapid, High-Level Introduction of Plasmid DNA

A range of recommended Nucleofector programmes were tested for their effect on GFP expression and cell viability after 24 hours; 1×10^6 cells were transfected with 0.4 μg of pmaxGFP plasmid. The most suitable programme, FF-158, is highlighted by the dashed box. On the left of the graph is a control, mock-transfected cells, showing typical viability and cellular autofluorescence (MFI). The data represent the mean and SEM of three biological replicate experiments, each with two technical replicates.

5.3.2.2. DNA Load and Cell Number Optimisation

The transfection platform required high levels of immediate gene expression for the screening tool to rapidly assess effects of genetic component overexpression. Transfected plasmid loads of above 0.4 μg would increase the numbers of gene

copies within cells which should lead to increased levels expression. The initial optimisation step tested a range of plasmid DNA loads within the manufacturer's recommended range, shown in **Figure 5.6**. Transfections, with 0.3 to 1.8 μg of pmaxGFP plasmid, all produced very high transfection efficiencies. As expected, increasing the transfected plasmid load increased the level of expression (as seen by the increase in MFI). This beneficial trend correlated with an undesirable decrease in culture viability. Intermediate DNA loads of approximately 0.9 to 1.2 μg were optimal in balancing transfection efficiency and viability, while providing a higher level of gene expression beyond that of lower DNA loads.

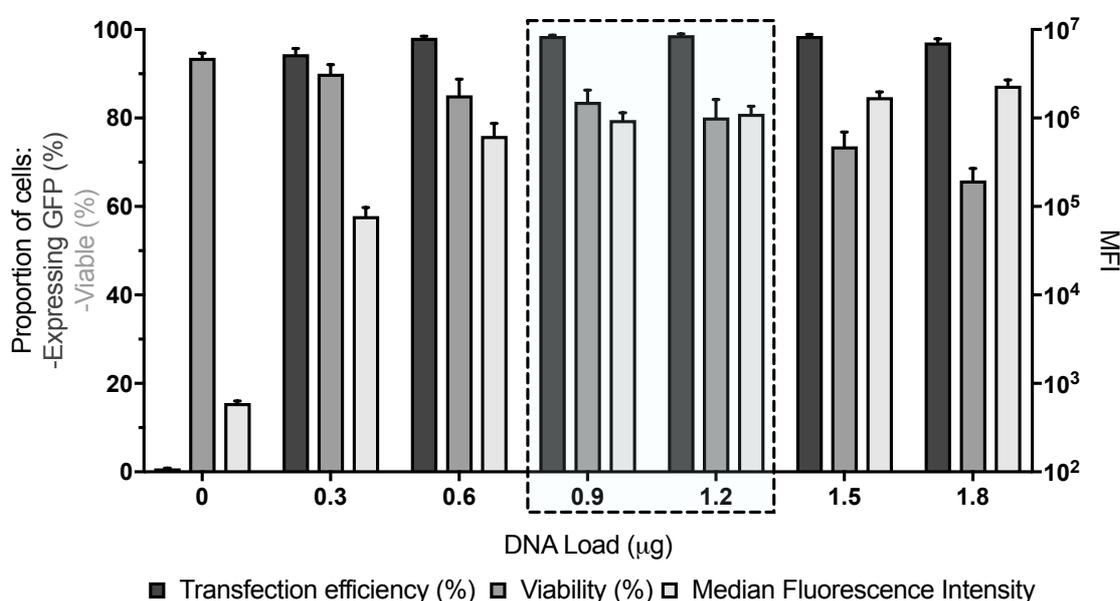


Figure 5.6: 0.9-1.2 μg of Plasmid DNA was the Optimal pmaxGFP Plasmid Load From a Range Based on Manufacturer's Recommendations

A range of plasmid DNA transfection loads were tested using pmaxGFP for their effect on expression and cell viability after 24 hours; with 1×10^6 cells transfected. The optimal amounts of plasmid DNA are highlighted by the dashed box. The data represent the mean and SEM of three biological replicate experiments, each with two technical replicates.

GFP, a small (molecular weight of 27 kDa) and fluorescent protein, not requiring PTM and ER processing, is a widely used reporter protein and an effective tool for studying expression. The transfection platform under development was to test the effect of genetic components within CHO cell factories producing a secreted recombinant protein. It was therefore of relevance to test and optimise the system's capability for transfection and expression of a multimeric, glycosylated,

secreted product, rather than just GFP. A model antibody, provided by MedImmune, Nip109 (herein referred to as ETE-mAb), was utilised for this purpose. Post-transfection viability, over a range of transfected DNA loads, was found to be unsatisfactorily lower after transfections with the ETE-mAb expressing plasmid, compared with GFP expressing plasmids. This is exemplified in a comparison using two DNA loads, shown in **Figure 5.7**. Across these two loads, on average around 40% further reduction in culture viability, after 24 hours, was seen from transfections with mAb plasmid. This effect on viability was seen immediately (2 hours) after transfection (data not shown).

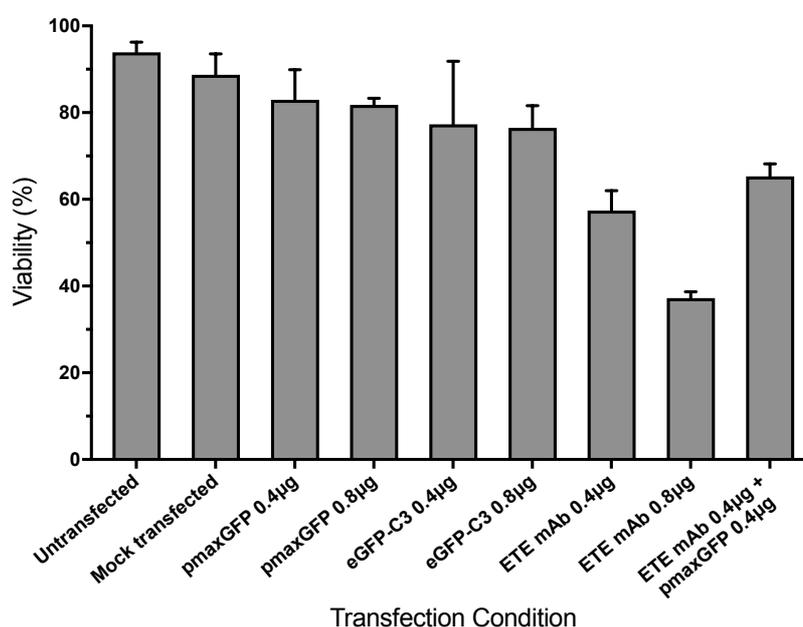


Figure 5.7: Transfections with mAb Encoding Plasmid DNA had Negative Effects on Post-Transfection Culture Viability

Differences were seen in post-transfection culture viability after transfections with mAb encoding plasmid when compared to those with two GFP encoding plasmids. 1×10^6 cells were transfected separately with 0.4 and 0.8 μg of either pmaxGFP, eGFP-C3 or ETE-mAb encoding plasmid or co-transfection with 0.4 μg pmaxGFP and of 0.4 μg ETE-mAb plasmids simultaneously. Cell viability assessed after 24 hours is shown. A substantial reduction in cell viability was seen after transfections containing mAb plasmid. The data represent the mean and SEM of two biological replicate experiments, each with two technical replicates.

The differences in post-transfection viability for different plasmids with the same DNA load was surprising. Very little information was provided in relation to the pmaxGFP plasmid used for ascertaining transfection efficiency (plasmid proprietary to Lonza). The pmaxGFP suspension buffer was not disclosed.

Changes to the buffer used to resuspend the ETE-mAb DNA for transfection (e.g. TE buffer or H₂O) did not have any effect on post-transfection culture viability (data not shown). Transfections with a well-characterised, commercially available GFP expressing plasmid, eGFP-C3, were undertaken for comparison. The size, sequences and components present on the eGFP-C3 plasmid were known (plasmid map shown in **Appendix Figure A.3A**). The eGFP-C3 plasmid was 4.7kb in size, around half the size to the ETE-mAb plasmid (9.5kb), and contained two ORFs – a GFP gene and a resistance marker (neomycin). The ETE-mAb plasmid expressed genes for the HC chain, LC chain and GS. As seen in **Figure 5.7**, effects on viability were similar from transfections with both GFP expressing plasmids. This suggests the increased plasmid size, and complexity of the products being generated, was resulting in the undesirably lower viability. This link could have been investigated further through undertaking transfections with various sized plasmids, encoding genes of differing complexities.

The poor cell viability post-transfection was not suitable for the transfection methodology being developed. To investigate a more optimal set of parameters a DoE approach was utilised, combining different amounts of plasmid with different numbers of cells being transfected. The range of cell numbers tested, 1-2x10⁶, was derived from the manufacturer's recommendation. Lower cell numbers were eliminated due to very poor preliminary transfection outputs resulting from these (data not shown). Having sufficient cells to seed adequate culture volumes to measure desired expression outputs was also kept in mind, an increased number of cells transfected being beneficial. The range of plasmid DNA investigated was also derived from the manufacturer's recommendation. Due to the large negative effect on viability apparent from initial transfections with an ETE-mAb expressing plasmid, a lowered range between 0.2 and 0.8 µg of plasmid was tested.

A design of experiments response surface model (DoE-RSM) methodology was used to predict optimal solutions between the two parameters (factors). The experimental setup was designed, and subsequently analysed, using Design-Expert® 10.0.6.0 software (Stat-Ease, Minneapolis, USA). A central composite design mapped responses over the ranges of the two parameters, with the levels of each factor tested summarised in **Table 5.1**. The input criteria, of 1-2x10⁶ cells

and 0.2-0.8 µg of DNA, set the boundaries for the design – the high and low testing points, each assessed in duplicate. The centre test point (1.5x10⁶ cells with 0.5 µg of ETE-mAb DNA) was replicated five times, allowing the models to provide an estimate of the pure error within the setup. These points, along with two test points outside of these boundaries (axial/star points) also included in duplicate, provided an estimate of curvature for any quadratic models generated.

Table 5.1: Levels of Factors Used in Central Composite Design Used to Optimise Transfection DNA Load and Cell Number

	Low	High	Centre	-Axial	+Axial
Cell Number (x10 ⁶)	1	2	1.5	0.8	2.2
DNA Load (µg)	0.2	0.8	0.5	0.07	0.92

Post-transfection viability and VCD after 72 hours were measured, along with the titre produced; these were used to calculate post-transfection IVCD and qP. These outputs were then added to the Design-Expert® software to produce statistical models, which established any relationships between the input factors and output responses. A model for each response was generated and analysed using the software’s guidance. Viability, titre and qP data were all fitted to quadratic models, whereas IVCD was fitted to a linear model. The following equations of coded factors, where A (x10⁶) is the cell number and B is the DNA Load (µg), were used to describe the models generated for each output.

$$(\text{Viability})^{1.32} = 264.68 - 5.6A - 85.31B - 10.55AB + 23.09A^2 - 15.34B^2 \quad (5.1)$$

$$(\text{IVCD})^{0.15} = 1.08 - 3.726e^{-003}A - 0.11B \quad (5.2)$$

$$(\text{Titre})^{0.2} = 1.09 - 4.11e^{003}A + 0.12B - 0.033AB - 0.041A^2 - 0.11B^2 \quad (5.3)$$

$$(\text{qP})^{-0.2} = 1.02 - 0.018A - 0.28B - 2.758e^{-003}AB + 0.034A^2 + 0.21B^2 \quad (5.4)$$

The summary of information relating to the models is in **Table 5.2**. Analysed by ANOVA, all resulting models were statistically significant (p<0.05). They also all had a significant “lack-of-fit”, meaning the proposed model fitted the measured

data well in terms of variance, and could be used. All models, apart from the model for titre, had good, reliable model predictability for the effect of different DNA loads and cell numbers transfected on the outputs (predicted versus adjusted R^2 values within 0.2 of each other; any outcomes derived from the titre model needed to be checked experimentally. The RSM graphs, in **Figure 5.8**, visualise each transfection output, in terms of the relationship between the cell number and the amount of DNA transfected

Table 5.2: Summary of Response Model Analysis of Outputs from the Central Composite Design Used to Optimise Transfection DNA Load and Cell Number

Factors A and B correspond to cell number ($\times 10^6$) and DNA Load (μg), respectively. Normal plots of residuals demonstrated normality of residuals for all models of viability, titre and qP (**Appendix Figure A.4**).

Response	Transform (λ)	Factor	Sum of Squares	P Value (<i>significant factors in bold</i>)	Model Predictability (Pred/Adj R^2)
Viability	1.32	Model	65007.20	0.004	0.79/0.90
		A	250.75	0.5394	
		B	58220.47	<0.0001	
		AB	445.34	0.4183	
		A ²	3707.73	0.0421	
		B ²	1636.66	0.1433	
Titre	0.2	Model	0.20	0.0003	0.67/0.90
		A	0.0001352	0.7850	
		B	0.11	<0.0001	
		AB	0.004279	0.1548	
		A ²	0.012	0.0338	
		B ²	0.077	0.0003	
qP	-0.2	Model	0.90	<0.0001	0.97/0.99
		A	0.002453	0.1873	
		B	0.37	<0.0001	
		AB	0.000030	0.8738	
		A ²	0.007599	0.0396	
		B ²	0.18	<0.0001	
IVCD		Model	0.78	<0.0001	0.95/0.97
		A	0.000094	0.5428	
		B	0.075	<0.0001	

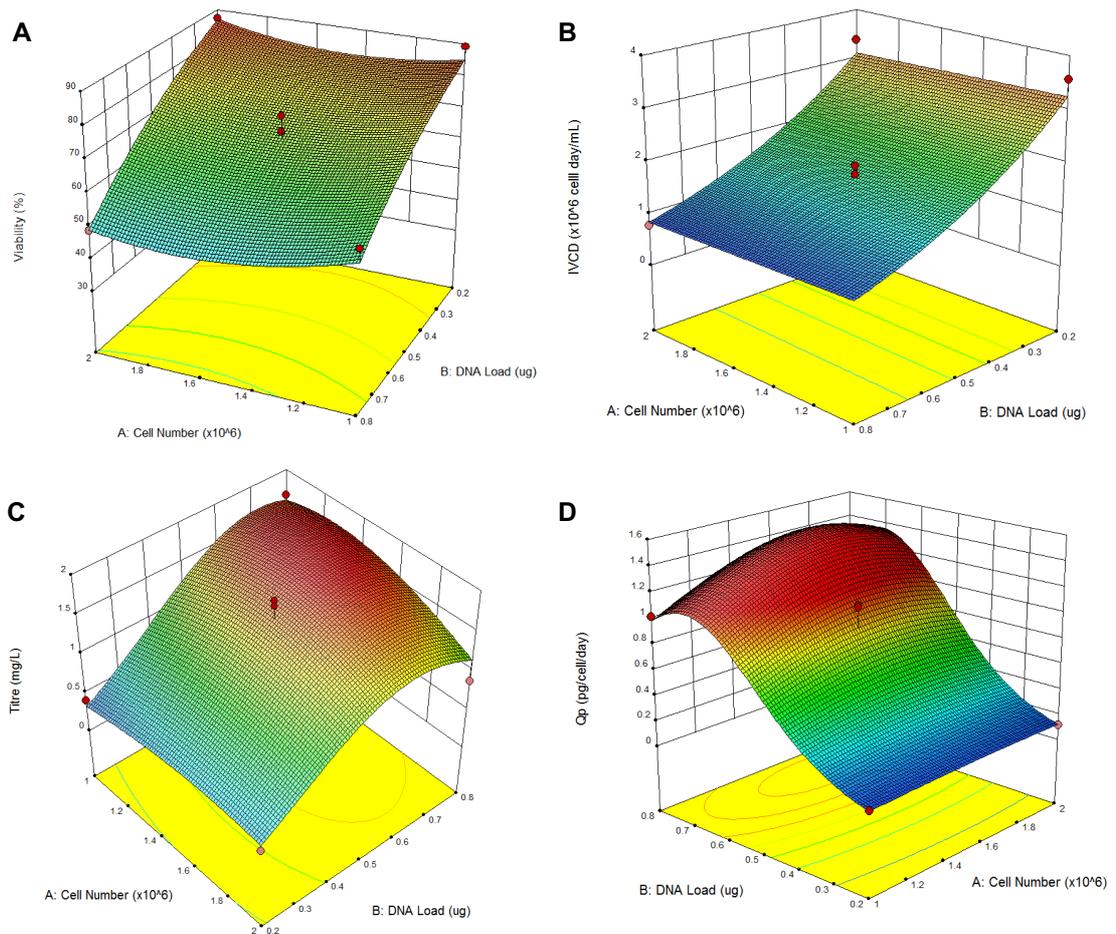


Figure 5.8: Three-Dimensional Surface Graphs Derived from DNA Load and Cell Number Optimisation Using DoE-RSM

Three-dimensional surface graphs from the DoE RSM optimisation process, where the number of cells and the amount of plasmid DNA transfected were altered over ranges of $1\text{-}2 \times 10^6$ cells and $0.2\text{-}0.8 \mu\text{g}$ of DNA, respectively. The graphs show the relationship between these input factors and the outputs, measured after 72 hours, of **(A)** culture viability (%), **(B)** IVCD ($10^6 \text{ cell day mL}^{-1}$), **(C)** Titre (mg L^{-1}) and **(D)** qP ($\text{pg cell}^{-1} \text{ mL}^{-1}$).

The DoE-RSM shows a clear negative correlation between transfection efficiency (inferred through productivity of transfected genes) and cell toxicity inferred through IVCD and viability). The significant factor causing this relationship was found to be the amount of plasmid DNA being transfected into cells. Across the range of cell numbers tested, lower amounts of DNA transfected resulted in improved culture viability and growth post-transfection, but a reduced productivity level throughout. The software could be used to undertake theoretical, numerical optimisation, combining the information from the analysed models. The model for titre was ignored due to its unpredictability (predicted versus adjusted R^2 values

had a difference of greater than 0.2). The criteria for the optimal solution derived could be adjusted depending on the desired outcome. For the transfection method being generated, the importance was primarily on post-transfection expression capability while maintaining cell growth as best as possible (i.e. high qP and IVCD); therefore, these attributes were given priority weighting. Viability was an important output, too much cell death would be detrimental to any screening tool's ability to test components, a viability threshold of > 65% was used, although a viability higher than this was desirable. An optimal solution proposed to achieve these criteria were transfections conditions of approximately 1.4×10^6 cells and 0.5-0.6 μg of plasmid DNA (desirability=0.63). These parameters formed the basis for the final experiments undertaken to finalise a satisfactorily optimal cell number and DNA load to use for the high-throughput screening platform. Lower DNA loads would have improved post-transfection viability, but would have been unfavourable to productivity levels, which needed to be kept sufficiently high for functionality screening and to make sure they could be detectable and measurable.

There were found to be discrepancies in the cell concentrations present during plate-based transfection setup. This primarily stemmed from inherent difficulties handling small volumes, a downside to a high-throughput micro-scale approach. It was determined that the primary source of the inconsistency arose after centrifugation, when cells were removed from their culture media and re-suspended in Nucleofection solution prior transfection. It was very difficult, when re-suspending cells in small volumes ($15 \mu\text{L well}^{-1}$), to completely remove the supernatant and for this residual liquid to not significantly affect the final volume. Consequently, cells were found to be diluted further than projected during the transfection process; 40% extra cells were required to overcome these discrepancies (based on setups for high-throughput transfections where whole plates, >96 samples, are being transfected). Therefore, to achieve transfections of approximately 1.4×10^6 cells per well, the number of cells that were centrifuged and isolated needed to be 2.33×10^6 cells per well.

The transfection setup would involve a combination of effector gene plasmids and recombinant product genes. Therefore, finalisation of the DNA load to use did not use just recombinant mAb plasmid. Once the 'Expression-Plasmid', described

previously in **Section 5.3.1** had been successfully developed and modified to encode for eGFP (to mimic effector gene expression), it was used, along with the recombinant mAb plasmid, to finalise total DNA load for transfection. The DTE-mAb plasmid was used as this represented the more 'worst-case' transfection scenario, and was a plasmid of interest for the screening platform. The recommended cell number, from the DoE experiments, 1.4×10^6 cells per well, was used for transfections herein. A range of DNA loads, with different ratios of eGFP 'Expression-Plasmid' and DTE-mAb plasmid, were transfected; the DNA load of 0.6 μg derived from the DoE experiment was the lowest load tested. Transfections with higher amounts of DNA would be more logistically versatile to implement co-transfection of multiple plasmids simultaneously and higher copies of the mAb genes would generate higher titres that would be more detectable. This is of particular relevance for a short-term screen with a DTE-mAb, where titres will be inherently low. In **Figure 5.9**, total plasmid loads of 0.6, 0.8 and 1.0 μg are shown, made up of equal and 2:1 eGFP 'Expression-Plasmid' to mAb-expressing plasmid ratios, respectively. A total load of 0.8 μg of plasmid generated the most optimal combination of transfection outputs, with adequate post-transfection viability ($\sim 70\%$), sufficient expression of both plasmids (represented through titre and MFI) while not being too detrimental to cell growth (represented by IVCD). Either ratio of mAb plasmid to eGFP 'Expression-Plasmid' was acceptable. It was decided the ratio of plasmids used should be recombinant product and screen specific, depending on the product's dynamics, the burden to cells and the requirements of the screen.

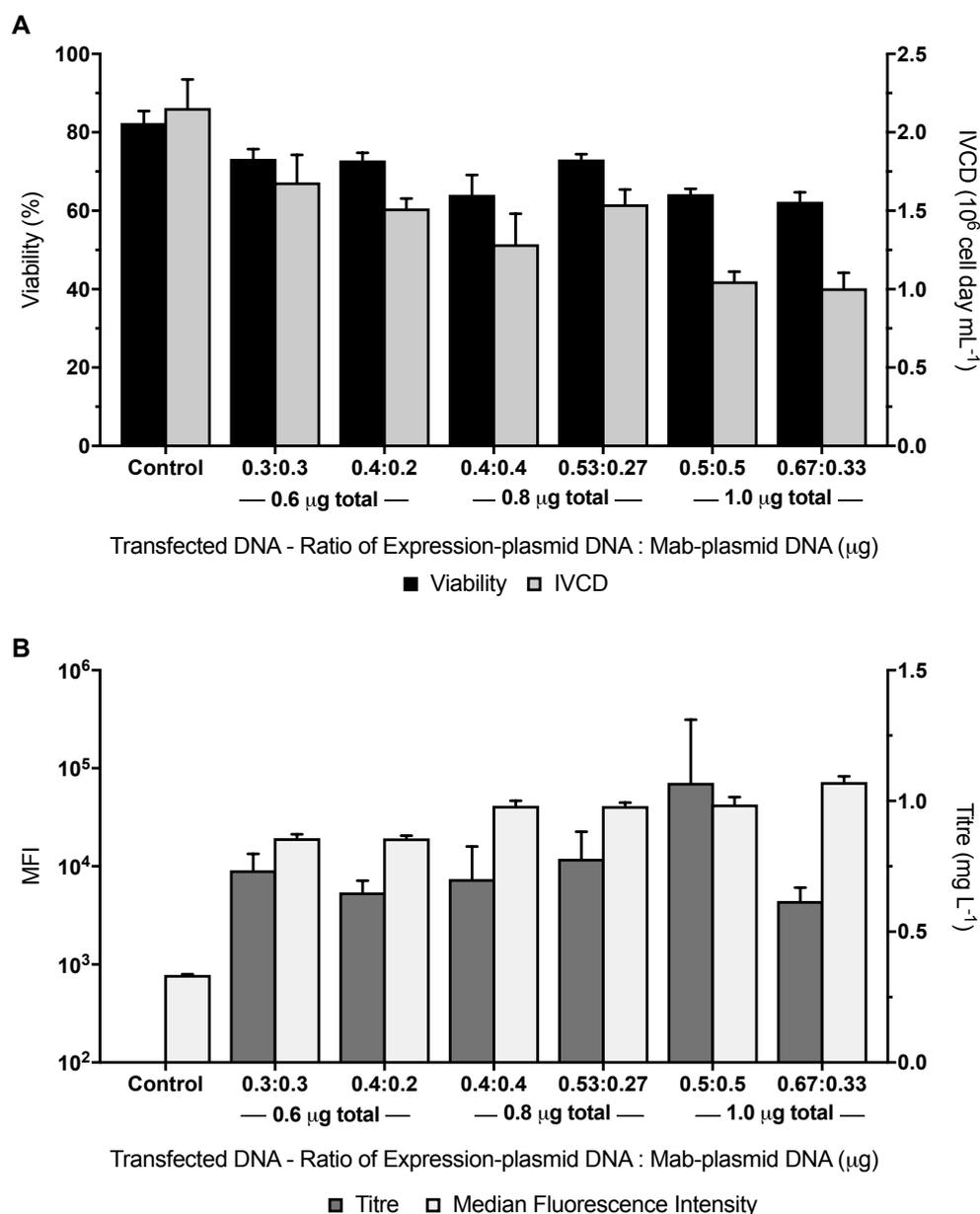


Figure 5.9: Multiplexing Transfected Plasmid Type to Determine Final Transfection DNA Load

To finalise the amount of DNA used for well-based Nucleofection, co-transfections with mAb-encoding DNA and ‘Expression-Plasmid’ DNA at different total loads were compared. Total transfection loads of 0.6, 0.8 and 1.0 μg of DNA are shown. These loads were split between eGFP ‘Expression-Plasmid’ and DTE-mAb plasmid at ratios of 1:1 and 2:1. For each transfection 1.4×10^6 cells were transfected. Of-interest transfection outputs are shown: **(A)** culture viability and IVCD after 72 hours, **(B)** MFI from GFP expression and titre measurements from DTE-mAb production, after 24 and 72 hours respectively. The data represent the mean and SEM from two biological transfections, each with three replicates.

5.3.3. Comparison of Plate-Based Electroporation to Plate-Based Lipofection

A comparison with an in-house, optimised lipofection protocol was undertaken to investigate the developed platform against a prevalent alternative approach. The transfection approach needed to rapidly cause high levels of expression of the recombinant genes being transfected, for their effect to be functional and measurable after only a short duration. The two optimised transfection approaches were compared, through transfections with both GFP and SEAP, a model secreted protein that can be easily quantified, to confirm the developed electroporation setup best met this requirement. The results from the two techniques are compared in **Figure 5.10**.

Transfection with a GFP-expressing plasmid showed that after 24 hours, electroporation transfections reproducibly resulted in almost all (>98%) of the population expressing GFP; compared to 71% from lipofection based transfections. Both methods resulted in reductions in viability, but there was a more negative effect from lipofection based transfections (77% viability for lipofection as opposed to 88% for electroporation). In terms of relative gene expression, the MFI demonstrated a higher expression of GFP from electroporation (in part attributed to the higher transfection efficiency). Transfection with a plasmid encoding SEAP (plasmid map shown in **Appendix Figure A.3B**) showed the specific production capabilities (amount each cell produces) of electroporated cells was approximately 70% higher than cells which had been transfected using Lipofectamine[®]. These trends were also seen after 72 hours (data not shown). Results from electroporation transfections were more consistent and less variable. 20% less DNA per electroporation transfection was required (though this was paired with a higher cell number requirement); with 0.8 µg per 1.4x10⁶ cells and 1 µg per 2x10⁵ cells for electroporation and Lipofection based transfections, respectively. A lower plasmid DNA requirement is a useful trait in a high-throughput screen.

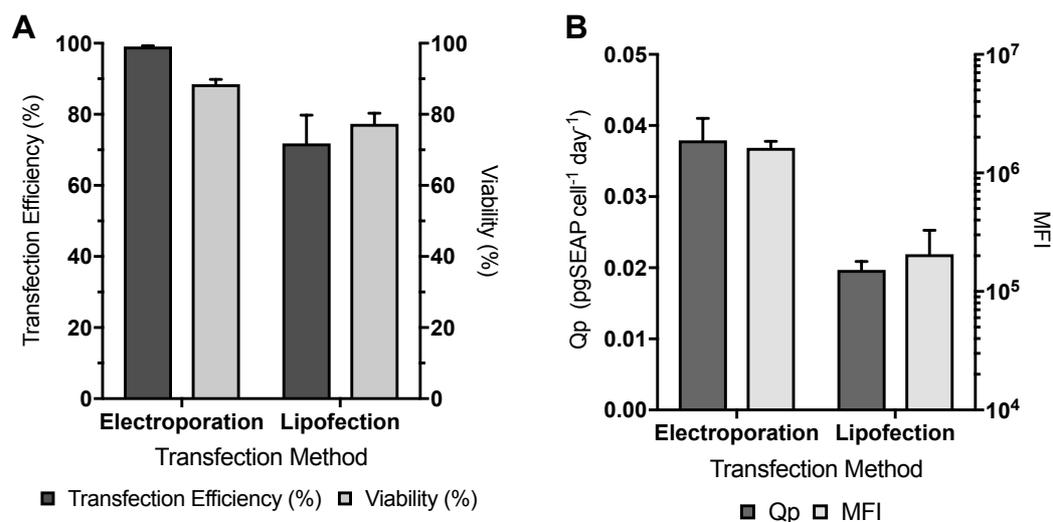


Figure 5.10: Comparison of Electroporation and Lipofection Transfection Methods Demonstrates Electroporation is a More Suitable Transient Transfection Approach for Rapid, High Level Gene Overexpression

The two transfection approaches were compared for their immediate transfection efficiency and effect on plasmid DNA expression after 24 hours. For the optimised lipofection setup 2×10^5 cells were transfected with a ratio of 3 μ L of Lipofectamine[®] reagent to 1 μ g of plasmid DNA. **(A)** The transfection efficiency, viability and **(B)** MFI from transfections with a pmaxGFP plasmid and **(B)** the specific production from transfections with a recombinant secreted product (SEAP) are shown. The data represent the mean and SEM from three biological transfections, each with three replicates.

5.3.4. Post-Transfection Plate-Based Culture

Multi-well plate-based culture is synonymous with, and a down-stream requirement of, the micro-plate-based transfection setup described so far. Plate-based post-transfection cultures enabled higher throughput potential and an easier handling of samples. Use of 96-well plates had compatible geometry and provided convenience with plate-based Nucleofection and the PrestoBlue[®] and ELISA methods described **Sections 4.4** and **4.6**.

Static, micro-plate based culture was used as it is the most extensively utilised and straight-forward multi-well plate-based culture approach. 96-well tissue culture treated plates were used with culture volumes of 100 μ L per well. The surface-area of the plate (~6 mm) and the use of static culture were restrictive to confluency and biomass accumulation. The screening tool being developed

needed to assess component effect after three to five days if a rapid assessment was to be made. To heighten the effect within this period, particularly to increase titre accumulation aiding measurement for DTE products, a range of seeding densities ($0.2 - 1 \times 10^6$ cells mL^{-1}) were compared. A balance between increasing short-term biomass accumulation without adversely impacting cell viability was required. Growth from optimal seeding densities ($0.2 - 0.6 \times 10^6$ cells mL^{-1}) is shown in **Figure 5.11**, with viability shown over the period of interest, of three to five days. As culture viability began to decline after day three ($<90\%$), culture duration was not extended beyond this period. A seeding density of 0.4×10^6 cells mL^{-1} provided the most optimal biomass accumulation by day three while cells remained in an exponential phase of culture. Increasing the seeding density caused cultures to reach stationary growth phase and decline in culture viability and biomass quicker; unsuitable for the developing screening tool where changes to cell growth are being investigated.

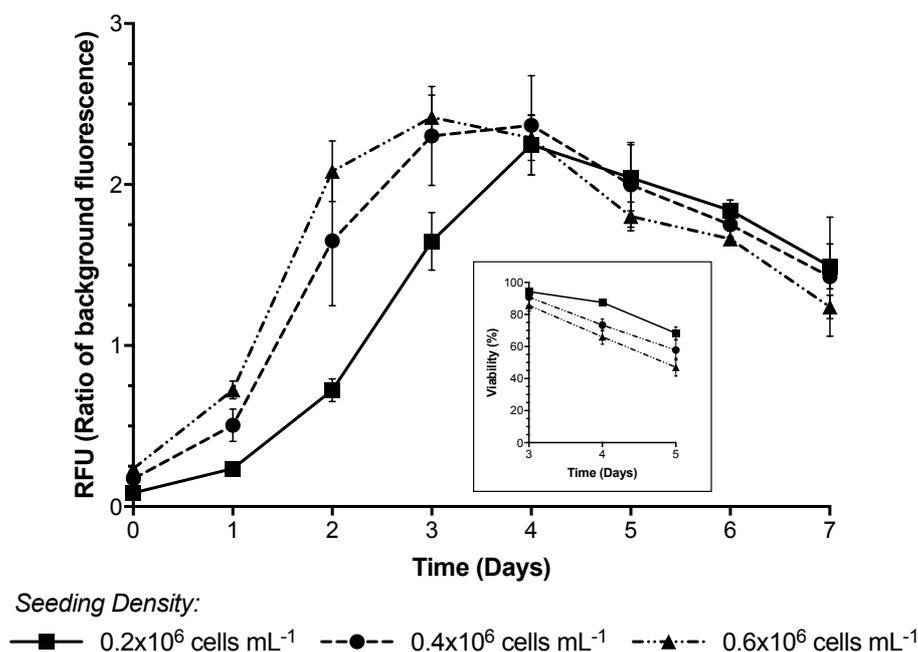


Figure 5.11: 96-Well Micro-Plate Static Culture Duration is Limited

A range of seeding densities ($0.2, 0.4$ and 0.6×10^6 cells mL^{-1} shown), at $100 \mu\text{L}$ volumes, were seeded into 96-well culture micro-plates, grown under static conditions for seven days. Daily PrestoBlue[®] analysis was used to assess cell biomass accumulation then decline over the seven-day period. The insert shows declining culture viability from day three through to five, measured using the Vi Cell. The data represent the mean and SEM from two biological transfections, each with three replicates.

Employing a shaking plate-based culture setup would introduce several benefits over a static micro-plate culture setup. Use of shaking deep-well 96-well plates increased culture volumes, improved cell culture performance and extended culture durations; while maintaining compatibility to the developed platform. A semi-optimised shaking-plate culture format, under development in the David James (DCJ) laboratory at The University of Sheffield, provided shaking 96-well plate-based culture for MedImmune's transient host cell line (CHO-T2). CHO-T2 expressed EBNA-1, enhancing transient expression and longevity (Daramola et al., 2014). Screening for functionality using this cell line would provide added benefit, to help maintain transient expression, if longer culture durations were to be investigated. To harness the shaking-plate based culture format, the transient host cell line needed to be sufficiently and comparably transfected in the transfection platform. Transfection efficiency and post-transfection viability from GFP transfections were explored. Shown in **Figure 5.12** is a comparison, from transfections with pmax-GFP, between the CHO-T2 and MEDI-CHO cell lines. Both cell lines had almost identical post-transfection results, in terms of transfection efficiency, MFI and viability at 24 hours and an equally similar comparison is seen from transfections with the eGFP-'Expression-Plasmid' (data not shown). These results demonstrate the transfection process and effect was very similar between the two cell lines and further optimisation for CHO-T2 was not required. CHO-T2 cells had better recovery and expression levels than MEDI-CHO by 72 hours (data not shown). With transfections using the 'Expression-Plasmid' a higher viability, proportion of cells expressing the transfected gene (GFP) and higher expression levels (higher MFI) were measured. The expressed-plasmid contains an OriP element which, along with the shaking-based culture mode, would have contributed to the enhanced post-transfection performance.

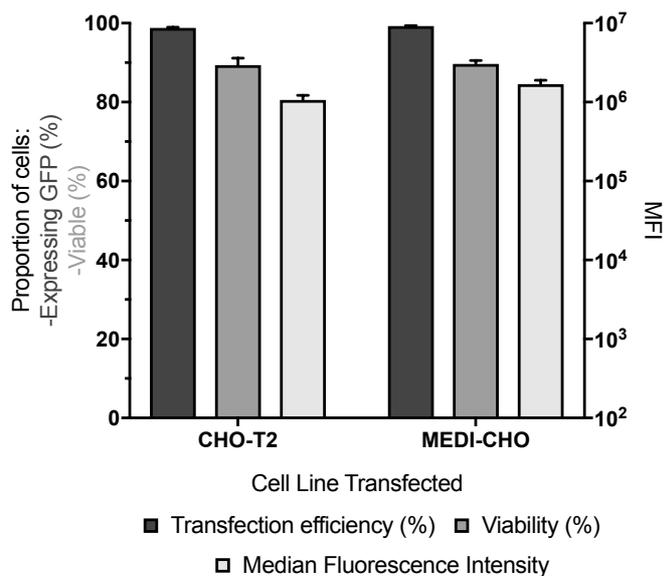


Figure 5.12: Comparison Between Transfections Using MEDI-CHO and CHO-T2 Cells Demonstrates Very Close Comparability of Transfection Between Cell Lines

The two cell lines, MEDI-CHO and CHO-T2, were compared for their immediate transfection ability, efficiency and effect on viability, using the developed platform, with cells transfected with 800ng of pmax-GFP plasmid. CHO-T2 cells were cultured using a semi-optimised deep-well plate-based shaking culture mode and are compared to static micro-plate culture of MEDI-CHO cells. The transfection efficiency, viability and MFI measured after 24 hours are shown for two biological replicates, each with three technical replicates. Error bars show SEM.

5.3.5. Plate-Based Parameter Measurements

Due to the nature of the micro-scale transfection and culture methodologies, only small sample volumes (<500 μ L for all measurements) were available to measure the cell culture performance resulting from the transfection process and expression of the introduced genes. The platform aimed to increase screening throughput, resulting in potentially hundreds of samples to measure simultaneously. The key parameters of interest to assess cell culture performance deriving from effector gene expression were recombinant protein titre (productivity) and cell growth (IVCD). 96-well plate-based assays compatible with the transfection and culture setups were required. To measure titre a Human IgG FastELISA kit was used (identical to that used in Johari et al., 2015). The binding of IgG Fc regions within sample supernatants was proportional to an enzymatic reaction which generated a measurable colour change.

To measure cell growth a PrestoBlue® viability dye method was utilised. The dye was added directly to cells, and only required straightforward mixing and a short incubation period, it was an irreversible process so only suitable for an endpoint assay. The fluorescence signal generated was relative to the amount of biomass (the number of cells) present. The method was compared, to the industry standard cell counting tool (**Figure 5.13A**), a Vi-Cell XR viability analyser, which uses an automated trypan blue exclusion dye method. This normal approach was not suitable for the screening setup here as the cell numbers and volumes being measured were too low and the relatively slow duration, of approximately three minutes, per sample, was too slow for the high-throughput nature of the platform. Viable cell density measured using PrestoBlue® was consistent to that measured using the Vi Cell. The methods correlated well ($R^2 = 0.982$). PrestoBlue® could measure cell densities well within the linear range for the setup duration and volume; it had higher sensitivity than the Vi cell, lower cell numbers could be more accurately measured. As the relative fluorescence measurement, derived using PrestoBlue®, was proportional to the concentration of viable cells, the relative fluorescent units (RFU) generated by a sample of cells could be directly substituted for VCD in calculations determining IVCD and qP.

One major disadvantage of PrestoBlue® was the inability to ascertain culture viability (percentage of viable cells within the total population), a prevalent assessment for cell culture performance. To measure cell viability, a plate-based flow cytometry method could be utilised, harnessing TO-PRO®3, a viability dye. The dye, excluded by living cells, binds to nucleic acids and results in fluorescence detected by a far-red laser on a flow cytometer. TO-PRO®3 has maximal excitation/emission at 642/661 nm, suitable for detection using a 633 nm laser with a 660/20 bandpass filter. The proportion of the population emitting far-red fluorescence directly reflected the proportion of the population that was not-viable. The assessment of culture viability was not exactly the same as for trypan blue exclusion, but the flow cytometry method was still fairly comparable to the Vi cell (R^2 of 0.935) (**Figure 5.13B**). There was more variability in measurements using TO-PRO®3, but viability can be ascertained over a wide dynamic range from the low volume cultures.

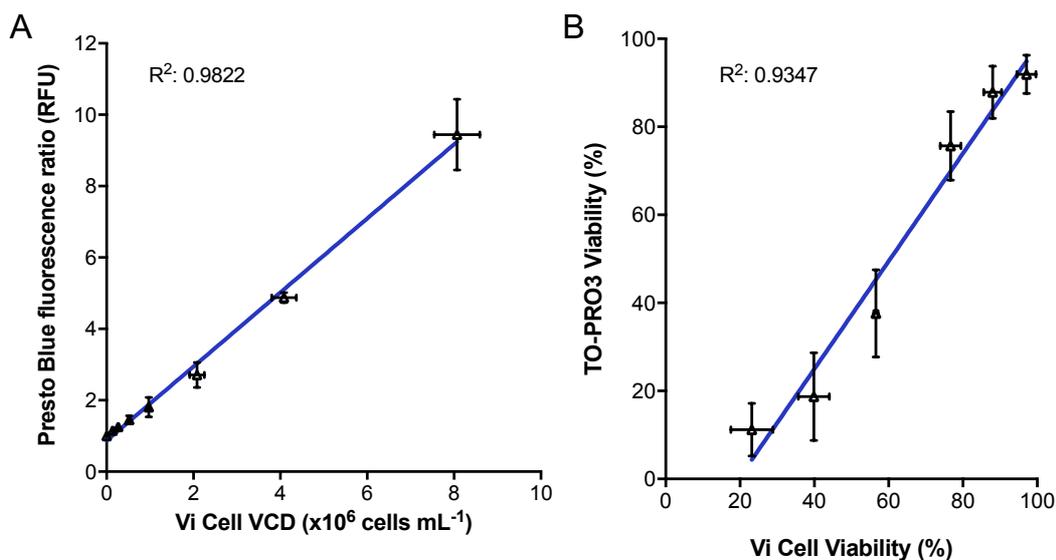


Figure 5.13: Positive Correlation Between Vi Cell Trypan Blue Exclusion and (A) PrestoBlue® Resazurin Assay and (B) TO-PRO³ Viability Assessment.

There is a good correlation between the gold standard trypan blue exclusion method and (A) PrestoBlue® and (B) TO-PRO³/flow cytometry, with R² values of 0.9822 and 0.9347, respectively. The data represent the mean and standard deviation from three biological transfections, each with three replicates.

5.3.6. Platform Variation

The developed transfection setup comprised of multiple in-parallel transfections, each taking place at low volumes in separate wells within the 96-well Nucleofection plate, this facilitated an increased throughput for screening. The handling and set up of transfections improved over time with practice; despite this, the high-throughput format and use of small volumes (20 μ L per transfection) are inherently more variable and likely to having increased amounts of error, than those on a larger scale. Investigations into the variation within the setup were made. Some of this variation could be reduced, by using an appropriate number of replicate transfections and by using controls within each transfection plate to make comparable conclusions.

The experimental outline used to investigate transfection variation is outlined in **Figure 5.14**. Three identical transfection runs were undertaken on separate occasions, each comprised of transfections of entire Nucleofactor plates. All wells on the 96-well plate were each co-transfected with DTE-mAb expressing plasmid and empty-‘Expression-Plasmid’ (267 ng and 533 ng, respectively), and were

each used to seed six 96-well culture plates. This enabled three discrete measurements for post-transfection cell growth (μ) (after 72 hours) for each transfection well. This provided an indication of variation stemming from different transfection runs (set up at different times with a different passage of cells), along with that derived from seeding multiple plates from the same transfected well of cells and that from separate micro-scale transfections (well-to-well variation).

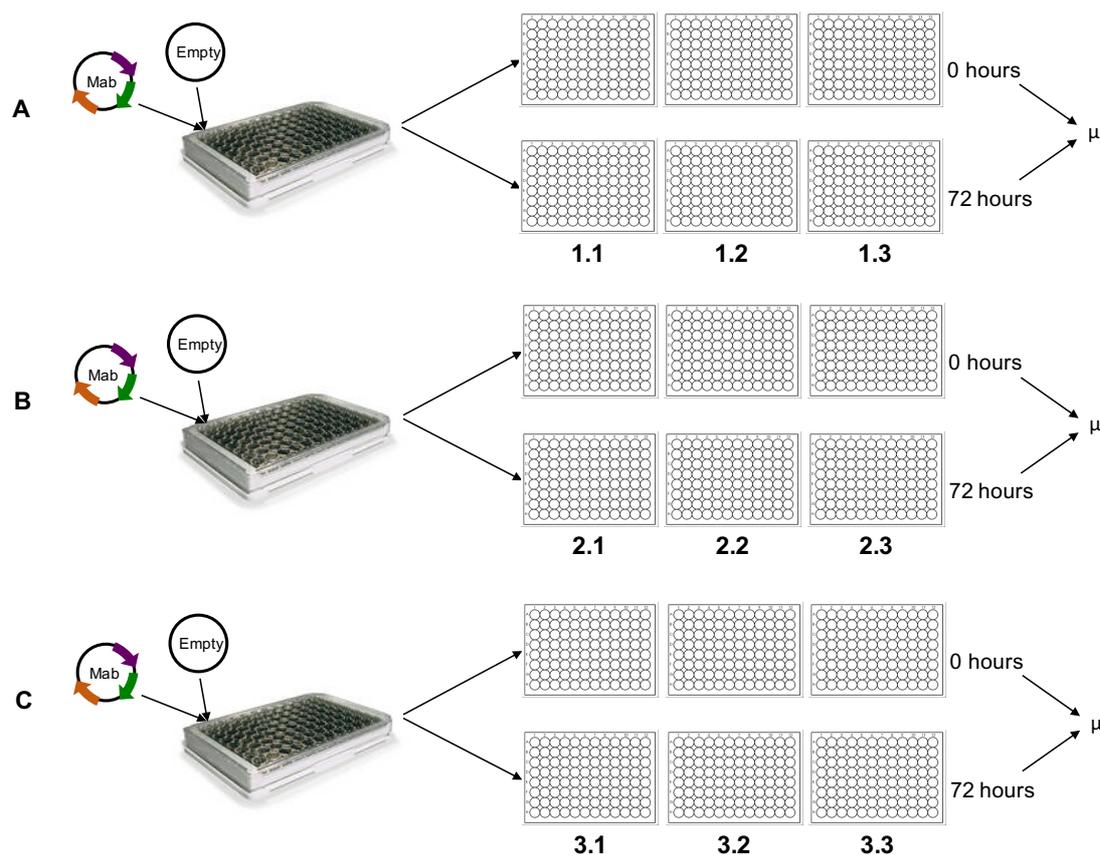


Figure 5.14: Experimental Outline to Investigate Platform Variation

Outline schematic of the experiment used to investigate variation within the platform setup. Three separate transfection runs were set up (A, B and C) at different times. All wells on each transfection plate were transfected with 533 ng of mAb-encoding plasmid and 267 ng of empty-‘Expression-Plasmid’. After transfection, each transfection well was seeded into six separate 96-well shallow culture plates. Three plates were measured immediately using PrestoBlue[®] analysis, the other three were assessed using PrestoBlue[®] after 72 hours of static growth. The measurements were used to calculate three growth rates (μ) per transfection well (e.g. 1.1, 1.2, 1.3).

The growth rates derived from three independent transfection runs, with each transfected well seeded into three separate culture plates, is represented in the box plot in **Figure 5.15**. The growth rates from the three plates from the first

transfection run are labelled in box A 1.1, 1.2, 1.3; with the second and third denoted in boxes B and C, respectively. The dashed lines (whiskers) show the range of growth rates, between the minimum and maximum on each plate, excluding outliers (points designated above and below, the upper and lower quartiles respectively, by beyond 1.5 times the interquartile range). The box represents the interquartile range, with the middle 50% of growth rates for each plate within the box. These regions are quite substantial and varied between plates and transfection runs; the largest contributor to variation appears to be derived from this transfection-to-transfection (well-to-well) variation. The bold black bars represent the average growth rate for each culture plate. There are differences in average growth rates between plates, but these are not as substantial as the differences in the range of growth rates. Some of this variation stemmed from the different transfection runs setup at different times with different passages of cells (i.e. between A, B and C). Some inherent biological variation of this nature is to be expected. There is also some variation stemming from seeding different culture wells on different culture plates, from the same transfected well (run) (e.g. differences between plates 1.1, 1.2, 1.3). Again, some variation of this nature is to be expected. Small (micro-scale) volumes are being handled to seed cells into multiple wells, making it very difficult to distribute cells precisely and evenly. Separate cultures will have inherent biological growth differences.

Differences stemming from slight changes in seeding densities will be exacerbated during culture. There is evidence of plate to plate variation in **Figure 5.15**, the correlation between culture wells seeded from the same transfection well was probed further. Comparisons made between all growth rates, from cultures seeded from the same transfection well, are shown in **Figure 5.16**. The growth rates correlated relatively well between equivalent culture plates (average $r = 0.694$). A high growth rate from a particular transfection tended to rank highly on all three culture plates. This suggested that seeding differences and culture-plate variation were not too considerable and culture plates could be assumed to be equivalent. The primary way this could be reduced was by only seeding one culture well per transfection well. When setting up cultures for static growth, cells were first seeded at the required concentration into larger volumes in 24-well shallow plates, then 100 μL volumes were transferred to separate 96-well micro-

plates for culturing and for post-transfection assessment. An additional benefit of utilising a deep-well culture setup was that only one culture needed to be used. The culture volumes involved were larger and sufficient for multiple output measurements to be made from the one culture; multiple cultures did not need seeding.

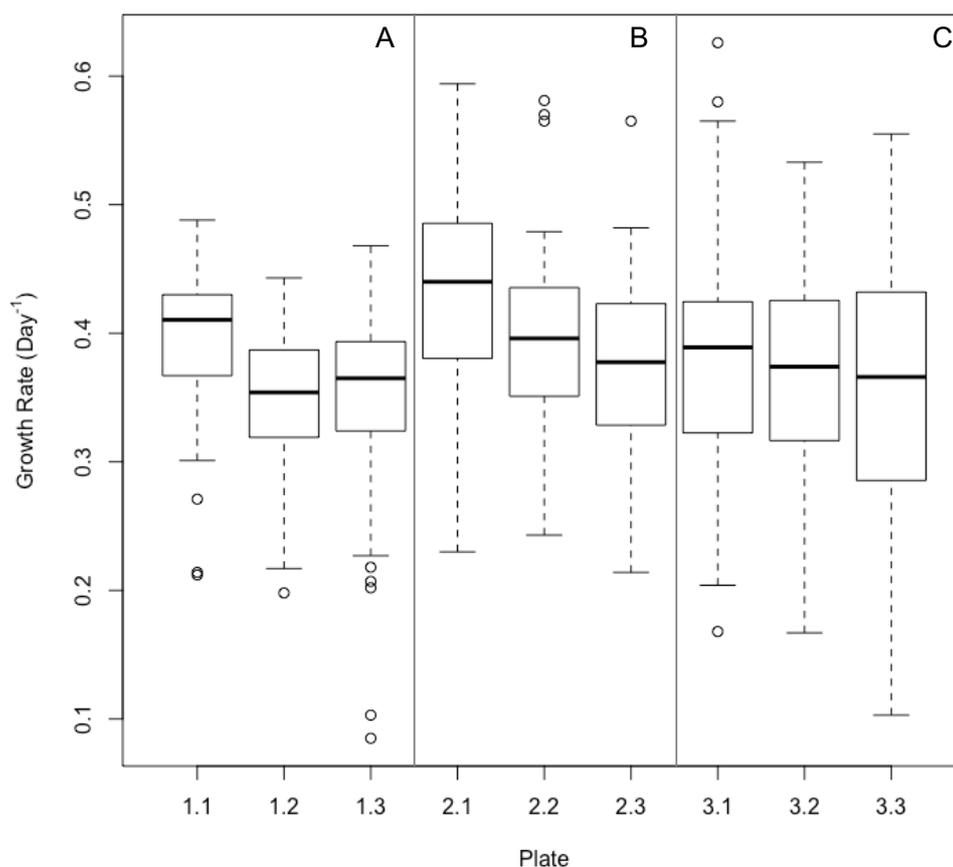


Figure 5.15: Box Plots Representing Growth Rates from Three Separate Whole-Plate Transfection Runs with Each Transfection Seeded into Three Separate Culture Plates

(A), (B), (C) represent culture plates setup from three independent, though identical, transfection runs 1, 2 and 3. For each transfection run, three separate culture wells, each on a separate plate, were established from the same transfection well (e.g. 1.1, 1.2, 1.3). Box plots were generated using R. The bold bars, the boxes and the dashed lines represent the mean, interquartile range and the min/max (range) of growth rates derived from each culture plate, respectively. The circles show outlying growth rates.

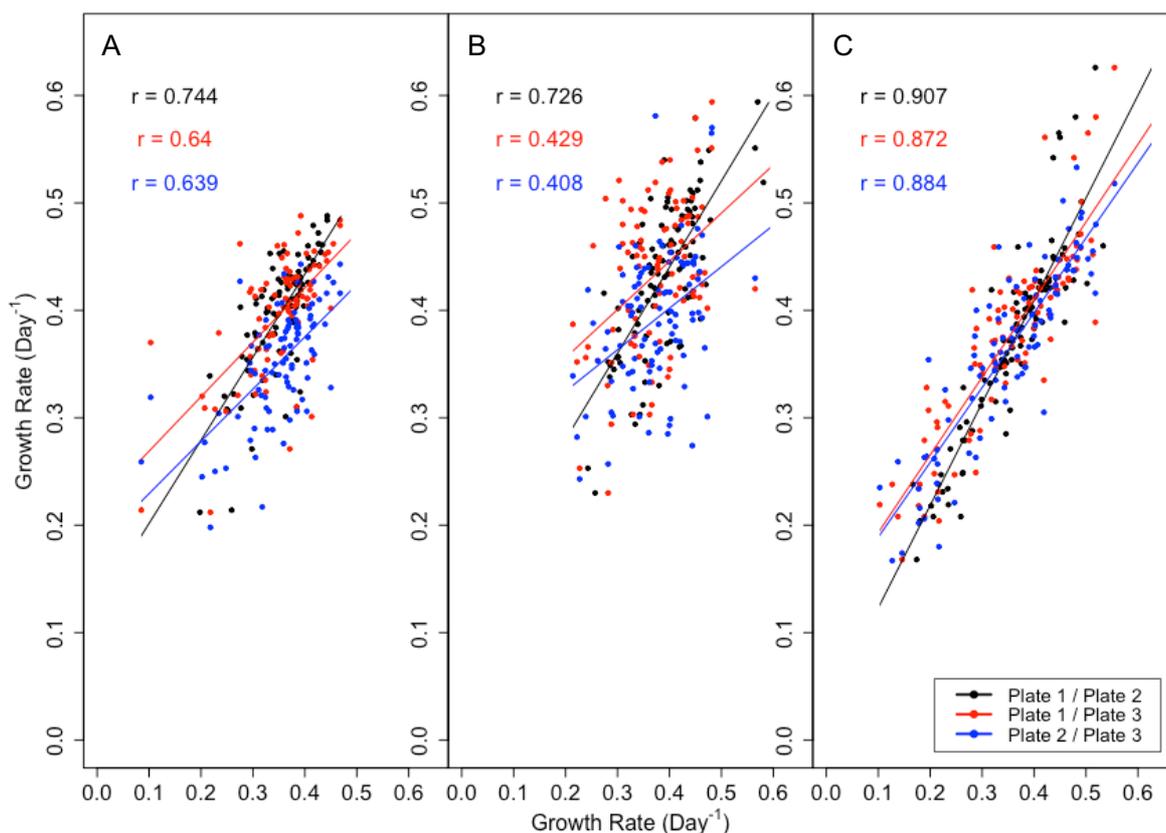


Figure 5.16: Correlations in Growth Rate as A Result of Seeding Multiple Culture Wells from the Same Transfection Well

Pearson's correlation (r) is shown between growth rates derived from seeding separate culture wells (on separate culture plates) from the same transfection well. All correlations were statistically significant (p value < 0.05). The black line and datapoints are growth rates on plate 1 compared to the same 'well' (from the same source transfection) on plate 2 (e.g. 1.1 vs 1.2), the red line/datapoints: from plate 1 and plate 3 and the blue line/datapoints: from plate 2 and plate 3. **(A)** is transfection run one, **(B)** transfection run 2 and **(C)** transfection run 3.

Increasing the number of repeat transfections per plate (number of replicate transfection wells) should have the potential to reduce the variation introduced through well-well transfection differences. To determine the number of replicates that would sufficiently reduce variation, while aiming to maintain as much throughput per plate as possible, the CV was calculated from randomly assigned replicates within each plate. The number of replicates tested ranged from between one and six. The process to generate groups of replicates was undertaken a hundred times to avoid any bias introduced from the random process. The plate CVs calculated were then pooled; it should be noted the level of transfection well-to-well variation (CV) did vary from plate to plate. The average

CVs for different number of replicates, derived from this process, is shown in **Figure 5.17**. Using singular transfection wells, with no replication, resulted in an approximate CV in growth rate of just under 20%. Setting up three transfection wells in-parallel wells (triplicates) would reduce the CV to around 11%, significantly lowering the variation from that present with the use of singular transfections, to a more acceptable level. Increasing the number of replicated transfections per run higher than triplicate did not substantially reduce variation further (for use of six replicates the CV only reduced to approximately 8%). In addition to this, the use of more replicates per transfection run would further reduce the number of parallel comparisons that could be made within a screen. Therefore, it was not deemed worthy to increase the number of replicates above the use of three, to preserve throughput as best as possible within the platform.

In addition to setting up triplicate transfection wells within a transfection run, each transfection run would be set up and undertaken on three separate occasions. This would help reduce biological variation (introduced partly through CHO cell cultures all being inherently different) and transfection and transfection set up variation. In addition to this, an inter-transfection run control within each plate would be included. This control would be cells transfected with the recombinant product gene plasmid, but no functional components (no treatment). This would enable normalisation within each transfection-run and would help to reduce the variation introduced from biological, set up and analysis assay differences.

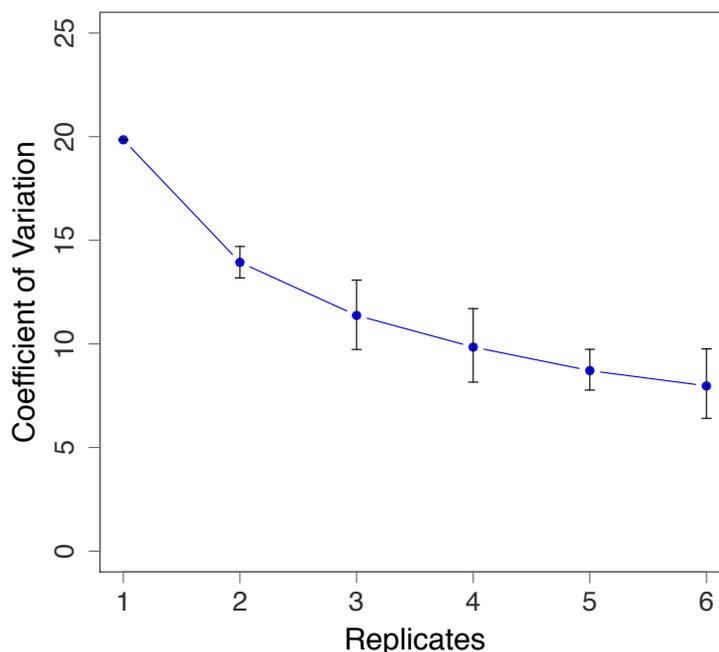


Figure 5.17: Increasing the Number of Transfection Replicates Within a Transfection Run Reduces Transfection Variation (CV)

Well-to-well transfection differences were deemed to be one of the main sources of transfection variation. To reduce this error, replicate transfection wells can be setup per transfection run. The CV derived from using different numbers of replicates is suggested in the figure. The data points represent the average CV calculated from randomly assigned groups of replicates (totalling from one replicate through to six) from each transfection plate (1.1 through to 3.3 from above). The error bars show the 95% confidence interval.

5.4. Discussion

It is clear in the literature as highlighted in **Chapter 3**, that cell engineering strategies have potential substantial benefit to improving CHO cell performance and recombinant protein production. Particularly for the manufacture of DTE proteins, which often have rate limiting processes downstream of translation; likely to be improvable through overexpression engineering of genetic components relating to protein folding, assembly and secretion. There are many genes with potential benefit, some with previous recorded investigation, others untested. Their effect is likely to be dependent on the level of their expression (Hansen et al., 2017). Simultaneous modulation of multiple genetic effectors, a multi-combinatorial solution with components operating in synergy, is likely to be

of additional benefit. To create multi-component synthetic DNA solutions, two fundamental design criteria should be addressed: a selection of the specific and minimal combination of components that operate in synergy to achieve the desired objective; and determination of their optimal relative functional stoichiometry.

To work towards testing the functionality of multiple components including in combinations and at different gene doses and stoichiometry, a screening platform would be of significant benefit and potential (Johari et al., 2015; Pybus et al., 2013). A high-throughput transient transfection platform for standardised, simultaneous assessment of genetic component functionality, suitable for MedImmune cell lines and recombinant products, was the primary objective of this chapter. The platform developed needed to incorporate a transfection method that enabled many (potentially hundreds) of relative transfections in-parallel. To enable this level of throughput, plate-based small-volume transfections appeared to be most appropriate. A plate-based transfection platform of this nature should reduce the time, resource and labour intensity and increase the capability and potential to screen multiple genetic effectors for their functionality. Culture approaches and assays to measure cell culture performance, that are compatible with the micro-scale and throughput aspects, also required consideration. The volumes associated with the setup (100s of μL at most), were limiting and made liquid handling throughout more challenging; there was a need for 96-well compatible approaches with sufficient sensitivity.

5.4.1. Developed Transfection Platform

For clarity, the parameters and outcomes from the platform development are summarised as follows. Within the platform, cells were transfected on day three of culture using the Nucleofection 96-well shuttle add on device with programme FF-158. Approximately 1.4×10^6 cells per well were transfected. To account for discrepancy introduced from the low volumes present within the setup, 2.33×10^6 cells per well, were prepared. Cells were mixed and transfected with 800ng of DNA per well. This DNA was made up of a mixture of recombinant-product expressing plasmid and 'Expression-Plasmid', encoding either genes of interest, or remaining empty of an ORF, to keep DNA loading constant. Immediately after

transfection, cells were seeded at approximately $0.35\text{-}0.4 \times 10^6$ cells mL^{-1} into micro-or deep-well multi-well plates and grown under a static or shaking mode, respectively. After 24 hours, transfection efficiency was above 94% (99% for transfections with the pmaxGFP plasmid). Post-transfection viability by 72 hours was, on average, 85% for co-transfections with mAb-encoding and 'Expression-Plasmid' plasmids (in a 2:1 ratio). The combination of these two post-transfection attributes demonstrated the transfection platform was efficient at delivering DNA to the whole population of cells, while not introducing too many toxic side effects. There was also a high resultant MFI, indicating high levels of gene expression.

The size of the transfected plasmid appeared to affect the efficiency of its introduction to cells and the resulting viability. This needed to be accounted for, as transfections with larger plasmids encoding for multimeric recombinant proteins were essential for this platform. Hence, the DNA load may need to be optimised (lowered) in the future if the platform were to be used for larger multigene transfection, transfections with plasmids potentially encoding for the recombinant gene of interest and an effector gene, or two, all in tandem.

The development of an 'Expression-Plasmid' for the expression of different effector genes of interest enabled standardisation and modulation within the platform. Different genes of interest could be straightforwardly cloned into the same plasmid backbone, this ensured different genes were transfected and expressed at relative levels to each other. The use of separate plasmids introduced screening flexibility, different genes could be straightforwardly interchanged at different levels. This was further investigated and is outlined and described in the next chapter.

A DoE approach was used for part of the optimisation process. Use of DoE-RSM has been shown to be a useful optimisation tool when there is more than one factor requiring optimisation within a transfection approach (Daramola et al., 2014; Mozley et al., 2014; Thompson et al., 2012). The methodology incorporates modelling the responses measured, based on multiple factors input, and analysing the models statistically for significant factors and their variance. Additionally, theoretical optimal combinations of factors could be proposed. Use of the approach enabled both cell number and DNA load to be investigated simultaneously. They were likely to be intrinsically connected. In this scenario,

the amount of DNA being transfected was more significant to the post-transfection responses measured than the cell number transfected. The methodology provided a basis to finalise the transfection parameters once the 'Expression-Plasmid' had been developed. In future optimisations, for example if a different cell line derivative was to be utilised and the platform required different conditions, a DoE approach would be very useful. It would enable multiple factors to be tested systematically and in-parallel, using a reduced number of experimental runs, and would produce statistical analyses and predictions of optimal conditions.

5.4.2. Comparison of Transfection Setup

The developed transfection setup was compared to an alternative high-throughput plate based transient transfection platform, used within DCJ laboratory at The University of Sheffield, to affirm its suitability. The alternative approach was an optimised lipofection-based transfection setup. The two methods differ in their process to introduce DNA into cells. Electroporation harnesses physical, electrical pulses to alter cell membranes, allowing DNA to directly and immediately enter into cells and potentially enter directly into the nucleus. Lipofection uses a cationic chemical to form positively charged complexes with plasmid DNA, these are attracted to, then fuse with cell membranes. The comparison demonstrated electroporation had elevated ability to rapidly introduce plasmid DNA to the whole population of cells and generate high levels of immediate gene expression. Electroporation demonstrated the desired requirements of the transient transfection screen, of interest here, better than the alternative lipofection setup. It is an industry relevant process, for example, electroporation is frequently used for cell line development by MedImmune (MedImmune, personal communication). One of the main limitations of the electroporation approach used here, compared to a chemical based transfection method, was that it was not scalable *per se*. The setup was constrained to the use of 96-well micro-plate format and the associated well size and volume, resulting in 20 μ L transfection volumes. There is potential to increase scale five-fold, though paired with reduced throughput, by utilising Nucleofector 4D technology (Lonza). This technology is deemed to enable directly scalable

transfections to the plate-based technology used here (Amaya, Lonza, personal communication). The platform development was aimed at high-throughput micro-scale transfections but there is some potential if larger transfected volumes are required later on in the engineering process. Larger numbers of transfected cells would be required to progress applying cell engineering strategies to regular biopharmaceutical practises and scales, like those used for SGE, to facilitate more in-depth investigations to be undertaken. There are promising developments enabling larger scale electroporation, for example. Maxcyte technology - involving 1L continuous electroporation (Steger et al., 2015) or the Nucleofactor LV unit (Lonza), which can transfect up to 10^9 cells (in 1mL successive electroporation volumes) in a potentially comparable way to the Nucleofection process used within this chapter.

The developed transfection platform can also be compared to a similar transfection platform recently published, also applicable for screening the effect of many target genes on recombinant protein production in CHO cells (Hansen et al., 2015). Hansen similarly describes a micro-scale, 96-well plate-based transfection approach. The transfection, culture and output response assays all utilise 96-well plate formats and rapid assessments of gene expression are made after two to three days of gene expression. Both the target gene and the recombinant product gene of interest are straightforwardly interchanged between transfections. In contrast to the platform described here, the setup uses lipid-based (Freestyle™ Max reagent) transfection and is aimed at the expression of non-mAb products of human origin. Titre is assessed using an indirect split GFP complementation method, where product genes are modified to contain an S11 tag, any target gene effects of interest would likely need to be validated (partly to ensure the tag sequence does not introduce unwanted side effects). The setup will offer a lower throughput (at least sixteen transfections per gene compared to nine). In contrast here there is a focus upon being able to test and identify multi-functional co-expression of target genes. To work towards this approach, incorporated into the platform, is standardisation, with interchangeable plasmids all utilising a common core backbone (helping to enable testing of gene dosage); transfection viability effects and efficiency are high and consistent; multigene stoichiometry is verified (**Chapter 6**); and a multifactorial DoE approach is

considered as an approach to test combinations of genes (**Chapter 7**). Gaining an indication of how target genes affect growth, important for longer duration and stable production performance, is an important output response in addition to titre. The inclusion of an OriP/EBNA transient expression system will help to enhance and extend TGE culture duration, enabling the TGE system to also be used to test effects later on in culture.

5.4.3. Plate-Based Assays to Assess Post-Transfection Cell Culture Performance

The development of a micro-plate based transfection protocol came with a need for compatible, appropriate assays to measure the effects resulting from the transfection and the overexpression of genetic components. The micro-plate based system comprised of many parallel low-volume cultures, which translated into many low titres and cell numbers to measure. Considering this, any output assays needed to be sufficiently sensitive and maintain compatibility with the 96-well plate format. The primary method to assess viable cell density and viability was a Vi Cell XR cell counter, but this required at least 500 μL of sample, containing at least 1×10^5 cells (to be within the dynamic detection range) and has poor throughput of nine samples at a time, with up to three minutes duration per sample; incompatible with assessing a micro-scale high-throughput culture setup. The alternative method used to assess viable cell populations was a PrestoBlue[®] viability assay. It correlated very well with the viable cell density measured using a Vi cell XR cell counter. It was a scalable plate-based approach; suitable and compatible with 96-well microtitre scale and was very straightforward and rapid to implement, with at least ninety-six 100 μL samples measurable within 35 minutes. A limitation is that it does not provide a direct measurement of VCD or viability. Viability could be assessed separately using a TO-PRO3 viability dye and flow cytometry analysis, but this was an additional assay, requiring additional, separate sample and time to take measurements (>1 hour per 96-well plate). For future platform development, if cost and resource allowed, an improvement would be to implement an alternative, direct method to assess both cell culture VCD and viability simultaneously and rapidly. Technology is emerging to enable this. A high-throughput plate-based microscope system, such as the

Nyone/Cellavista (Synentec, Elmshorn, Germany) would enable automated plate-based brightfield imaging of cultures. Twinned with a trypan blue exclusion method, both VCD and viability measurements can be taken using small sample volumes and only taking a few seconds per well; this method was recently used in study of antioxidant effects within CHO cells (Camire et al., 2017). Alternatively, a Cytonote HT cell counter (Iprasence, Clapiers, France) uses very small (>3 μL) culture volumes to rapidly assess (20 seconds per sample, 48 samples) both VCD and viability. It can simultaneously provide VCD and viability without requiring samples to be stained or labelled. Cell viability is inferred by light diffraction, with living and dead cells having different light diffraction profiles. The very low sample volume, would easily enable multiple assessments of cell culture performance throughout the duration of culture.

The other primary cell culture performance attribute of interest was a measurement of the titres generated. A 96-well FastELISA assay, allowed quick (<1 hour) assessments of IgG levels in culture supernatants. The method required a minimum of 10 μL of sample which needed to be diluted into the dynamic range of the assay. Alternative plate-based approaches to measure accurate IgG levels, if available, could be high-performance liquid chromatography (HPLC), an Octet (Pall ForteBio LLC, Portsmouth, UK), or a ValitaTITER assay (Thompson et al., 2017; Valitacell, Dublin, Ireland). These have potential for more direct IgG quantification, with less sample dilution requirements and increased accuracy and ease of use and throughput. The method chosen needs to be sensitive enough to accurately measure low titres; prevalent in the platform, as titres are being measured early on in culture and the products of interest are likely to be DTE products with inherently low titres. Alternative assays would need to be used if the recombinant product of interest did not contain an Fc region, making it incompatible for detection with the FastELISA approach used here.

5.4.4. Variation Within Platform

The biggest contributor to variation within the setup was between separate transfections (separate wells within a transfection run). The handling of small volumes (total volume per transfection of only 20 μL), was likely to be a major

contributor to the variability inherent within the system. Well-to-well transfection variation was expected. Differences in transfection effect would have occurred when the number of cells and the amount of DNA involved differed slightly, the electroporation pulse and duration may not have been exactly the same between transfections, or the times involved in the process were uneven. When setting up many micro-scale transfections simultaneously it will be very difficult to maintain exact consistency throughout all samples. Differences will occur and will introduce variances in the transfection process, but these should be minimised. To help account for variation and reduce the assay CV, three replicate wells within each transfection plate for each condition being tested would be used. The assay CV could be further reduced with the use of more replicates within a run, but this would further reduce the throughput potential within transfection runs. It was deemed that using more replicates would not be a significant improvement to reducing CV further. Developing a high-throughput screening platform was the priority. The use of high-throughput systems requires a balance between quality and quantity, between their throughput and their sensitivity. Transfections would also be repeated on three occasions, as is common, to help account for inherent biological variation as well as reduce transfection and transfection setup variation.

5.4.5. Developments in Shaking Plate Based Culture

Post-transfection culture, for compatibility and to accommodate the high number of in-parallel sample and small volumes and numbers of cells concerned, was to be in multi-well plates. The most straightforward and widely applicable culture method was use of shallow multi-well plates in a static culture mode. The use of 96-well culture plates maintained compatibility with the transfection setup, but the culture volume, cell densities, batch-mode and duration were significantly limited. The surface area for cell growth was limited, gas transfer poor and evaporation considerable (particularly in outer wells, though these were not used here). A three-day transfection screening setup was possible, but post-transfection culture improvements were of interest and of much potential benefit. With the cell line used here, culture was limited to approximately four days, making capabilities to investigate cellular effects later on in culture very limited.

Per transfection sufficient cells were transfected to enable moderately larger culture volumes to be set up (approximately volumes of up to 2.5 mL if cells are seeded at a concentration of 0.4×10^6 cells mL^{-1}). In-parallel developments, within the DCJ laboratory at The University of Sheffield, to enable a plate-based suspension CHO cell culture format were beneficial here, for the development of the plate-based transfection platform. Shaking plate culture was not straightforward to implement, a more complex culture setup was required. The successful developments enabled transfected cells to be seeded into deep-well 96-well plates and be cultured in a shaking incubator. Humidification and the Duetz system limited, otherwise considerable, culture evaporation and permitted low-culture volumes. The Duetz system clamps kept the plates secure for the high-speed rotation required to keep cells in suspension. The setup provided improved culture conditions; compatible in terms of geometry and scale with the developed electroporation setup. Use of suspension culture has benefits over static culture and is more suited to the culture of the cells of interest here – suspension adapted CHO cells. It provides improved gas transfer and accessibility to nutrients, beneficial for CHO cell growth and productivity and enables higher cell densities and longer duration cultures to be attained with confluency/surface area limitations being less of a problem. The total volumes per well, in deep-well 96-well plates, are up to 2 mL, with working volumes approximately 25% of this volume. These larger volumes would reduce variability otherwise introduced from parallel cultures seeded from the same transfection well; larger culture volumes would allow synchronous assessments of cell growth and titre from the same culture. The increased culture volume would enable alternative cell culture performance attributes to be investigated if required; for example, product quality. To accompany the advancement of shaking-based culture enabling longer culture duration, transfecting MedImmune's transient CHO cell line would facilitate enhanced, longer duration transient transfections. Advantageously, using the developed platform the CHO-T2 cell line transfected equivalently to the MEDI-CHO cell line. This indicates that other MEDI-CHO cell line derivatives could be transfected effectively using the platform, with potentially other CHOK1 derived cell lines also being successful.

5.5. Summary

Here a multi-well micro-plate based electroporation transfection setup has been described and developed. Reported was the optimisation of the platform, capable of hundreds of simultaneous, micro-scale transfections of CHO cell cultures, each with multiple plasmids utilising a common core backbone. The plate-based transfection results in high levels of immediate gene expression, with a high transfection efficiency, surpassing that achieved using an alternative optimised lipofection transfection method. Variability within the approach primarily stemmed from well-to-well transfection variation, likely to be from the handling of the small volumes involved, a disadvantage of a micro-scale setup like this. There were likely to be uneven volumes and cell numbers being handled for example, which would have altered the effect of the transfection process. They could be reduced by sacrificing some throughput by setting up triplicate transfections within each transfection run. There was a need for compatible approaches downstream of transfection within the platform. Low volume, high-throughput approaches to measure VCD and viability simultaneously are developing and becoming more widespread in their use; they should be implemented into the platform when possible to improve the accuracy and ability to measure post-transfection cell culture performance.

The platform incorporates standardised co-transfection of multiple plasmids to express multiple genes simultaneously. In the next chapter the ability for the platform to transfect individual cells at a pre-specified stoichiometry of genes is quantitatively investigated. Subsequently, in **Chapters 7** and **8** potential applications of the developed transfection platform are highlighted. In particular its ability to be used for high-throughput screening of the functionality of an effector gene library, including with genes titrated at different doses, is explored.

6. Validation of Multiple Gene Transfection

For improvement in CHO cell production capability, optimal product-specific solutions are likely to incorporate more than one effector gene, therefore, a unique balance of multiple genetic components will be required. This chapter describes multiple gene expression using the developed transfection platform defined in the previous chapter. A fluorescent reporter approach, combined with flow cytometry single-cell analysis was used to assess multiple gene transfection. The co-transfection of multiple plasmids resulted in individual cells displaying a closer adherence in expression to the ratio in which genes were transfected compared to use of a multigene plasmid, containing all genes of interest. Desirable stoichiometry can be attained by altering the ratio at which the separate plasmids are transfected. From this work, separate plasmids encoding different genes can be co-transfected simultaneously to screen for combinatorial effector gene function.

Acknowledgments

The multigene plasmid used for this work was developed and provided by Yash Patel (University of Sheffield). The initial mCherry plasmid used for proof-of-concept was provided by Dr Andrew Peden (University of Sheffield). The R script used for analysis was developed by Dr Joseph Cartwright (University of Sheffield).

6.1. Introduction

A primary application of the developed transient platform, outlined in **Chapter 5**, is to identify combinations of functional effector genes leading to the generation of cell factories with improved expression capabilities. The platform intends to screen the functionality of numerous effector genes both on their own, at different

doses, and, perhaps of more significance, their effects when used in combination. If multiple effectors are beneficial, there is likely to be an optimal stoichiometry, balancing the relative levels of the individual components. A reliable prediction and generation of stoichiometry is required for model-based DoE optimisation to be applied (Johari et al., 2015; Pybus et al., 2013). As a screening tool the high-throughput platform will enable multiple simultaneous transfections. Prior to investigating the platform for its potential in testing effector gene functionality, the efficiency of multiple plasmid transfection was confirmed. Confidence that the transfection setup is delivering multiple genes to individual cells is investigated along with how gene expression corresponds to the transfected ratio of the multiple plasmids.

To introduce multiple effector genes simultaneously into cells there are two main approaches that could be implemented within the transfection setup. (i) Multiple genes could each be encoded on separate, standardised plasmids, co-transfected together into cells; the simplest, most straightforward approach. Once plasmids are generated to express particular genes, different combinations and levels of genes can be introduced to cells through straightforward adjustment in the combination and amounts of plasmids being transfected. Using this method, there is the potential for a discrepancy between individual cells in the transfection of the different plasmids, and therefore, it is of interest to explore whether cells are all sufficiently receiving the correct mixture of transfected genes. (ii) The main alternative to transfecting multiple genes simultaneously into cells is to introduce all the genes of interest on one plasmid, referred to here as a multigene plasmid. The genes are each expressed in tandem on one plasmid under separate, but identical, promoters. The benefit of this is that only one plasmid needs to be transfected into cells and, if delivered successfully, all genes will be present within transfected cell population. A limitation with this approach is the generation of many plasmid variants that would be required to screen different combinations of effector genes, needing many rounds of time-consuming cloning. Additional implementation would be needed to vary the level of expression of each gene on multigene plasmids, to alter the stoichiometry, something that is not, as it stands, most straightforward. Harnessing precise control of gene expression is only beginning to be attainable and implementable, for example with the progression

of transcriptional control through using synthetic promoters (Brown et al., 2014). The use of separate plasmids is likely to be more feasible to screen libraries and different combinations of effector genes.

Using the developed transient transfection platform, it is unknown how effectively transfecting separate plasmids mimics a singular multigene plasmid to deliver multiple genes robustly into individual cells. The purpose of this chapter is to investigate and validate multigene transfection, with the developed transfection platform, to be used for effector gene screening. The extent of cellular variation in expression of three genes and how tightly individual cells adhere to the transfected ratio is explored. Undertaking analysis at a single cell level gives an insight into how closely individual biological units are conforming or deviating in plasmid expression ratio from the transfected ratio of plasmids. In addition, the aim of the chapter is to confirm that stoichiometry can be robustly achieved by altering the amounts at which separate plasmids, each encoding different reporter genes, are transfected. The variation in the expression of the multiple effector genes needs to be low enough to ensure different stoichiometry can be implemented using this approach.

To investigate these concepts, the approach taken utilised fluorescent proteins; proteins which emit different, measurable wavelengths of fluorescence, when exposed to particular wavelengths of light. They are commonly used as reporters for transfection efficiency and to tag genes, for example for protein subcellular locations to be identified. To mimic simultaneous multiple gene transfection, three different fluorescent protein genes were used. Their different fluorescent emission and excitation spectra enable quantitative analysis of their expression, the fluorescence produced can be measured using flow cytometry. Flow cytometry is a laser-based, powerful and widely used tool incorporating rapid, high-throughput analysis, capable of measuring the fluorescence of thousands of individual events (cells) in a matter of seconds. Multiple parameters can be measured concurrently, and it is commonly used to quantify the expression of intracellular and cell surface molecules. Use of this approach, easily allows the level of expression of the three fluorescent reports in individual cells to be measured, and how these correspond to the plasmid ratio transfected.

The 'Expression-Plasmid', outlined in **Section 5.3.1** was used to separately encode each of the three fluorescent reporter genes. These plasmids were co-transfected into cells at equal and varying ratios to each other. For comparison, a multigene plasmid, encoding the three fluorescent proteins in tandem, was also transfected. In all cases the resulting total plasmid DNA concentration was kept the same to try keep the dynamics of the electroporation process consistent including the transfection efficiency and cytotoxicity introduced. The level of fluorescence was measured, using flow cytometry, after 24 hours, to allow time for sufficient gene expression. Cellular fluorescence intensity values provided an insight into how individual cells conformed in the expression of the ratio of genes in comparison to the ratio at which they were transfected. The analysis, of tens of thousands of cells, was undertaken using R software.

Introducing genes on separate plasmids could result in a discrepancy of individual cells not receiving all three plasmids. The variation in relative expression of the three fluorescent reporters between individual cells was also explored using single cell analysis. To confirm if stoichiometry is reliably achieved, using separate plasmids, the ratio of plasmids transfected was altered and analysed.

6.2. Experimental Approach

To study the concept of multigene transfection within the developed transient platform a range of transfections were undertaken, summarised in **Table 6.1**. In all cases the total concentration of plasmid DNA transfected was kept constant at the previously optimised amount of 800 ng (**Section 5.3.2.2**). To achieve this, an empty-'Expression-Plasmid' containing a CMV promoter, but no ORF, was used where needed. To ensure sufficient gene expression fluorescence levels were analysed, using flow cytometry, after 24 hours. Different amounts (0-800 ng) of each fluorescent protein encoding-plasmid were set up to check the dynamics of expression and detection of each fluorescent protein, and to aid further analysis. Co-transfection, with an equal loading (1:1:1 ratio) of each fluorescent protein plasmid with and without the addition of recombinant protein (DTE-mAb) expression (referred to as Triple and MabTriple, respectively), was undertaken to form the basis of the multiple gene co-transfection analysis.

Transfections with a multigene singular plasmid, encoding all three fluorescent proteins in tandem, each under separate CMV promoters, were undertaken to compare the dynamics and variance from a multigene plasmid transfection approach. Employing multiple separate plasmids and co-transfecting them into cells in combinations allows for the ratio of plasmids being transfected to be easily altered. To ensure stoichiometry is being achieved using this approach, and to investigate inherent variation, several transfections with different combinations of each fluorescent protein encoding plasmid were undertaken.

Table 6.1: Transfection Setup Used to Investigate Multigene Transfection

The amount and combination of genes used in each transfection is highlighted. With the amount of plasmid used in ng (total 800 ng). The ratio of fluorescent protein plasmids used are highlighted with TagBFP (blue), eGFP (green) and mCherry (red).

Transfection	TagBFP (ng)	eGFP (ng)	mCherry (ng)	Empty- 'Expression- Plasmid' (ng)	DTE- mAb (ng)	Multigene (ng)
Triple (1:1:1)	177.8	177.8	177.8	266.7	0	0
mAbTriple (1:1:1)	177.8	177.8	177.8	0	266.7	0
Multigene	0	0	0	266.7	0	533.3
Triple (1:2:4)	76.2	152.4	304.8	266.7	0	0
Triple (3:2:1)	266.7	177.8	88.9	266.7	0	0
Triple (1:2:1)	133.3	266.7	133.3	266.7	0	0
Triple (3:1:1)	320.0	106.7	106.7	266.7	0	0
Triple (1:3:1)	106.7	320.0	106.7	266.7	0	0
Triple (1:1:3)	106.7	106.7	320.0	266.7	0	0

6.2.1. Fluorescent Protein Expression Plasmids

For reference the nucleotide sequences for the three fluorescent proteins used for this work can be found in **Appendix Table B.1**. The sequences were optimised for *C. griseus* expression and modified as described in **Section 5.3.1.2**.

Each gene was cloned using *AgeI* and *PstI* restriction sites into the expression-plasmid and sequence verified to ensure integrity, like that described in detail in **Section 5.3.1.2**. Plasmid maps of the final plasmids are shown in **Figure 6.1**. The length of gene insert encoding each fluorescent protein was approximately the same, in terms of number of base pairs, for all three proteins. This means that all three plasmids were approximately equal in size (between 6020 and 6038 bp). The same weight (in ng) of DNA has approximately equal numbers of copies of each of the plasmids. It could be assumed there was no difference or bias in delivery between the three plasmids and the same transfection efficiency.

Figure 6.1 also shows the multigene plasmid used. The expression components on the plasmid were identical for each gene. The genes for the fluorescent proteins were identical in sequence to those on the single gene plasmids and, similarly, were all downstream of a MedImmune derived CMV promoter. On the multigene plasmid, the eGFP gene was positioned upstream of a TagBFP gene, which was upstream of an mCherry gene. Along with the coding regions of the three fluorescent protein genes, the plasmid differs from the single expression plasmids in that it encodes GS, enabling stable expression if required with the use of MSX selection. Cells transfected with this plasmid has the additional 'burden' of expressing the GS protein. The multigene plasmid was 11.5 kb in length, approximately twice as large in size as the single expression plasmids, but not that dissimilar compared to some recombinant protein encoding plasmids (MedImmune, personal communication).

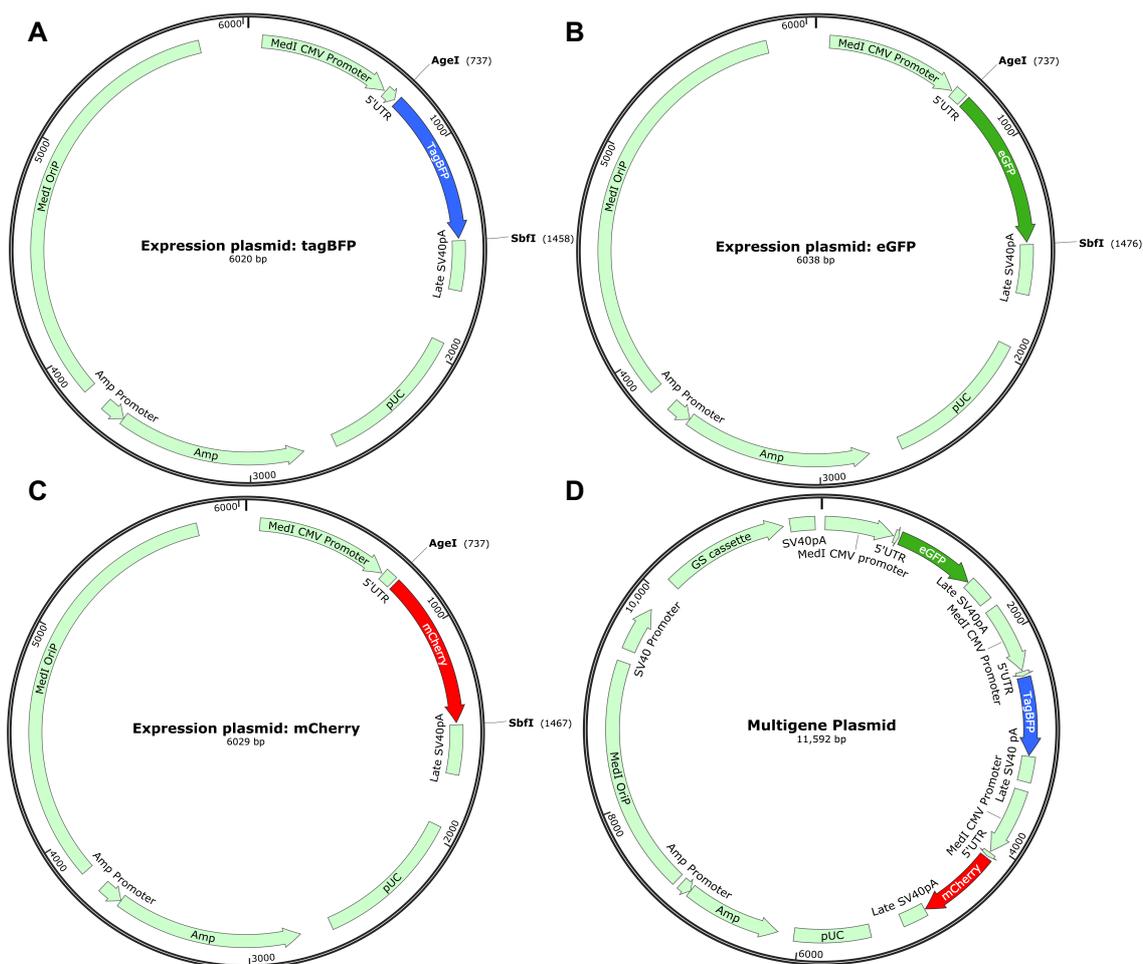


Figure 6.1: Fluorescent Protein Plasmid Maps

The plasmid maps encoding (A) TagBFP (6020 bp) (B) eGFP (6038 bp) and (C) mCherry (6029 bp) within the previously described ‘Expression-Plasmid’, with the total plasmid length in brackets. The sequences encoding the three fluorescent proteins were also expressed in-tandem in a multigene plasmid (11592 bp); all genes expressed under separate, but identical components, including CMV promoters.

6.3. Results

6.3.1. Development of the Fluorescent Reporter Strategy

Fluorescent proteins are prevalent tools for measuring gene expression and were employed for this work as reporters to enable analysis of plasmid expression (inference of plasmid delivery). Three fluorescent proteins, with distinctly separate excitation and emission spectra, were desirable. This was to enable independent measurement and analysis of the levels of the proteins when they were multiplexed and co-expressed together.

The scope for the combination of fluorescent proteins used was limited by the flow cytometer optic availability (laser and detection setup). The emission wavelengths of the fluorescent proteins needed to be matched with the detection filters available, and the use of narrow filters minimises detection overlap. The range of optics on flow cytometers is increasing, but many basic flow cytometers primarily only have blue (488 nm) lasers and a combination of violet (~405 nm), and/or red (~633 nm) lasers, limiting the proteins that can be excited and measured. The optics available were a major consideration in selecting which fluorescent proteins were used in the experimental design. The fluorescent proteins chosen should also ideally have some similarity in structure, to minimise differences being introduced from dissimilar rates of folding and assembly; use of monomeric proteins would reduce this. Proteins with similar brightness would improve comparability and detectability. Proteins with prior mammalian expression were only considered, to ensure their suitability and absence of cytotoxicity. In general, fluorescent proteins have broad spectral fluorescence curves, over a wide wavelength range. To reduce spectral overlap, proteins with spread out, and to the best extent, separate spectra, with one another were sought. Harnessing separate lasers to excite each protein at different parts of the fluorescence spectrum would help to prevent overlap of detection in different filters.

The original and most routinely used fluorescent reporter protein is GFP. An engineered version, eGFP, was altered for optimal flow cytometry use (Cubitt et al., 1995; Heim et al., 1995; Telford et al., 2012), it is maximally excited and efficiently detected using widely available flow cytometry optic, a 488 nm laser. The bright protein monomer has maximal excitation and emission at 488 nm and 507 nm, respectively. From previous work, including here in the development of the high-throughput transfection platform, it can be successfully expressed in CHO cells, and with the gene construct already available, eGFP was a practical and clear fluorescent protein to use for multiplex co-expression investigations.

In the flow cytometer used for analysis, a 405 nm laser was present, to harness this a protein excited close to 405 nm was sought. TagBFP (mTagBFP), a monomeric blue fluorescent protein (BFP), with maximal excitation and emission at 402 nm and 457 nm, respectively, met the experimental design criteria. It is a

brighter protein than similar alternatives (e.g. EBFP2), which is particularly important due to the prevalence of autofluorescence around the associated emission wavelengths. There was evidence (Evrogen, 2008; Subach et al., 2008) of successful mammalian expression, without a cytotoxic effect, and minimal overlap in spectra to eGFP. Other fluorescent proteins excited by a 405 nm laser; cyan fluorescent proteins, have an increased overlap in emission fluorescence with eGFP, than that for TagBFP.

The use of a protein excited by a red laser (633 nm) would enable good separation, in terms of fluorescence excitation and emission, from that of TagBFP and eGFP. Unfortunately, very few far-red fluorescent proteins have been characterised. E2-crimson (Strack et al., 2009) is a far-red protein with maximum excitation and emission at 611 nm and 646 nm, respectively. The fluorescence spectra for the protein can be found in **Appendix Figure B.1**. It is described as the one of the most efficiently excited (by a 633 nm laser) far-red fluorescent proteins available (Hawley et al., 2017; Strack et al., 2009). Other far-red proteins would only be excited at low levels, at approximately 10% efficiency. E2-crimson is a bright protein, with evidence of mammalian expression and no cytotoxicity, however, the protein disadvantageously is a tetramer. When testing E2-crimson expression within this system, it became apparent there were problems when multiplexing E2-crimson and TagBFP. E2-crimson was undesirably excited, at low levels, by the 405 nm laser. This was combined with, without prior indication, fluorescence emission at low wavelengths, detected by the 450/50 bandwidth filter, used for TagBFP detection. An example of this problematic overlap is shown in **Figure 6.2**. The substantial overlap could not be removed by compensation, a standard process (mathematical method) used to correct for spectral overlap when measuring fluorescence. It was speculated that the overlap was attributable an increased autofluorescence induced by the transfection process; with increased autofluorescence not removed with compensation (Guglielmo Rosignoli, MedImmune, personal communication). Detection of E2-crimson when co-expressed with TagBFP could not be suitably separated, so a different fluorescent protein was required.

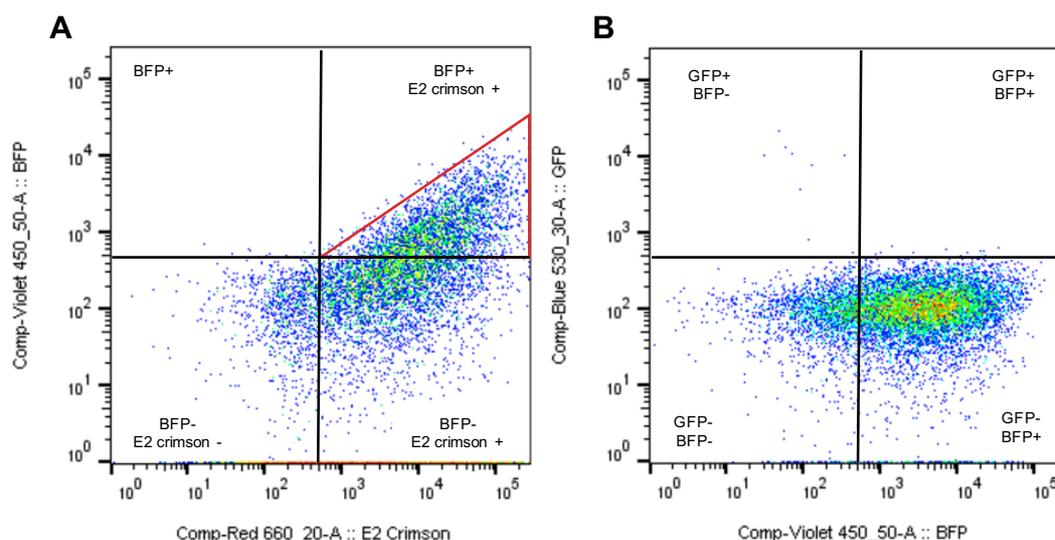


Figure 6.2: E2-Crimson is Not a Suitable Reporter Protein for Multiplexing with TagBFP Expression

Flow cytometry analysis from individual expression of E2-crimson (**A**) and TagBFP (**B**). For E2-crimson expression (**A**) there should have only been events in the bottom half of the plot (BFP-), as demonstrated by TagBFP expression in (**B**). When excited with a 405 nm laser and detected with the 450/50 filter, used for TagBFP excitation and emission, there was substantial E2-crimson fluorescence detected in the 450/50 filter (red triangle), with large amounts of false positives for TagBFP expression introduced. This spillover in fluorescence could not be removed through compensation. TagBFP and E2-crimson expression could not be independently analysed when co-expressed.

There did not appear to be any definitive far-red fluorescent protein alternatives to E2-crimson, so it was decided that a different flow cytometer configuration, not utilising a 633 nm laser, was required. A flow cytometer with a 532 nm laser (MedImmune) was available, and a plasmid encoding mCherry. mCherry is a monomer, widely utilised and often multiplexed with GFP. It is a bright protein, with good photostability and suitable for mammalian expression without cytotoxic effects. Its excitation and emission maxima are 587 nm and 610 nm, respectively; but it can be sufficiently excited using a green 532 nm laser. Beneficially there was little excitation, and overlap, with the 488 nm laser used for eGFP excitation. In summary, the three fluorescent reporter proteins used for this study, were TagBFP, eGFP and mCherry. Their excitation and emission spectra, and the flow cytometer configuration used, is shown in **Figure 6.3**. Small amounts of compensation were applied to overcome minor spectral overlaps.

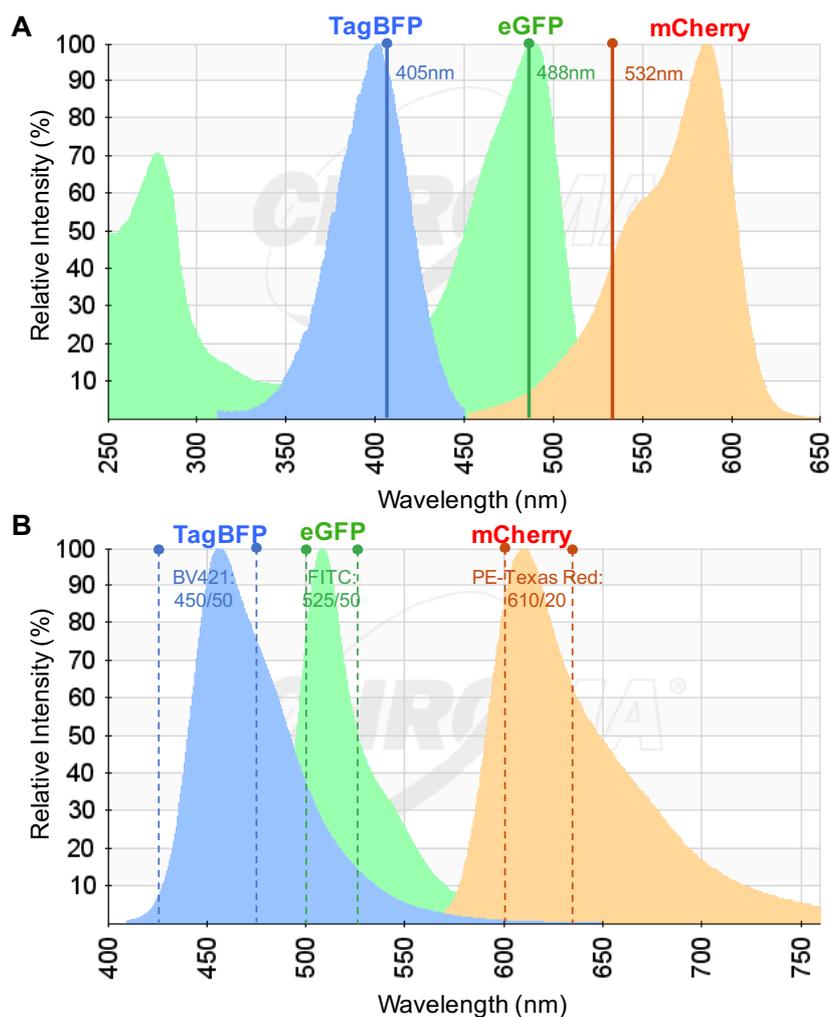


Figure 6.3: Excitation and Emission Spectra and Flow Cytometer Configuration Used for TagBFP, eGFP and mCherry Detection

(A) The excitation spectra for the three fluorescent proteins and the wavelength of the three lasers used for excitation of each protein. (B) The emission spectra for the three fluorescent proteins with the bandpass filters used for detection. Adapted figure from Chroma Spectra Viewer (www.chroma.com, Chroma® Technology Corporation, Vermont, USA).

6.3.2. Fluorescent Protein Flow Cytometry Detection and Analysis

Flow cytometry allows for multi-parametric analysis of single cells in a rapid and quantitative manner. Gating strategies are used for effective and correct analysis of flow cytometry data and enables only the cells (events) of interest to be analysed. The gating approach used here, outlined in **Figure 6.4** and described

below, enabled isolation of cells positive for all three fluorescence parameters. These triple-positive events were then analysed further.

The gating strategy used eliminated events of no interest; fragmented cells and debris, from further analysis (approximately 25% of events eliminated; **Figure 6.4A**). These events were identified based upon their size and complexity/granulation (FSC and SSC), inferred using a mock-transfected population to identify healthy cells. The Fortessa flow cytometer (BD BioSciences), used for data collection, can result in clusters of cells and cell doublets being analysed as a single event, not just single cells. As flow cytometry analysis of single cells is of particular significance for this investigation, the doublets/clusters were removed from further analysis based upon their size and time they take to pass the laser (represented by the area of the pulse produced – side scatter area (SSC-A); approximately <10% of events eliminated). Doublets/cell clumps become evident when the area of the pulse is displayed against the height of the pulse (side scatter height (SSC-H)) generated by the same event, as seen in **Figure 6.4B**. Normal cells have a better correlation in terms of their SSC-A and SSC-H, cells outside this correlation were excluded.

To ensure viable cells were analysed, a viability dye (DAPI) was added to samples immediately prior to analysis. The DAPI dye enters compromised membranes, of dead or dying cells, and binds to double stranded nucleic acid (DNA). When bound it emits a strong blue fluorescent signal, excitable by a 355 nm laser and detectable by a 450/50 bandpass filter, enabling exclusion from further analysis (**Figure 6.4B**; approximately <7% of events eliminated). The overlap from using DAPI with TagBFP was removed by applying low amounts of compensation.

To isolate the triple positive fluorescent population of cells, negative control cells, mock-transfected with empty-‘Expression-Plasmid’ were used to infer positive populations. To overcome the issue of spectral overlap, explained subsequently in **Section 6.3.2.1**, the peak generated from autofluorescence of mock-transfected cells was positioned at an approximate relative fluorescence intensity of 1 (graphically represented as 10^0) for all fluorescent parameters (BV421 (TagBFP), FITC (eGFP), PE-Texas Red (mCherry)). Gates were positioned to infer and measure positive events for each fluorescent parameter, as seen in

Figure 6.4 D, E. Transfections with each separate fluorescent protein encoding plasmid were included in every flow cytometry run (equivalent to single ‘stained’ samples) to ensure fluorescence was not being detected in the incorrect channel for each parameter and to enable any small levels of compensation required to be applied (**Appendix Figure B.2**). This removed any undesirable fluorescence overlap and prevented false positives being analysed. Each fluorescent protein should be only detected in its corresponding filter (parameter). For example, the occurrence of TagBFP expression should only shift populations along the BV421 axis, and not along any of the other axes (demonstrated for all proteins in **Appendix Figure B.2**).

To ensure only definite triple positive populations were being analysed, cells were first isolated based on their positivity for mCherry fluorescence (approximately <60% of events eliminated), then their positivity for TagBFP and eGFP fluorescence (approximately <10% of events eliminated). This is shown in **Figure 6.4F**, with gating hierarchy (backgating) outlining the gating and population ancestry of how the final triple positive populations were inferred. For further analysis, the population-level fluorescence statistics (MFI) and the fluorescence intensity values for each parameter measured for each cell (event) within the triple positive population (approximately 40% of viable event population) were exported from the flow cytometry analysis software. The recorded relative fluorescence intensities (scale vales) for all the compensated parameters were exported in CSV formats.

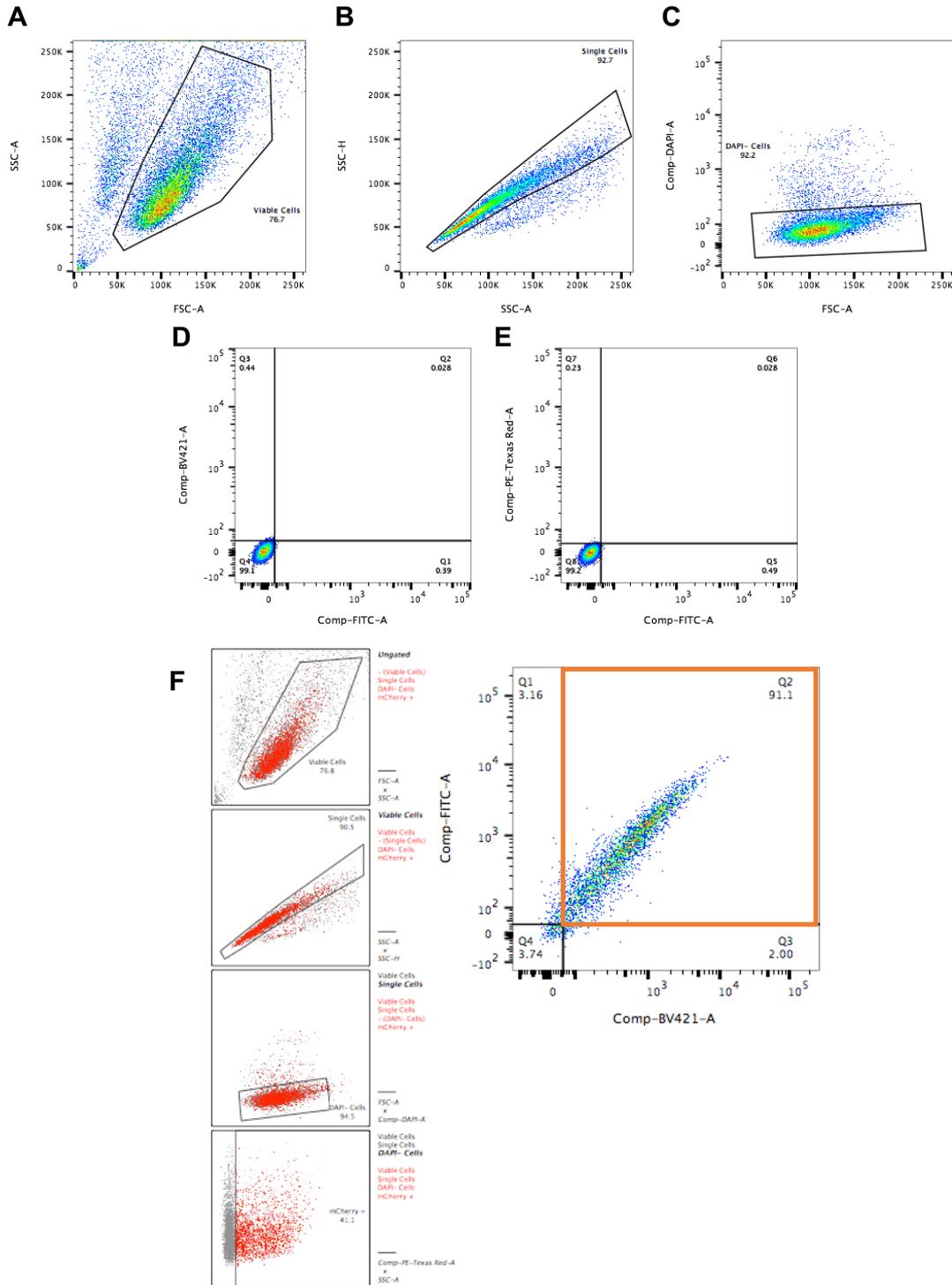


Figure 6.4: Flow Cytometry Gating to Isolate Triple Fluorescent Events

(A) Debris/dead cells were removed based upon their size (FSC-A) and granulation (SSC-A). (B) Single cells were isolated, based upon granulation, to remove cell clusters/doublets from further analysis. (C) A DAPI viability dye isolated viable cells. (D, E) Mock-transfected cells, used to position and gate non-fluorescing populations to infer positive populations. (F) Triple positive fluorescent events (Q2 - orange box), isolated based upon PE-Texas Red (mCherry) signal; then based upon both a BV421 (BFP) and a FITC (GFP) signal. The full gating approach is summarised through backgating plots to the left in (F) (ancestry of the triple population in red).

6.3.2.1. Challenges and Limitations in the Detection of Three Fluorescent Proteins

When eGFP expression was measured using flow cytometry there was found to be overlap emission, in the BV421 filter used to measure TagBFP expression, at high fluorescence intensities. This can be seen by the right-hand bend in the population in **Figure 6.5A** at high GFP intensities. The overlap could not be removed through compensation, the normal approach to overcoming fluorescence spectral overlap. The overlap may have been due to the fluorescence levels being very high and out of the linear range of detection for the flow cytometer used (MedImmune, personal communication). This is an undesirable scenario, as an overlap in fluorescence would introduce false positive TagBFP (BV421) fluorescence. To overcome the overlap, the voltage used for eGFP excitation was reduced. This reduced the sensitivity and lowered the fluorescence intensity detected in the FITC channel; an example of which is shown in **Figure 6.5B**. This approach introduced false negative results, more events were deemed as 'negative'; with their fluorescence intensity measured similar to levels of autofluorescence and located in the negative gate (Q8). As a result, this reduced the sensitivity and affected the measured transfection efficiency (percentage of the population expressing the transfected genes), but it ensured that only true positive fluorescent events were measured. This was deemed a more appropriate scenario, than introducing unknown false positives. For consistency, the voltages of all three lasers were lowered so the highest recorded events within each parameter were located at approximately 10^4 fluorescence intensity units.

There was a very small spill over of fluorescence from mCherry into the BV421 channel (**Figure 6.5C**). This was only evident at higher fluorescence intensities, derived from transfections involving 400-800 ng of mCherry-encoding plasmid (higher than that used in the majority of transfections). The overlap was only present at very low levels (<4% of the population detected) and could not be removed through compensation. It must be noted there is a chance of a small amount of false positive TagBFP signal introduced from mCherry excitation.

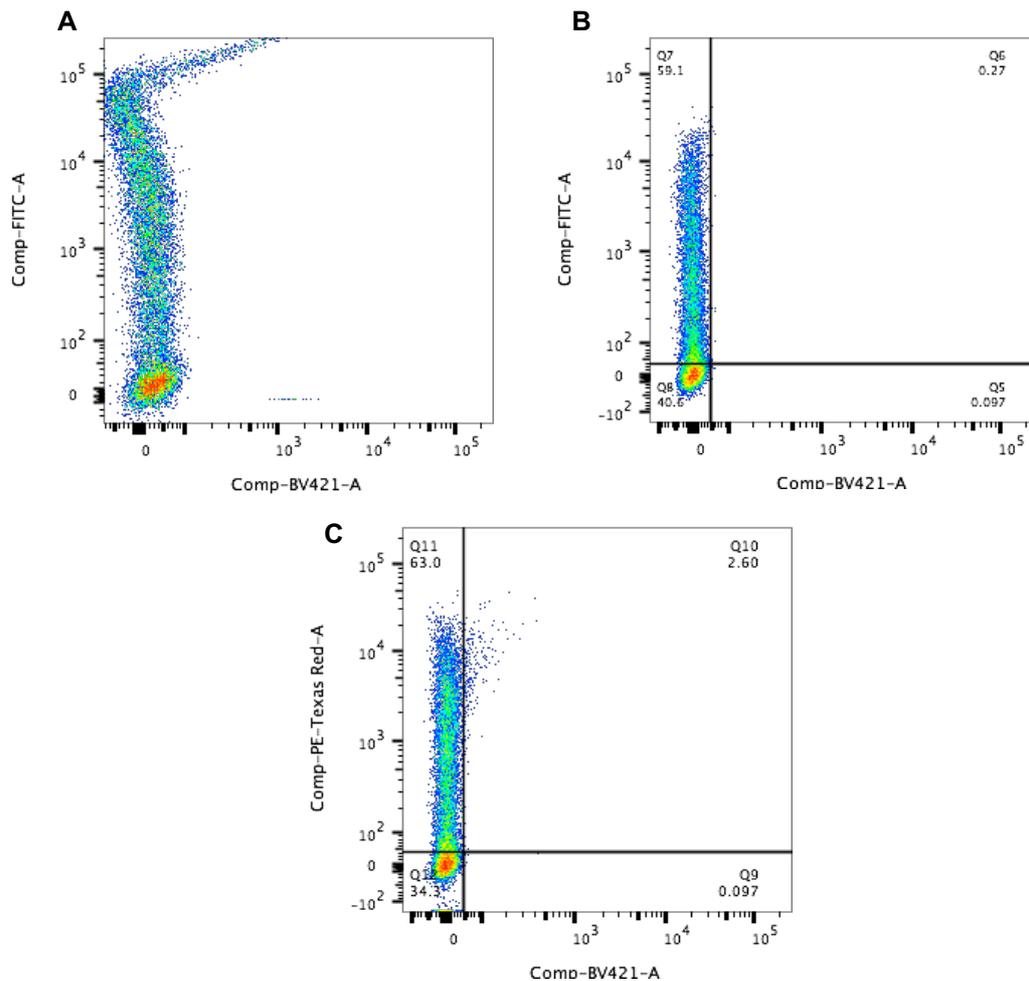


Figure 6.5: Limitations in Detection from Multiplexing Fluorescent Protein Reporters

At high fluorescence intensities ($>10^4$) eGFP fluorescence bled into the BV421 channel, introducing false positive BV421 fluorescence, seen by the right-hand bend (A). Lowering the voltages to position the highest FITC fluorescent events, at just below 10^4 fluorescence units, prevented BV421 bleed through (absence of right-hand bend). (B) At relatively high fluorescent intensities (10^3 - 10^4) there was a small level ($<4\%$) of spill over of mCherry fluorescence into the BV421 channel, introducing potentially small levels of false positive TagBFP fluorescence. The amount of overlap was deemed negligible and not significant to affect subsequent analysis (C).

6.3.3. Normalisation to Adjust for Inherent Differences Between Fluorescent Reporter Proteins

As outlined in the previous section the fluorescent proteins, particularly mCherry, were not all equally or optimally excited and detected. These disparities in quantification need to be minimised to enable relative comparisons of each

fluorescent protein to be made and allow multigene transfection to be studied using this experimental approach.

Transfections with each individual fluorescent protein, at a range of plasmid loads, provided an indication of the differences in the relationship between transfection DNA load and fluorescence levels between the three fluorescent proteins. The MFI from different amounts of each plasmid are shown in **Figure 6.6**. The titrations demonstrate there is a linear relationship between DNA load and fluorescence intensity, and that the setup is sensitive to identify changes in expression using a DNA load between 50 ng and 800 ng.

The titrations show for eGFP the same amount of DNA produces a greater fluorescence, elevated beyond that of TagBFP and mCherry; it appears as a 'brighter' protein. Fluorescent proteins are known to have different levels of brightness; derived from the product of the proteins extinction coefficient and its quantum yield (Hawley et al., 2017; Shaner et al., 2005). As the lasers and detectors are not necessarily optimal or equal between the proteins, excitation and detection is not equivalent between the three proteins. This is particularly apparent for mCherry, as the laser used for excitation is not near to maximal (**Figure 6.3**), one of the limitations of the flow cytometry setup. There is variation and disparity introduced with how the flow cytometer is set up; the alterations to prevent the problematic overlap at high fluorescence intensities, the compensation applied, and the gating strategy used. In order for expression based comparisons between the three proteins to be made it is important to account for this discrepancy to make fluorescence intensity levels more directly comparable.

To improve in making direct comparisons between expression levels of the three fluorescent proteins, disparities from the flow cytometry setup were standardised to reduce their prevalence. Statistics of triple fluorescent populations (mean, median and CV), were used to gauge the relationship in the detection set up between the three proteins. On average eGFP was 1.5 and 1.52 time brighter than mCherry and TagBFP, respectively. For subsequent analysis, these values were used to adjust for the differences in brightness between the different proteins. For each triple positive event, the raw fluorescence intensities, from the three channels, were adjusted to account for this relative difference in eGFP

detection. All raw BV421 (TagBFP) values were adjusted by 1.52X and PE-Texas Red (mCherry) values by 1.5X. This adjustment improves single cell fluorescence analysis, normalising for fluorescence and detection differences between the three proteins so expression levels can be evaluated more comparatively.

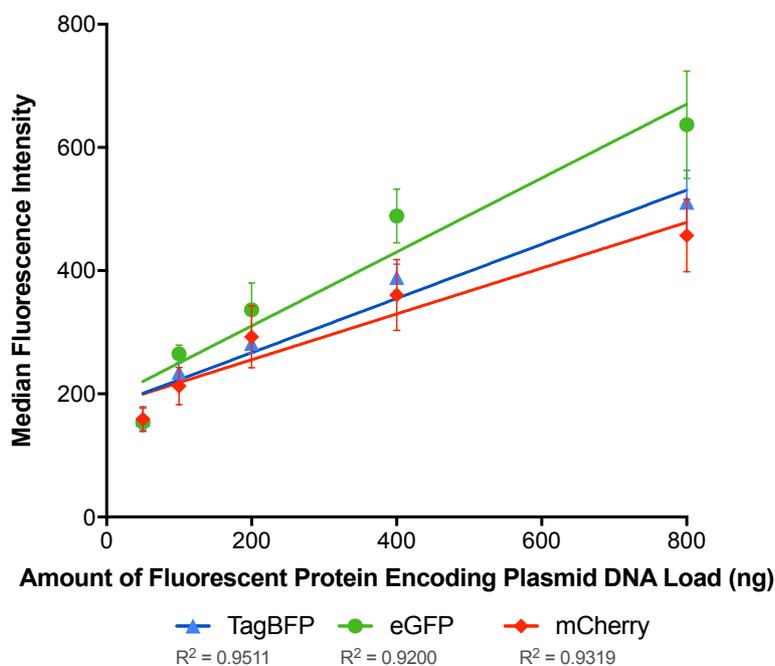


Figure 6.6: Fluorescent Reporter Plasmid Titration

The relationship between the amount of fluorescent protein-encoding plasmid transfected into cells and the MFI (measured within the positive gate for each fluorescent protein) is shown for TagBFP (blue), eGFP (green) and mCherry (red). Empty expression-plasmid was used to maintain total DNA load transfected. The data points represent the average and SEM of three biological and three technical replicates.

From the development of the high-throughput transfection platform, the data demonstrate the transfection efficiency, is consistent between transfections, shown in **Figure 6.7A**. The transfection efficiencies of two different GFP encoding plasmids, pmaxGFP and eGFP-‘Expression-Plasmid’, across eleven different transfections, all with 800 ng of DNA, are shown. The average transfection efficiency is high for transfections with either plasmid, 99.2% and 92.4%, for pmaxGFP and eGFP respectively, and the variation is fairly small, represented with low CV values of 0.5% and 3.6% respectively. The three plasmids used for this work, encode the eGFP, TagBFP or mCherry, outlined in **Figure 6.1**, are very similar in size, with at most 18 base-pair differences in size. Each plasmid comprises of the same identical backbone (‘Expression-Plasmid’), resulting in

sequence similarity of at least 88%. All proteins have identical promoter and polyA components. It is assumed, as long as the plasmids are well mixed, plasmid uptake and expression, on average, are equal between all plasmids. At the population-wide level it is assumed the average ratio in expression will be the same ratio as that delivered to cells during the transfection setup, with no bias between the plasmids used here. This is validated by an example of simultaneous measurement of the three plasmids using a different flow cytometer (FACs-Aria (BD Sciences); available for temporary use at MedImmune). The FACs-Aria has a 561 nm laser, which enables more optimal mCherry excitation and detection. It did not create the overlap problems between eGFP and TagBFP, meaning voltages and therefore sensitivity did not have to be reduced. Demonstrated in **Figure 6.7B**, is the high transfection efficiency attainable for all three fluorescent proteins (90.9%, 87.4% and 88.1% for TagBFP, eGFP, mCherry, respectively) with a high transfection efficiency for all three genes within the same cell being 88%.

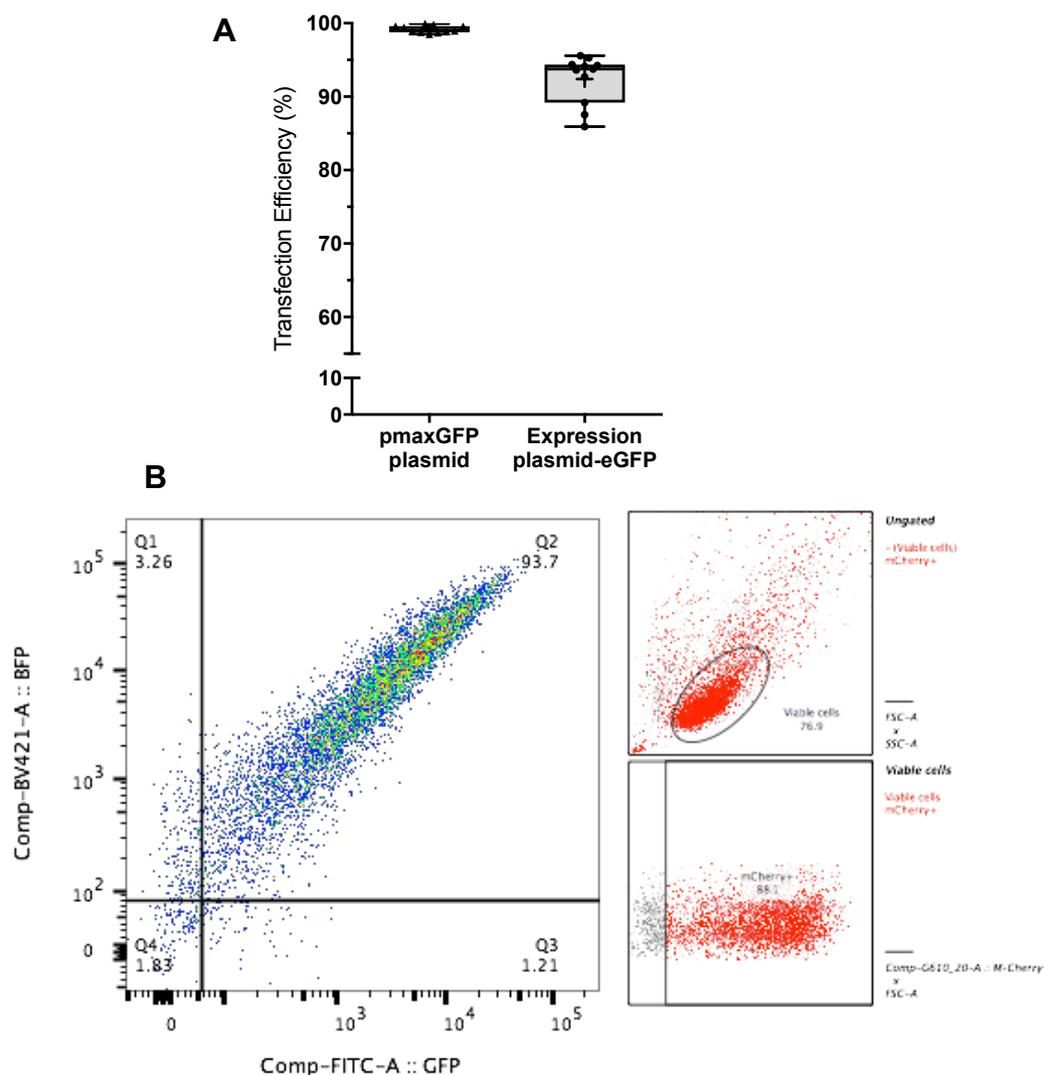


Figure 6.7: Transfection Efficiency is Consistent Between Transfections and Plasmids

(A) Box and whisker plots representing the transfection efficiency (% GFP expression) after 24 hours, measured from eleven transfections with pmaxGFP plasmid and eleven transfections with expression-plasmid encoding eGFP, all transfections had equal plasmid loading of 800 ng. The mean transfection efficiency (and CV) from transfections with pmaxGFP and expression-plasmid: eGFP are 99.2% (0.5%) and 92.4% (3.6%), respectively. (B) Example of a scatter plot and backgating from an alternative, better suited flow cytometry setup (FACs-Aria) for detection and analysis of the three fluorescent proteins, demonstrating high transfection efficiency is being attained, and is approximately equal, for all three plasmids.

There is a wide range, over several orders of magnitude of fluorescence intensity values measured (10^0 - 10^4), exemplified with the logarithmic scaling on the flow cytometry plots shown in **Figure 6.4**. This is primarily derived from cells taking up different amounts of total plasmid during transfection, resulting in wide ranges

of total protein expression. To account for the wide ranges of fluorescence intensities for analysis, to enable direct comparisons in relative expression between individual cells to be made, the three fluorescence intensities for each cell were represented as ratios of one another. An example of this is shown of this in **Table 6.2**. To convert all the ratios into a similar scale (x:1:y ratio), the median (middle fluorescent value) for each cell (highlighted in bold and underlined in **Table 6.2**) was used as the common denominator. All three fluorescence values in each cell were normalised against their median value. This put the dataset into an x:1:y ratio format, permitting comparisons of how each individual cell differs in ratios, of the three fluorescent proteins, to the transfected equal ratio (1:1:1) of the plasmids.

Table 6.2 Example of Normalisation Undertaken to Adjust for Total Fluorescence and Convert the Dataset into Comparable Ratios

The median (middle) fluorescence intensity (FI) for each cell was used to convert cellular fluorescence intensity into comparable terms, into a ratio (x:1:y) of each other, to help overcome differences in total cellular fluorescence across the dataset. The three fluorescence intensity values measured for each cell were divided by the median value for that cell (underlined and in bold for each cell (row)), converting the dataset into comparable terms (ratios).

Event (Cell)	BFP FI	eGFP FI	mCherry FI
1	<u>95.8</u>	72.2	110.9
2	3.1	<u>4.2</u>	4.7
3	<u>1003.4</u>	898.5	1219.7

↓ Normalisation using MFI for each cell ↓

Event (Cell)	Normalised BFP	Normalised eGFP	Normalised mCherry
1	<u>1.000</u>	0.7754	1.158
2	0.738	<u>1.000</u>	1.120
3	<u>1.000</u>	0.895	1.216

In addition, there is between-cell variation and fluctuation in the total amount of DNA being delivered to cells. This is exemplified with the variation in the fluorescence measured between different cells, across several orders of magnitude (10^0 - 10^4). The total fluorescence measured per cell (eGFP, TagBFP and mCherry fluorescence combined) will vary across several orders of

fluorescence intensity. It is assumed that cells with higher total fluorescence, as a result of higher plasmid delivery, will be more likely to adhere in uptake and expression to the ratio of the plasmids being transfected. The diverse amounts of total plasmid delivery will therefore affect adherence to the transfected ratio being investigated here. This assumption is based upon the notion that, when plasmid delivery is higher there is a greater chance of even distribution of plasmid uptake, similar to the transfected ratio of plasmids; derived from the central limit theorem: when sample size increases, the likelihood of a normal distribution in the averages from the data derived increases. It is assumed that a well-mixed distribution of plasmids and cells are being transfected and that there is no bias in the uptake of particular plasmids, something that is aided here through the use of the same 'Expression-Plasmid' backbone throughout. The cells with higher levels of fluorescence can be thought of as higher 'contributors', having more influence within the population in terms of gene expression. This is of interest for this transient expression platform as cells with higher gene expression will be greater 'contributors' in terms of the production output of the population. For example, for cells transfected with a plasmid to express a recombinant product, cells with more plasmid delivered are likely to be higher expressers, and therefore higher producers, of the recombinant product. Titre, the primary attribute considered for producing cells, is the measure of accumulation of recombinant mAb during culture, higher producing cells will contribute more significantly to the resultant titre. Analysis here is weighted towards these 'higher expressers', as they would be higher 'contributors' to this important production attribute, their ratio in expression of transfected genes will contribute greater to the resulting responses measured.

Total cellular fluorescence is accounted for when investigating the variation in expression. When undertaking the analysis and calculating averages and variance, in adherence to the transfected ratio, from single-cell measurements, weighted statistics (mean, median and CV), established from the total cellular fluorescence of each cell, were used, calculated using R software.

6.3.4. Expression Adherence to Transfections with Equal Copies of Genes

Multiple transfections with equal loading (1:1:1 ratio) of individual plasmids, each encoding separately for the three fluorescent proteins, were undertaken, with and without the co-transfection of the recombinant DTE-mAb. For these, either 266.7 ng of DTE-mAb plasmid or empty 'Expression-Plasmid' were transfected into cells along with 533.3 ng, made up of the fluorescent protein-encoding plasmids (177.8 ng of each) (**Table 6.1**). Transfections with a multigene plasmid, encoding all three fluorescent proteins in tandem on one plasmid enabled comparisons in transfection setup for simultaneous transfection of multiple genes. 533.3 ng of multigene plasmid, with 266.7 ng of empty 'Expression-plasmid', was used for each transfection (**Table 6.1**). Of interest were how individual cells were adhering in expression, of the three reporter genes, to the transfected plasmid ratio. The ratio, representing the expression ratio between the three proteins, was derived for each cell, as previously explained, in **Section 6.3.3**. For each fluorescent protein, the normalised expression levels for all individual cells was compared. The results derived for each protein are visualised in the histograms in **Figure 6.8** with the average and variance tabulated in **Table 6.3**.

In **Figure 6.8**, the data show populations all transfected with equal copies of eGFP, TagBFP and mCherry. Analysis of co-transfections with separate plasmids (**Figure 6.8A**) shows individual cells are, on average (both in terms of mean and median) receiving and expressing the genes in the same ratio as they were transfected (1:1:1 ratio). With a mean ratio, in terms of TagBFP: eGFP: mCherry expression, of 1.038:1.021:1.043. There is variation in adherence away from this average ratio. The average CV across all three proteins, depicting relative variation by expressing the standard deviation of data as a ratio to the data mean, is 29.7%.

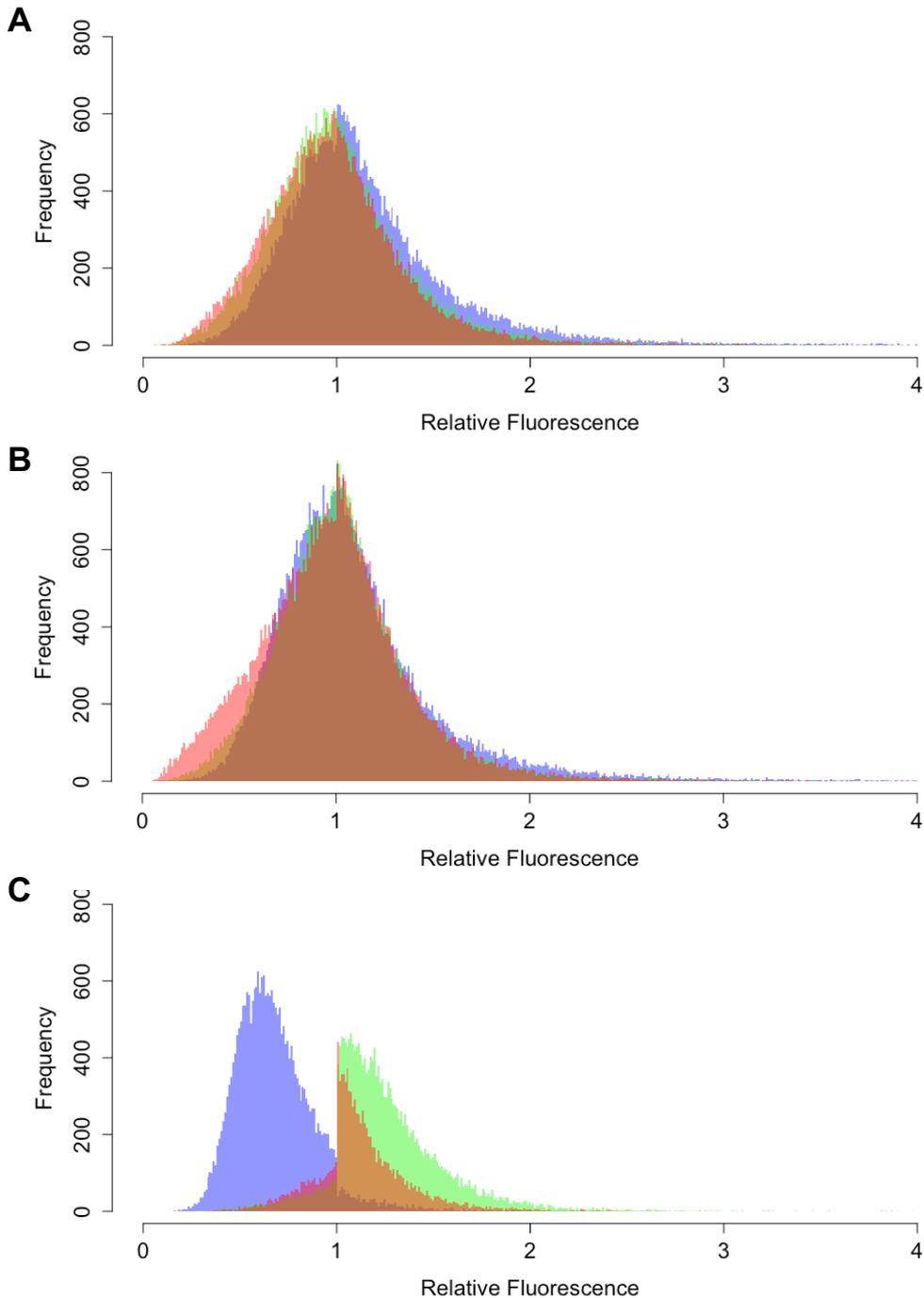


Figure 6.8: Histograms Representing the Spread of Fluorescent Protein Expression at a Cellular Level

The spread and average expression for each protein (TagBFP=blue, eGFP=green, mCherry=red) on a cellular level, in terms of relative fluorescence, from transfections with (A triple) three separate plasmids, each encoding a separate fluorescent protein, transfected at equal amounts (1:1:1 ratio); (B mAbTriple) the same three plasmid co-transfections but with the additional expression of a recombinant mAb (DTE) rather than empty-‘Expression-Plasmid’ and (C Multigene) the same three fluorescent proteins transfected and expressed from a multigene plasmid encoding all three genes in tandem. The data represent results from 18 separate transfections (A) and (B) and nine separate transfections (C).

Table 6.3: The Average and Variance in Fluorescent Protein Expression at a Cellular Level

In conjunction with Figure 6.8, the table displays the mean and median relative fluorescence and CV (spread of fluorescence results) of fluorescent protein expression at a cellular level. Data is shown from the different types of transfection setup – co-transfection with equal amounts of three separate plasmids (Triple), in conjunction with mAb expressing plasmid (MabTriple) and with all three proteins being expression from the same plasmid (Multigene). The average CV of co-transfection with separate plasmids is 28.85% and using a singular plasmid is 20.70%.

		Transfection Ratio	Mean	Median	CV (%)
Triple	BFP	1	1.038	0.986	32.0
	GFP	1	1.021	0.998	26.3
	mCherry	1	1.043	1.026	30.7
MabTriple	BFP	1	1.030	0.984	29.9
	GFP	1	1.040	1.018	24.6
	mCherry	1	1.002	0.988	29.6
Multigene	BFP	1	0.602	0.586	25.7
	GFP	1	1.228	1.179	17.4
	mCherry	1	1.181	1.126	19.0

The additional biosynthetic burden to cells expressing and producing a recombinant product gene (in this case a DTE-mAb) does not appear to adversely affect the individual cell expression to the transfected plasmid ratio. As seen in **Figure 6.8A/B** and **Table 6.3**, the average expression ratio and variance were not substantially different for transfections that did not contain mAb encoding plasmid and empty-‘Expression-Plasmid’ was transfected instead. The mean ratio across the three fluorescent proteins when mAb plasmid was co-transfected was 1.030:1.040:1.002 (B:G:C) and the average CV was 28.1%.

Results from transfections containing a multigene plasmid encoding all fluorescent proteins, rather than co-transfections of separate pieces of DNA, show a different trend in expression of the three fluorescent proteins. The results are represented in **Figure 6.8C** and **Table 6.3**. Overall equal expression between the three fluorescent proteins is not observed when the genes are all encoded on the same plasmid. In these scenarios, unanimously equal copies of each gene are being introduced into each cell successfully transfected. The unequal expression between the three proteins results in a ratio of 0.602: 1.228: 1.181 (of B:G:C), which is somewhat from the desired, and expected, 1:1:1 ratio. These

data suggest that there is some degree of interference, dampening expression of the middle gene on the multigene plasmid, (in this case the TagBFP gene; with the eGFP gene positioned upstream and the mCherry gene downstream). Between individual cells there is less variance from this average expression ratio. The average CV, of 20.7%, is lower than that derived from transfection with separate singular expressing plasmids. Expression conformity is tighter, cells are adhering more closely to each other and the average ratio. In other words, while individual cells are transfected in a similar way to each other, they are all more likely to demonstrate discrepancy in terms of expression relative to transfection gene ratios. There is a difference between an ability to deliver DNA to the cell at a particular ratio and achieving a particular stoichiometry in expression. Transfection ratios with the multigene plasmid will be clearly more accurate, when the plasmid is transfected into cells all the genes are being delivered equally (in a 1:1:1 ratio) as they are all contained in the one piece of DNA. However, the ratio of gene expression does not reflect this because of an unknown level of interference between the transcriptional units on the plasmid, resulting in uneven expression.

To further probe the variation in adherence, at an individual cell level, to the ratio at which plasmids were transfected, the data were categorised into different ratios. Represented in **Figure 6.9** are the frequency (percentage of cells) which met a particular criteria, i.e. were expressing the three fluorescent proteins within a particular ratio. The data show the expression between all three proteins to within a particular ratio. 81.7% of the triple expressing cells conformed in expression of all three fluorescent proteins to within a 0.5:1:1.5 ratio (represented by a black dashed line in **Figure 6.9**). This proportion does not alter with additional recombinant protein production burden (DTE-mAb production), with 80.3% of the population adhering to within this ratio (purple dashed line in **Figure 6.9**).

As expected, from the discrepancy in a 1:1:1 ratio using a multigene plasmid, a lower proportion, namely 64.8% (green dashed line in **Figure 6.9**) of cells were within a 0.5:1:1.5 ratio. This reduction in conformity demonstrates the use of separate plasmids to transfect multiple different genes into cells is the more

appropriate approach resulting in a greater conformity to the transfected, desirable ratio.

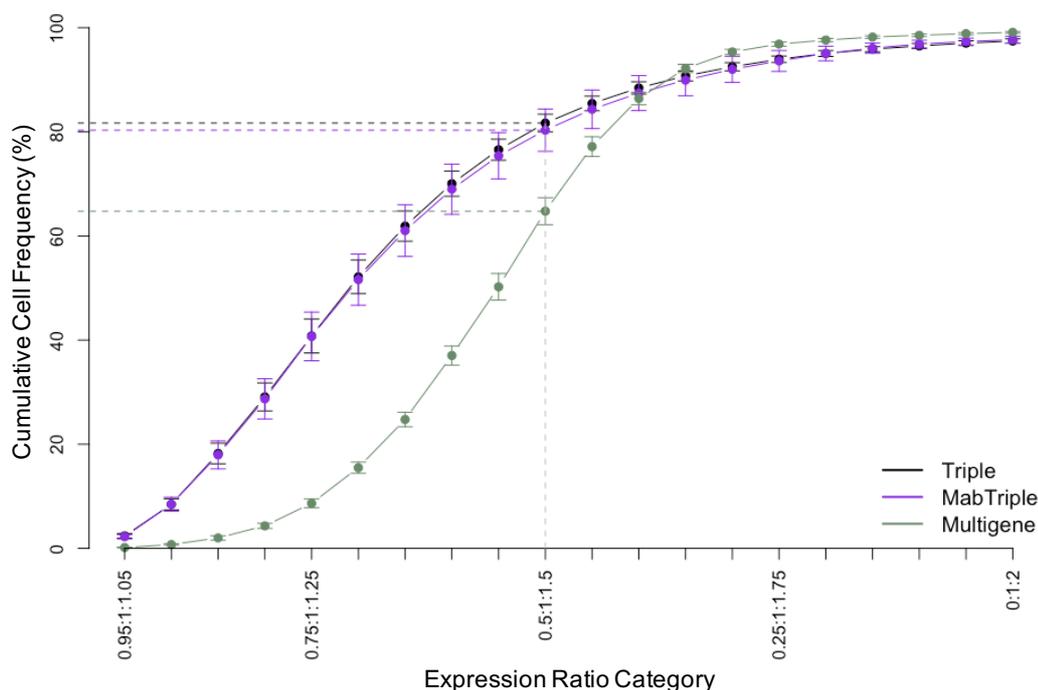


Figure 6.9: The Adherence of Cellular Expression to the Transfected Plasmid Ratio (1:1:1)

The data show how tightly, at a cellular level, expression of the three fluorescent proteins adheres to the ratio at which the plasmids were transfected. The dashed lines show expression within a threshold of a 0.5:1:1.5 ratio, from co-transfection of proteins encoded on different plasmids (Triple and mAbTriple) or from transfection of a single multigene plasmid expressing all three proteins in tandem. The data points represent the mean and SEM from either six (Triple and mAbTriple) or three (Multigene) biological replicates each with three technical replicates.

6.3.5. Alteration in the Ratio of Plasmids Transfected

When carrying out simultaneous expression of multiple genes, it is relevant to be able to alter the ratio (stoichiometry) between the components. For example, to be able to adjust expression levels to find and implement an optimal balance between effector genes. When transfecting with separate plasmids, the simplest approach to adjust ratios is to alter the amount of each type of plasmid in the transfection mix. To confirm this approach results in alterations in stoichiometry with similar trends in cellular adherence and variation to a 1:1:1 ratio, transfections with different ratios of the fluorescent protein plasmids were

performed. Six different ratios were used, covering higher and lower expression levels for all three proteins (**Table 6.1**). In all cases 533.3 ng of fluorescent protein encoding plasmid transfected was used in total; with 266.7 ng of empty-‘Expression-Plasmid’ used to equalise total DNA load in all cases.

To enable direct comparative analysis, the fluorescence intensities, measured for each cell were converted into an x:1:y (1:1:1) style format. This allowed for direct comparisons between different ratios to be made and for variance, in the expression of each individual fluorescent protein, to be compared. An example of the normalisation is shown in **Table 6.4**. Firstly, the three fluorescence intensities in each cell were divided by their relative quantity in the transfected plasmid ratio mix. For example, the results shown in the table are derived from cells which were transfected with four parts mCherry plasmid, two parts eGFP plasmid and one part TagBFP plasmid. The fluorescence intensities from mCherry and eGFP, will always be, approximately, four and two times as bright than those for TagBFP, respectively. To account for this, and put the fluorescence intensities in direct relative terms to each other, all mCherry and eGFP fluorescence intensities for cells transfected with this plasmid ratio were divided by four and two respectively. This approach was implemented for all the different ratios examined. Subsequently, in a similar respect, as described in **Table 6.2**, to account for the wide range in total cellular fluorescence, and to put cellular fluorescence into relative terms between the three proteins, the three fluorescent intensities were normalised for the median value for each cell. This put the datasets into relative x:1:y ratios, similar to that previously explained.

The results, shown in **Figure 6.10** and tabulated in **Table 6.5**, demonstrate that specific ratios of genes can be taken up and expressed by altering the amount of each plasmid that is transfected into cells. Based on expression results from individual cells, the average (both in terms of mean and median) ratio of the three proteins matches the relative plasmid ratio that was transfected into cells. The peaks on the histograms in **Figure 6.10**, are at the same ratio the plasmids were transfected at. For **Figure 6.10** the derived ratio for each cell has been converted back to the represent the ratio in terms of the original transfected ratio (e.g. reverted into a 4:2:1 ratio, rather than the 1:1:1 ratio used for analysis) to aid visualisation of the results across the different ratios. The results in **Table 6.5**

summarise how the average (both mean and median) expression within cells, for each of the six different transfection ratios tested, matches its transfected plasmid ratio satisfactorily. When rounded to the nearest integer all expression ratios exactly correspond to the transfected ratio.

Table 6.4: Example of Normalisation Undertaken to Adjust for Differing Fluorescence Intensities Derived from Altering the Transfected Plasmid Ratio

In the example below cells were transfected at a ratio of 1:2:4, of TagBFP, eGFP and mCherry-encoding plasmids respectively. For direct comparison to be made, the fluorescent intensity (FI) for each fluorescent protein was divided by its relative part in the transfected plasmid mix. In this example, fluorescent intensities from mCherry and eGFP were divided by four and two, respectively. Next the median fluorescent intensity value for each cell was used to convert cellular fluorescence into a comparable ratio (x:1:y) format, to adjust for the differences in total cellular fluorescence across the dataset.

Event (Cell)	BFP FI	eGFP FI	mCherry FI
1	95	204	416
2	3	7	13
3	767	1557	3010

↓ Normalisation for differing fluorescence scales due to unequal stoichiometry ↓

Event (Cell)	Normalised BFP	Normalised eGFP	Normalised mCherry
1	95	<u>102</u>	104
2	3	3.5	<u>3.25</u>
3	<u>767</u>	778.5	752.5

↓ Normalisation using MFI for each cell ↓

Event (Cell)	Normalised BFP	Normalised eGFP	Normalised mCherry
1	0.931	1.000	1.020
2	0.923	1.077	1.000
3	1.000	1.015	0.981

For example, cells transfected with a 1:2:4 ratio of Tag BFP, eGFP and mCherry plasmids, respectively, had a mean cellular expression ratio of 1.160:1.974:4.096. The variance across the population, represented here by the

CV, is also shown in **Table 6.5**. The average CV, across all ratios, is 30.3%, almost identical to that derived from cells transfected with an equal ratio of plasmids. The CV is fairly consistent between all the results from all the ratios tested.

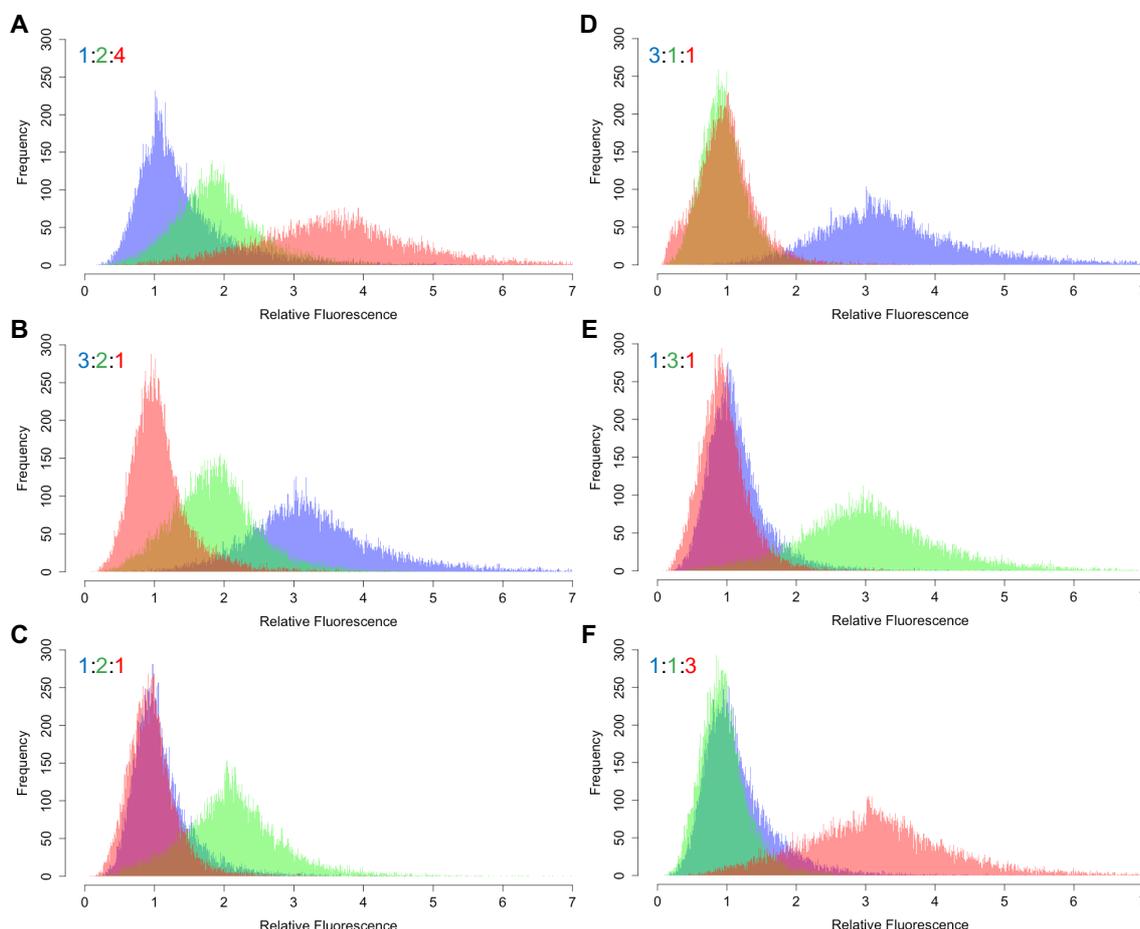


Figure 6.10: Histograms Representing Cell Fluorescent Protein Expression from Alterations in the Plasmid Ratio Transfected

Three fluorescent protein encoding plasmids were co-transfected into cells at differing, randomly chosen, ratios. The spread and average expression on a cellular level, in terms of relative fluorescence, is shown for six different ratios (represented as ratio of BFP (blue) to eGFP (green) to mCherry (red) expression): **(A)** 1:2:4, **(B)** 3:2:1, **(C)** 1:2:1, **(D)** 3:1:1, **(E)** 1:3:1 and **(F)** 1:1:3. Each ratio represents results from nine separate transfections.

Table 6.5: The Average and Variation in Cellular Fluorescent Protein Expression from Alterations in the Ratio of Plasmids Transfected

In conjunction with **Figure 6.10** the table displays the mean and median relative cellular fluorescence ratio and CV of fluorescent protein expression from alterations in ratio of plasmids transfected. The average CV is 30.3%.

		Transfection Ratio	Mean	Median	CV (%)
A	BFP	1	1.160	1.074	39.3
	GFP	2	1.974	1.932	26.3
	mCherry	4	4.096	3.992	30.8
B	BFP	3	3.154	3.034	26.7
	GFP	2	1.961	1.946	24.5
	mCherry	1	1.066	1.026	33.0
C	BFP	1	1.003	0.949	33.7
	GFP	2	2.173	2.128	25.9
	mCherry	1	1.037	1.001	39.2
D	BFP	3	3.349	3.190	28.3
	GFP	1	1.001	0.987	27.1
	mCherry	1	0.972	0.933	34.2
E	BFP	1	1.024	0.971	33.1
	GFP	3	3.180	3.115	24.3
	mCherry	1	1.080	0.992	33.1
F	BFP	1	1.011	0.965	31.5
	GFP	1	0.969	0.951	26.9
	mCherry	3	3.462	3.371	27.3

To further investigate the cellular expression ratio derived from alteration in the ratio of plasmids transfected, the results were classified into cellular ratios. The data were kept in the normalised x:1:y format for direct comparisons to be made. The percentage of cells with ratios of the three fluorescent proteins within particular categories (e.g. categories of 0.95:1:1.05 and 0.9:1:1.1), is shown in **Figure 6.11**. All six transfection ratios tested produced similar trends of conformity in expression ratio. In a similar respect as seen from transfections with an equal (1:1:1) plasmid ratio, on average 77.8% of cells are within the equivalent of a 0.5:1:1.5 ratio of their expression of the three introduced proteins. As seen in the dashed lines, in **Figure 6.11**, very similar percentages of cells conform to within the different ratio threshold categories. This shows there is similar conformity when different amounts of individual plasmids are transfected into

cells. These data demonstrate that altering the ratio of separate plasmids transfected enables reliable and specific expression of different ratios of genetic components and does not introduce considerable differences in cell to cell variation.

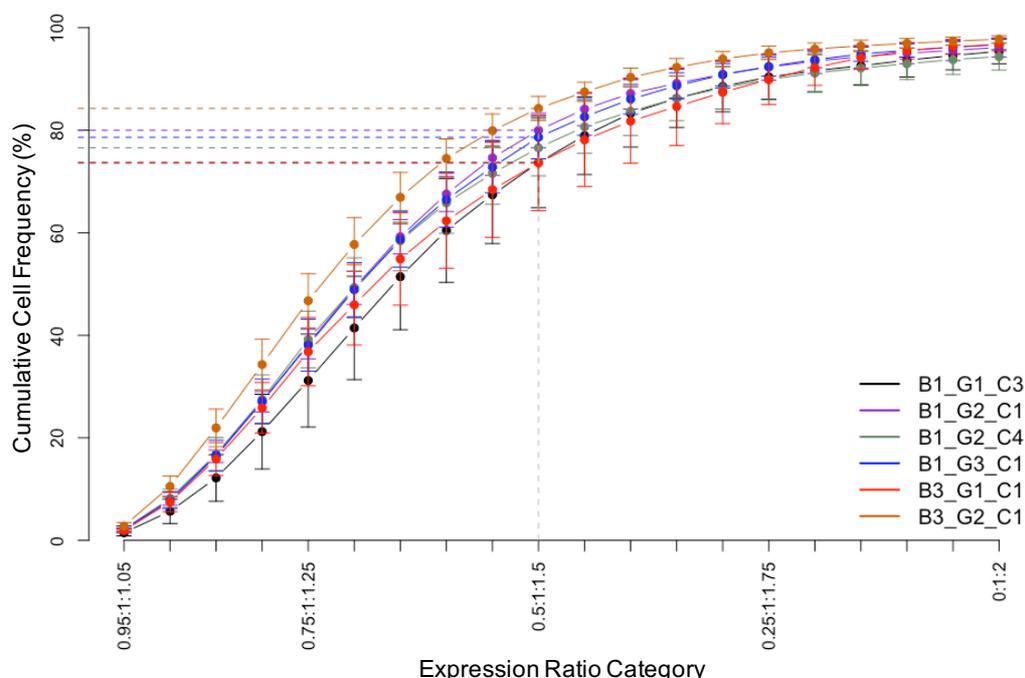


Figure 6.11: The Conformity of Cellular Expression with Alterations in Transfected Plasmid Ratios

The data show how tightly cell expression of the three fluorescent proteins conforms to the ratio at which plasmids were transfected at. The dashed lines show the amount of cells conforming within a 0.5:1:1.5 ratio threshold, from co-transfection of plasmids at differing ratios. The ratios represent plasmids for TagBFP (B), eGFP (G) and mCherry (C). As explained in the main text the data has been converted to be in a 1:1:1 style to enable direct comparison. The data points represent the mean and SEM from three biological replicates each with three technical replicates.

6.4. Discussion

It is likely that the optimal approaches to improve the expression capability of CHO cell factories will require product specific, multiple component solutions (Johari et al., 2015). The high-throughput transient transfection platform, developed as described in the previous chapter, aims to be utilised in this direction, to screen potentially beneficial genetic components. There is potential to screen for both their effect on the CHO cell factory in isolation and, of more

potential benefit, in combination with other components. Alongside this, the product gene being expressed can be interchanged to investigate product specific solutions. For most screens, the platform would incorporate transfection and expression of multiple genes simultaneously. These multiple (more than two) genes could be delivered to cells in two main ways: through the co-transfection of genes on separate plasmids; or via transfection of a singular multigene plasmid. Both approaches have advantages and disadvantages, but it is deemed substantially more straightforward to express each component using a separate plasmid (Mansouri and Berger, 2014). Titrating different amounts of each plasmid to alter expression levels and, more crucially, allow for different genes to be combined at different levels would be logistically easier than using a multigene plasmid. The purpose of this work has been to provide assurance that the transfection setup enabled appropriate multiple gene delivery and to confirm that differing stoichiometry is attainable. Of interest were cell based comparisons of co-transfections of multiple plasmids and transfections of a multigene plasmid, and the cell to cell variation derived.

6.4.1. The Use of Fluorescent Proteins to Analyse Multiple Gene Expression

To investigate multiple gene transfection, three separate fluorescent reporter proteins were employed. Fluorescent proteins are frequently utilised as tags when measuring cellular gene expression, and when combined with powerful flow cytometry approaches, allow for large-scale cell level analysis to be undertaken. As presented here, there are limitations to simultaneous measurement of multiple fluorescent proteins. Fluorescent proteins have wide excitation and emission spectra resulting in significant spectral overlap. The primary approach to overcome this is to find proteins with as differing excitation and detection wavelengths to each other as possible and, where feasible use a flow cytometer with many different optics (Hawley et al., 2017). This approach was taken here, although initially only a flow cytometer with 405, 488 and 633 nm lasers was available. Far-red lasers (630-640 nm) are very common lasers present on most flow cytometers, but there is a limited number of far-red proteins suitable for excitation with these lasers. At the time, E2-crimson was one of a few far-red

proteins available sufficiently excited by a 633 nm laser and characterised for mammalian expression (Strack et al., 2009; Telford et al., 2012). The limitations in multiplexing E2-crimson with TagBFP were not anticipated, and the spectral overlap only becoming evident from flow cytometry analysis, could not be overcome by applying compensation. Access to a flow cytometer with a green 532 nm laser enabled other fluorescent proteins to be potentially utilised, replacing the use of E2-crimson with mCherry, a protein frequently utilised in the literature and commonly multiplexed with eGFP. Along with the complications from E2-crimson there were other smaller spectral overlaps, that could not be compensated for, between eGFP and TagBFP at high fluorescence intensities. These were overcome by lowering the excitation voltage used, but this resulted in the introduction of false negatives and a lowered sensitivity. This was detrimental to enabling any direct transfection efficiency assessment to be made. To remove these limitations from the study, more fluorescent proteins would need to be tested to find a better combination of independently measurable reporters. Sufficient knowledge and predictability of the many fluorescent reporter proteins that could be combined is not available, and from the unanticipated outcomes described here, it is recommended that any potential combinations would need to be tested within cells, to confirm independent evaluation is achievable. Testing of this nature would be time-consuming and potentially expensive, with many fluorescent reporter gene constructs being required.

Despite not being the most optimal combination of three fluorescent proteins for the flow cytometry setup available, the combination of TagBFP, eGFP and mCherry for this work were adequately, independently measurable (**Appendix Figure B.2**) and the overlap limitations could be overcome sufficiently. For an improved detection setup in the future, a spectral flow cytometer would allow for the detection of multiple fluorescent proteins without the problems of spectral overlap (Hawley et al., 2017). This emerging technology removes separate fluorescence detection and can ascertain the abundance of different fluorophores from within a mixture. As the particular interest for this work is to evaluate the expression of multiple plasmids, as a result of altering plasmid type and stoichiometry, using the developed transfection platform, an alternative to measuring fluorescent reporter proteins could be to fluorescently tag plasmid

DNA. The intracellular levels of different overexpressed genes could then be measured. Advantageously this would be a more direct measurement of gene uptake, not affected by discrepancies introduced from cellular transcription and translation processes or inherent fluorescence, as with fluorescent proteins. Beneficially fluorescent DNA labels appear to have narrower spectra than those of fluorescent proteins, reducing potential overlap when combined.

6.4.2. Co-transfection with Multiple Plasmids Compared to a Single Multigene Plasmid

The co-transfection of multiple plasmids into cells is a widely used approach to introduce multiple genes into cells simultaneously. It is suggested that a transfection process introducing many plasmid copies to each cell with a high transfection efficiency, increases the likelihood of the number of independent plasmids being introduced (Jordan and Wurm, 2003). The developed transfection setup provides a consistently high transfection efficiency, as shown using a single reporter gene (eGFP), and the transfection efficiency should be maintained, with total concentration of plasmid DNA remaining consistent for all transfections. A validation of transfection using the three fluorescent proteins suggests a high transfection efficiency is robustly achieved for all three plasmids at approximately 88%, as seen in **Figure 6.7**. Despite not knowing the number of actual plasmid copies introduced into cells, the amount of cellular expression (MFI) is high, and the number of plasmid DNA copies transfected is far in excess of the number of cells being transfected, which suggests that a high number of copies of plasmids are reaching cells.

Analysis of the work undertaken in this chapter shows that co-transfection with multiple separate plasmids, using the developed transfection platform, introduces, on average, an equal number of copies of each plasmid to individual cells. As expected, there is variation in expression diverging away from the ratio of plasmids transfected. From transfections undertaken with equal plasmid copies being introduced, approximately 80% of triple-positive cell populations were expressing the three recombinant genes, to within a 0.5:1:1.5 ratio of each other. The electroporation setup used to introduce the DNA into cells will not be responsible for all the variation (average CV, of 28.85%). DNA delivery into cells

is considered to be a random Poisson process (Jordan and Wurm, 2003; Materna and Marwan, 2005), which will be a factor in this variation measured.

In addition, there will be inherent biological differences between individual cellular units; there will be heterogeneity between transcription, translation, protein folding and turnover rates, and cells will vary in terms of their total plasmid uptake. The use of very similar plasmids to encode for the three reporter genes, except for a few sequence differences, will help to alleviate some disparity in delivery, the plasmids will be approximately identical and indistinguishable to the cells in terms of their uptake. There will be differential fluorescence brightness and differences within the flow cytometry analysis contributing to the variation. The average CV, of 20.7% across the expression of each protein, derived from transfections using a multigene plasmid, is still relatively high. In theory, there should be very little difference within the level of each fluorescent protein, consistency should be high for multigene plasmids as transfected cells only receive copies of all three genes. This analysis method, using fluorescent proteins to investigate adherence of multigene transfection, appears to overestimate the variance of uptake and expression.

There is no substantial difference in expression of the reporter constructs when there is the additional transfection and expression burden on cells to co-express a recombinant product gene, alongside the three reporter proteins. The same cellular trend in conformity to the ratio of the three fluorescent reporters is apparent when the three plasmids are co-transfected along with a larger plasmid encoding for a DTE-mAb instead of an empty-‘Expression-Plasmid’. This is of importance, as this will be the likely scenario encountered in the screening platform setup.

The primary alternative to introducing multiple genes simultaneously is to use a single multigene plasmid. This has the advantage that when plasmid delivery is successful all the genes of interest will have been taken up and be present for co-expression within the cells. The multigene used for this work was generated using Golden Gate Assembly (New England Biolabs), which enables directional cloning of multiple DNA fragments simultaneously, and, at the time of writing, is still under development in collaboration between MedImmune and Yash Patel (Yash Patel, personal communication). From the transfections undertaken with

the multigene plasmid used here, as described in **Section 6.3.4**, it is clear there is unequal expression of the three genes present on the plasmid, despite all genes being expressed under the same elements and CMV promoters. There is a greater discrepancy in achieving the transfected gene ratio than from separate plasmids. Analysis here suggests that the middle gene on the plasmid, in this case TagBFP, has approximately half the expression level than its two flanking genes. Positional effects, resulting in uneven expression, are a well-documented phenomenon for multigene constructs expressing more than one gene (Kadesch and Berg, 1986; Proudfoot, 1986; Underhill et al., 2007; Yahata et al., 2005). The suppression of downstream genes is widely attributed to transcriptional interference; dampening expression of a downstream gene due to suppression from the close proximity of an upstream transcriptional process (Shearwin et al., 2005; Yahata et al., 2005). It is suggested that for this multigene plasmid the eGFP and mCherry genes are sufficiently separated to not be interfering transcriptionally; though the exact mechanism of uneven expression was not within the scope of the work here.

There have been successful developments in the progression of multigene plasmid construction for mammalian cells (Guye et al., 2013; Kriz et al., 2010), with recent constructs advantageously capable of equal expression of several genes simultaneously. The multigene plasmids developed were shown to introduce less variance in expression to that derived from co-transfection of separate genes. Guye and colleagues harnessed the use of insulator elements between each gene on a multigene plasmid. These spatially separate neighbouring elements alleviated transcriptional interference. This has been shown to be successful in other multigene constructs, but the relatively long length of the insulator, for example 1.2kb, is not ideal for the already inherently large multigene plasmid constructs (Hasegawa and Nakatsuji, 2002; Yahata et al., 2007). Despite these very encouraging developments, the novel multigene constructs to date do not have the ability to achieve a precise stoichiometry between the multiple genes present. Implementing the advances in synthetic promoters may have significant potential to facilitate precise control of each gene in a multigene circuit (Brown and James, 2016). This approach would require positional effects and interference within the multigene to be consistent and

predictable between transfection and the expression of different genes, for example different lengths of genes. If not it will be substantially more difficult to overcome and engineer.

The problems identified here with the multigene construct are likely to be overcome; interference between genes has the potential to be reduced and exact, accurate expression between components could be attainable with current synthetic engineering developments. However, based on the presently available multigene plasmid, that the use of separate plasmids to introduce multiple genetic components into cells is the best approach. The use of separate plasmids, although not guaranteed, introduces genes at approximately equal levels on average (1.038:1.021:1.043), into individual cells. Unequal expression, likely from transcriptional interference, is apparent from the current multigene plasmid (0.602:1.228:1.181). Although as expected, expression from a multigene construct is less variable, with a lower average cellular CV of 20.70% compared to 28.85%, for separate plasmids. Even with the prospect of improvements in mammalian multigene constructs, for screening setups it remains significantly more straightforward to alter combinations of genes being transfected by altering the mixture of plasmids being transfected, as opposed to undertaking numerous cloning experiments to generate multigene plasmids for all the permutations of interest.

6.4.3. Stoichiometry Attained Through Co-Transfection of Separate Plasmids

The work in this chapter demonstrates that varying the ratio of plasmids transfected is, on average, effective in altering the stoichiometry of gene expression present at a single cell level. This is an essential characteristic for a combinatorial effector screening platform. Altering the ratio of plasmids transfected does not appear to alter the cell to cell conformity or variation of plasmid expression. When three plasmids are co-transfected, cells, on average, express the plasmids at the same ratio at which they are transfected, with the average CV of 30.3% (and a standard deviation of 4.6%). Validating that alteration in stoichiometry is attainable with the developed transfection platform and increases confidence in the platform's suitability for subsequent co-

transfection effector gene screening. Altering the levels of separate plasmids being transfected is the most straightforward and flexible approach to implement altering gene stoichiometry. Precise control of the expression of genes could be achieved by altering the promoter upstream of each gene and therefore levels of gene transcription (Hansen et al., 2017). This is an emerging capability, with successful development of synthetic promoters with varied transcription strength, making this more likely (Brown and James, 2016). Implementation of this approach would introduce additional cloning steps, to change the promoter-gene combination and therefore level of expression, something less desirable for a high-throughput setup.

6.5. Summary

The platform aims to be used to screen for effector gene functionality, including the use of combinations, upon CHO cell factory expression. The work undertaken in this chapter aims to assess and confirm multiple gene transfection using the current transfection setup previously optimised. Transfections with three fluorescent reporter genes encoded by either separate plasmids or a singular multigene plasmid, were used to investigate multiple gene transfection. Singular cell analysis using flow cytometry demonstrated that, on average the use of separate plasmids resulted in cells conforming in expression, though with more variance, to the ratio that the plasmids were transfected, in a superior respect, to expression from a non-engineered multigene plasmid. Alterations to the multigene plasmid could progress towards equivalent expression between the genes present on the plasmid, but it would remain more straight-forward and flexible, for a screening tool measuring potential combinatorial effects, for separate plasmid levels to be altered for transfection. When changing the ratio of plasmids transfected the cell population conforms consistently, and stoichiometry of different genes can be attained. Using the developed 'Expression-Plasmid' and high-throughput transfection setup the platform can be implemented for simultaneous multiple gene delivery and expression. Screening of a range of effector genes for their functionality on an industrial recombinant mAb is explored in the next chapter.

7. Implementation of the Developed Platform for Effector Gene Screening

This chapter implements the high-throughput transient transfection platform as a screening tool to test many, simultaneous, co-expressed engineering strategies with the aim of improving DTE protein production. A library of effector genes, targeting functional cellular modules of ER folding and assembly processes and Golgi complex secretory pathways, was developed and then auditioned. The chapter demonstrates that engineering strategies can be screened simultaneously and rapidly to identify product-specific solutions to improve protein production, the amount of gene dose transfected alters effects measured and a DoE methodology will help to identify and optimise stoichiometry of combined effects of the co-expressed of multiple components.

Acknowledgments

The B cell/plasma cell transcriptomic analysis and in-house CHO transcriptomic and proteomic datasets were provided by Nabila Rahman, Dr Darren Geoghegan and Dr Joseph Longworth, respectively (University of Sheffield). Target data-mining and analysis literature CHO omic sources was undertaken in collaboration with Nicholas Barber (University of Sheffield).

7.1. Introduction

There is a developing repertoire of diverse and engineered recombinant products of industrial interest within biopharmaceutical manufacturing pipelines. There are increasing numbers of 'second-generation' and engineered molecules (Walsh, 2010; Walsh, 2014). With this trend comes an increase in the number of products described as DTE, characterised by poor production yields, from limitations in the

production processes that take place within the CHO cell factory itself. For these there is an increased requirement for process alterations and improvements to increase yields and make the products viable candidates in the biomanufacturing pipeline. Currently, one of the primary approaches to improve production yield involves comparing large numbers of cell lines for their inherent production and growth capabilities to screen natural heterogeneity and find individual cell lines with enhanced production properties (**Section 2.3.4**). An alternative, with potentially greater efficacy for yield improvement of some products, particularly those that are DTE, is to engineer the cellular factory itself (**Section 3.4**). Engineering the cell host refers to altering the inner workings of the cell, commonly through overexpression or down-regulation of particular genes or via the addition of chemicals, to adjust the levels of proteins and processes present. The approach directly changes the cellular machinery with the potential to implement strategies and develop phenotypes beyond that possible by just screening for natural heterogeneity. For DTE proteins that seemingly have limitations somewhere within their production process within the cell causing the associated characteristic poor yields, this approach is undeniably of significant potential. Particular processes and parts of the cell machinery can be targeted, including specific parts of the cell which may be limiting production capabilities or prominent parts of the recombinant protein production processes. For example, Johari et al (2015) show how, for a DTE fusion protein, a bespoke, biphasic engineered solution, of CypB overexpression to increase cell growth followed by PBA and glycerol addition to increase specific production later on in culture, was a more beneficial approach to fusion protein expression during a fed-batch transient production process, than screening the inherent biological capability of sixteen different clonal populations. In relation to this concept, the implementation of a high-throughput screening approach has been emphasised in the literature as being a useful and appropriate requirement to help work towards investigating, optimising and implementing product-specific, bespoke engineering solutions (Johari et al., 2015; Pybus et al., 2013). Different products will introduce different inherent biosynthetic limitations, therefore a platform approach should help to identify product-specific engineering solutions by rapidly swapping the recombinant product being investigated.

Within this chapter the use of the developed transient transfection platform, as described in **Chapter 5**, for screening an effector gene library is explored. A DTE-mAb was used as a model recombinant product for improvement of expression. This chapter provides a 'proof-of-concept' demonstration of the platform to rapidly co-express and audition the functionality of many components simultaneously, to identify potentially generic and product-specific engineering strategies to improve product expression.

Since it was not possible to test the many different target genes that could have been investigated, several information sources, primarily the literature and plasma CHO cell omic datasets, were used to provide a rationale for the generation of a practical library of genes. The sources of information helped to identify genes associated with a high antibody producing phenotype, or whose function is likely to play a role in doing so, the main desirable characteristic of interest here. The primary engineering foci were protein processing, folding, assembly within the ER, and secretion of the protein through the Golgi complex and out of the cell; crucial modules of recombinant protein production. Post-translational and secretory capacity often do not correlate with levels of recombinant gene, mRNA or intracellular protein (Fann et al., 1999; Ku et al., 2007; Mohan et al., 2008). This phenomenon is more likely in a high efficiency TGE system, where high levels of recombinant genes are immediately introduced and expressed. The amount of recombinant nascent polypeptide is likely to be high and require high levels of processing and secretory reactions; exacerbating any post-translational limitations. Additionally, previous overexpression strategies exemplified in the literature demonstrate that engineering protein processing, through co-expression of components of the ER and UPR transactivators, can be suitable approaches to improve the processing and secretory capabilities of the production host and improve the yields of DTE proteins generated (Cain et al., 2013; Le Fourn et al., 2014; Johari et al., 2015; Pybus et al., 2013). As the response from an effector gene is likely to be dependent on the expression level of that gene (Hansen et al., 2017), the high-throughput nature of this platform enabled different doses of each effector gene to be tested.

In the literature, there are limited examples of combining expression of multiple genes, however it is likely a combinatorial solution will be of added benefit (Cain et al., 2013; Johari et al., 2015). When combining expression of multiple genes, the components need to be balanced as the ratio used is likely to be crucial to the effect introduced. There is a need for approaches to test novel combinations of components in a standardised manner. When investigating combinatorial effects, the potential number of combinations of interest to test could be exponentially vast. Therefore, a higher-throughput approach to testing will be useful. In **Chapter 6**, the use of the high-throughput transfection platform to co-transfect and co-express multiple genetic components simultaneously was verified. In this chapter, a methodology to test the functionality from multiple co-expressed components is outlined. A DoE approach incorporating Design-Expert® software to empirically and statistically model multiple-component engineering strategies to attain optimal combinations is described.

7.2. Experimental Approach

To compare the functionality of multiple genetic components both singularly as well as in combinations, consistent and comparable expression was required. All synthetic parts (effector genes) were encoded separately, but transcription was driven by the same CMV promoter on an identical plasmid backbone (as described in **Section 5.3.1**). The component parts were interchanged and transfected proportionally alongside the recombinant DTE-mAb plasmid.

The screening approach taken, to compare and multiplex effector genes simultaneously, is outlined in **Figure 7.1**. Firstly, the effector genes were tested individually, across a range of four different gene dose levels, to provide indications of effective genes which improved cell line performance. Different gene loads should be tested as the effect on productivity is likely to be dependent on the amount of effector gene being expressed (Brown and James, 2016; Hansen et al., 2017; Johari et al., 2015; Tastanova et al., 2016). There is likely to be an optimal effective range of expression, which may be missed if only a singular expression level is tested. A high and a low impact level found for each functional gene would be useful when progressing to screening the effects of combinations of genes. The gene doses used were appropriate to enable the

same doses to be combined for co-transfections while maintaining consistent DNA load throughout.

When testing combinations of multiple genes, positive synergistic or additive effects are of interest. Implementing a DoE methodology is a powerful tool to search for interactive factors within experiments. When only one factor at a time (OFAT) is altered, it is difficult for interactions between factors to be identified. DoE provides a multifactorial approach, modelling responses and interactions mathematically, providing statistical analysis of significant interactions and variables tested simultaneously, meaning fewer experimental runs are required. A two-level factorial design approach can screen for interactions. Genes of interest, from singular gene screens, were assessed simultaneously, at low and high levels, for any significant positive interactions to the response outputs measured. Any significant interactions of interest could be investigated further using a DoE-RSM approach, where the factors (genes) are combined at different levels within a defined design space to further understand and visualise the impact of their interaction. Importantly, this technique can provide theoretical optimisation to help find ideal stoichiometric ratios of the interacting genetic effector components. The design space may need to be adjusted or narrowed in successive DoE runs to find this optimal region of interaction to best improve the cell factory performance.

To summarise, the screen enabled a large amount of functional effector components to be auditioned at different levels simultaneously. Each gene was tested individually (in triplicate) at four different levels with seven different components compared simultaneously per transfection plate. Effector genes could then be tested together, in a logical manner using a DoE methodology, with the aim to find and optimise any beneficial component combinations.

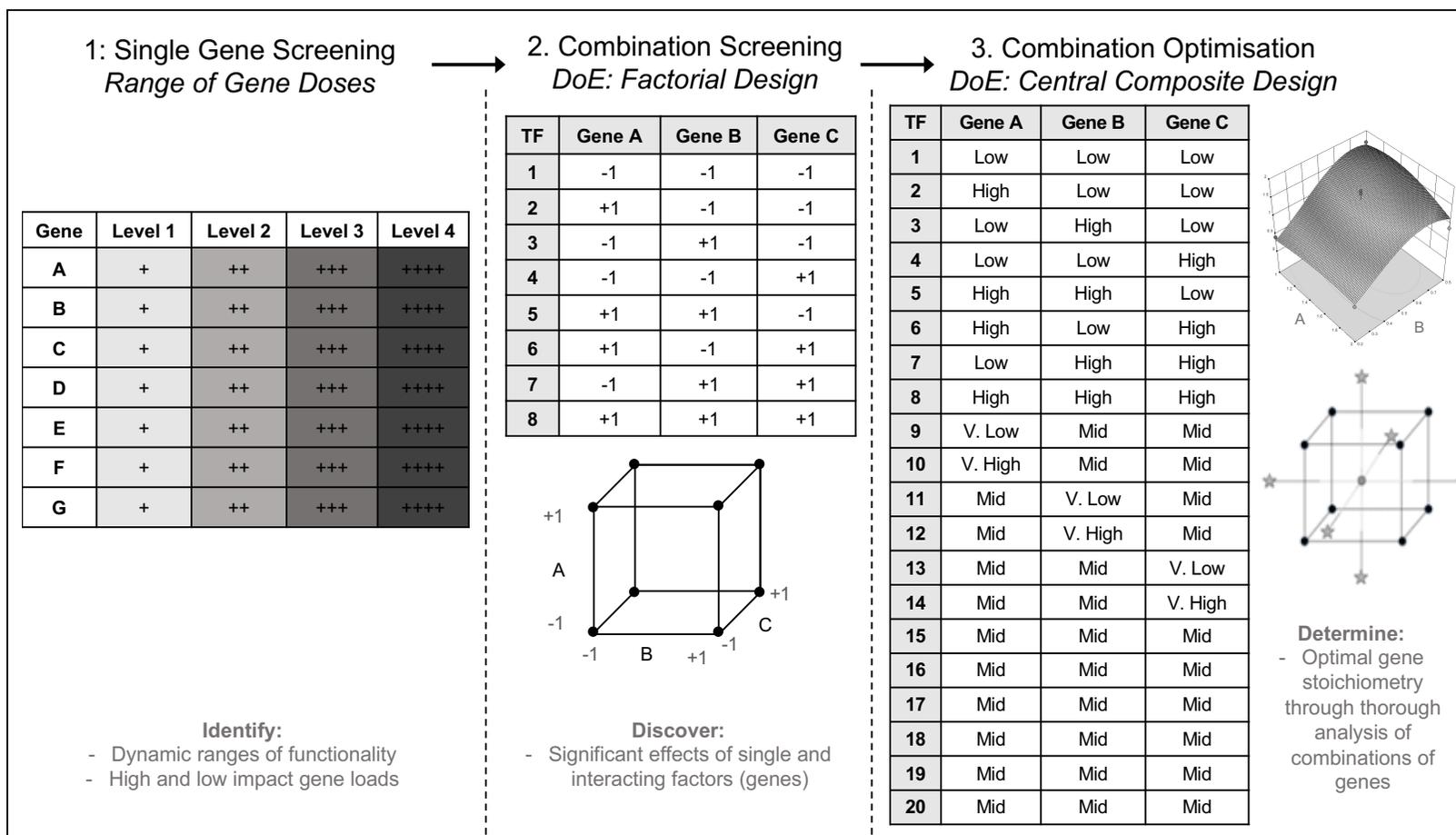


Figure 7.1: Effector Component Screening Strategy

The figure summarises the process used to screen many genetic components simultaneously. (1) Genes tested for their individual functional effects over a range of different doses. (2) Combinations of beneficial effector components are investigated in DoE two-level factorial designs. (3) Positive combinations of interest can be studied further and optimised using a DoE-RSM approach.

7.3. Results

7.3.1. Development of Effector Gene Library

As highlighted in the introductory chapters of the thesis (**Chapters 2 and 3**) there are numerous cellular components which partake, regulate and are associated with recombinant protein production. Knowing which components to target and test was not a straightforward task. Here, several sources of information (summarised in **Table 7.1**) were used to help provide a rational and logical process to generate a library of effector component targets to test within the high-throughput platform. One of the primary sources of information was the literature. There have been a wealth of previous studies (**Table 3.1-Table 3.5**), of varying success, describing overexpression of genetic components with the aim of improving recombinant protein production. These examples provided indications and some evidence of success of particular gene targets to support their inclusion in the library. In addition, the literature and online databases (for example NCBI/UniProt) provided an expansive knowledge about the function, or potential function, of particular proteins. Functions associated with protein processing and secretion could be of potential benefit if enhanced. The literature also provided information from CHO omic level analysis, with studies published which compare, highlight and describe transcriptomic and proteomic dataset analysis derived from producing and non-producing CHO cells (Carlage et al., 2012; Doolan et al., 2008; Harreither et al., 2015; Nissom et al., 2006; Yee et al., 2009). In most cases, producing cells were compared globally to non- or lower-producing cells. Components that changed between phenotypes, potentially as a cause or result of the change in production capability, can be inferred. Two in-house CHO omic datasets (from the University of Sheffield) were also used to provide similar indications, and evidence of presence, or absence, of particular genes/proteins. The other primary source of information used were two transcriptomic datasets derived from B cells and plasma cells, components of the natural immune system (B cells terminally differentiate into antibody-producing plasma cells). Plasma cells, being efficient and high-antibody secreting cell types (Nutt et al., 2015), are a good model for the phenotype being consistently sought after for CHO cell factories. The changes made, when B cells switch to plasma cells, are of interest

and help to define and explain the change to the desirable phenotype. Here, the comparative analysis was used to indicate genetic targets, which if upregulated might result in a higher producing phenotype.

A focus on targets relating to discrete functional modules of CHO cells (protein processing, folding and secretion), all significant for recombinant protein production, were used. Many previous examples have demonstrated that components of protein folding and assembly pathways can be altered to relieve cellular bottlenecks and improve recombinant protein producing capacity (Nishimiya, 2014; Schröder, 2008). Components were selected for study if they met more than one of the following criteria: if a gene had been previously overexpressed with successful outcomes; was associated with an improved production phenotype; was novel with a function of interest; or it correlated well to a high producing phenotype. The cost of DNA synthesis limited the library size generated. DNA synthesis is reducing in cost with time, demand and improved technology, but at approximately 18 pence per base pair, it was still a substantial cost when synthesising many genes.

Table 7.1: Multiple Information Sources Guided Effector Component Choice

Information Source		Reasoning
Plasma cell transcript abundance comparison to B cell transcript abundance		Plasma cells (natural antibody producing cells within the immune response (Nutt et al., 2015)) have a high capacity to secrete large quantities of antibodies. Plasma cells are terminally differentiated B cells; gene expression changes to B cells take place resulting in antibody secreting plasma cells. Transcriptomic datasets, derived from both RNA-Seq and Affymetrix analysis between B cells and plasma cells, identified changes associated with the antibody secreting phenotype of interest here.
Literature	Protein function	Evidence in the literature of a gene function highlighted potentially relevant targets to improve the production capability of the CHO cell factory.
	Previous engineered target	Evidence in the literature that a gene had been previously upregulated to result in productivity improvements demonstrated component potential.
CHO Omics	From the literature	CHO transcriptomic and proteomic studies published in the literature provided comparisons between: producing and non-producing CHO cell lines; low and high producers; and different points in culture. The studies provided evidence of components upregulated in association with increased production capacity.
	In-house proteomic dataset	An in-house proteomics dataset, derived from a mAb producing CHO cell line, provided increased evidence of presence/abundance of particular proteins in CHO cells.
	In-house transcriptomic dataset	An in-house transcriptomic dataset provided transcript abundances, at two time points in culture, between a MEDI-CHO producing cell line and the non-producing MEDI-CHO host.

The genes chosen and tested are summarised in **Table 7.2** and **Figure 7.2**, with further details about each gene function, reason for choice, species sequence source and gene ID summarised in **Appendix Table C.1-Appendix Table C.4** and nucleotide sequence in **Appendix Table C.5**. The genes were categorised into three principal functional groups. Many ER components are molecular chaperones, functioning in association with protein folding in the ER, many having holdase or foldase functions and/or involved with ER quality control. Other

components of this group were ER enzymes, which have roles in modifying and processing the recombinant protein to contain correct disulphide or peptide bonds and glycosylation patterns. Other components had roles in Golgi complex vesicular secretion, with functions stemming from export of proteins from the ER, vesicular transport through the Golgi complex (tethers/SNAREs) or to the plasma membrane or associated with Golgi complex lipid synthesis and expansion of the secretion system. A collection of transcription factors was also targeted. Transcription factors cascade their effects to alter many genes in concert. They are multigene engineering effectors, of interest as they may improve protein processing or secretion capacity in a global fashion.

Table 7.2: Effector Gene Library

The table lists the genetic effectors tested within the screening platform. The sequences either encoded a human (H) or CHO protein (C). Annotations show use of *=cleaved/cytoplasmic domain only, ^=active protein form (mutated to prevent deactivation) and ¹=components in first round of screening.

Transcription Factors	Protein Processing in the ER		Golgi Complex/Vesicular Secretion Processes		Other
ATF6 α c* (H) ¹	BiP (H) ¹	Grp94 (H)	Bet1 (C)	PREB (C)	Largen (H) ¹
CREB3L2* (C)	Calr (H)	HSPA1a (C)	Bet3 (C)	PRKD1 (C)	
TFE3 [^] (C)	CANX (H)	Hyou1 (H)	CERT [^] (C)	Rab1a (C)	
XBP1s (H) ¹	CypB (H) ¹	PDI (H) ¹	ERGIC53 (C)	Rab11a (C)	
XBP1s (C)	CRELD2 (C)	PDIA4 (H)	MCFD2 (C)	SGMS1 (C)	
YY1 (C)	Dad1 (C)	PRDX4 (H)	p115 (C)	Stx5a (C)	
	ERo1L α (H) ¹	SRP14 (C)			
	FKBP11 (H)	Tor1a (H)			

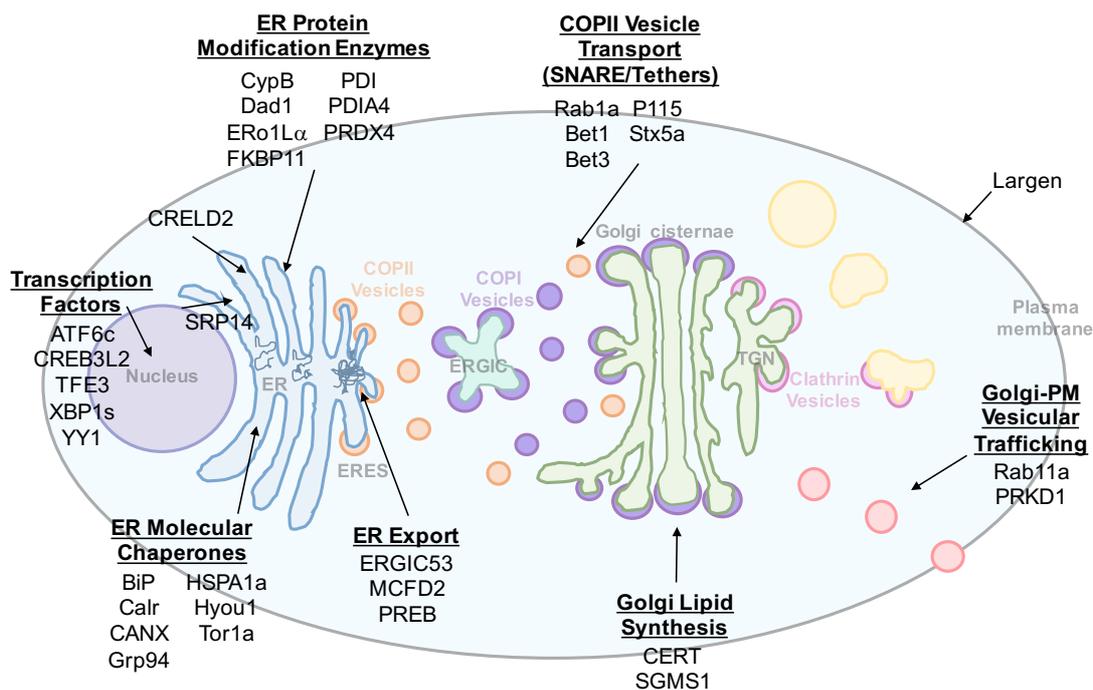


Figure 7.2: Overview of Effector Genetic Components

Genes from the generated library of effector components are grouped by their general function and displayed at their approximate functional site in the cell/secretory pathway (further functional information can be found in **Appendix Table C.1-Appendix Table C.4**).

The gene target sequences synthesised were derived from public online databases (NCBI, Uniprot and CHO Genome), with genes identified by their primary gene names. Protein and nucleotide sequences, for both human and CHO species, were compared for sequence confidence. For ATF6 and CREB3L2, only the DNA encoding the active, cytoplasmic region of the protein was synthesised. Similarly, for XBP1, the spliced active version was synthesised. For two proteins, TFE3 and CERT, a mutation in one amino acid was made to express the active, dephosphorylated form.

For the first seven genes tested (ATF6c, BiP, CypB, ERO1L α , Largen, PDI and XBP1s) encoded human protein sequences. The other genes synthesised were either human or CHO protein sequences. The proteins overexpressed previously in the literature were primarily human proteins. The advantage of using human sequences is that the annotation is currently generally better than for the *C. griseus* genome and proteome; meaning there is an increased confidence in a human gene sequence making a correctly functioning protein, but there is the

risk the protein will not function appropriately in CHO cells. The annotation and information available for the CHO cell genome is improving and the rodent connection to well-characterised murine sequences is a useful trait that can be used to increase confidence in CHO protein annotations. Many recombinant biotherapeutic proteins are human or humanised mAbs. Therefore, for target proteins which primarily associated and functioned directly with the recombinant protein itself, human genes were used; for example, ER chaperones, foldases and holdases (e.g. Grp94) which interact directly with the recombinant polypeptide chain during its processing. For proteins that mainly interact with the CHO cell machinery or other CHO cell proteins, *C. griseus* protein sequences were used to try and ensure the expressed component interacted with its associated cellular proteins and processes; for example, transcription factors and SNARE/tether proteins.

7.3.2. Range of mAb Expression with Platform Setup

Two model mAb-expressing plasmids were provided by MedImmune for the purposes of this project. The two human mAbs were both IgG1 molecules, one characterised as an ETE antibody (Nip109, referred to as ETE-mAb) and the other as a DTE antibody (DISCO, referred to as DTE-mAb) with lower titres (MedImmune, personal communication,). The two recombinant products had similar heavy chain constant domains (encoding gamma 1 isotype), but differed in their light chain sequences, kappa and lambda for the ETE and DTE molecules, respectively. The plasmids encoding the two molecules contained identical sequences for GS expression, identical promoters (CMV) for both the HC and LC and identical polyA terminating sequences for the HCs. The polyA termination sequence for the light chains differed between the two molecules. The ETE and DTE plasmids were similar in size, being 9537 and 9524 base pairs, respectively.

Transient expression, with varied of amounts of each mAb plasmid transfected (0-800ng), shown in **Figure 7.3**, demonstrated that the DTE-mAb had a significantly reduced production rate, when compared to the ETE molecule. There was an approximate eight-fold reduction in specific production rate, also mirrored in the volumetric titres measured (data not shown). For both mAbs, the same trend in post-transfection viability and growth was seen. The more

recombinant mAb plasmid transfected and expressed, a greater negative effect, albeit small, on culture viability and growth was observed; likely to be partly due to the increased biosynthetic burden being placed on cells. The fact that the DTE-mAb plasmid had the same effect within cells as the ETE-mAb plasmid implied that the difficulties in expression were not causing a cytotoxic/damaging effect on the production host under these conditions. The disparity in expression appears to descend from differences in the ability of cells to synthesise and secrete the DTE molecule compared to the ETE molecule. There was potential for improvement in production for both proteins, though particularly for the DTE molecule; the cell can function to accommodate increased protein production and secretion, with a saturation point not being reached with the tested range (0-800 ng; **Figure 7.3**). As mentioned, the DTE molecule will be the model mAb used within these experiments; the molecule having greater requirement and scope for improvement.

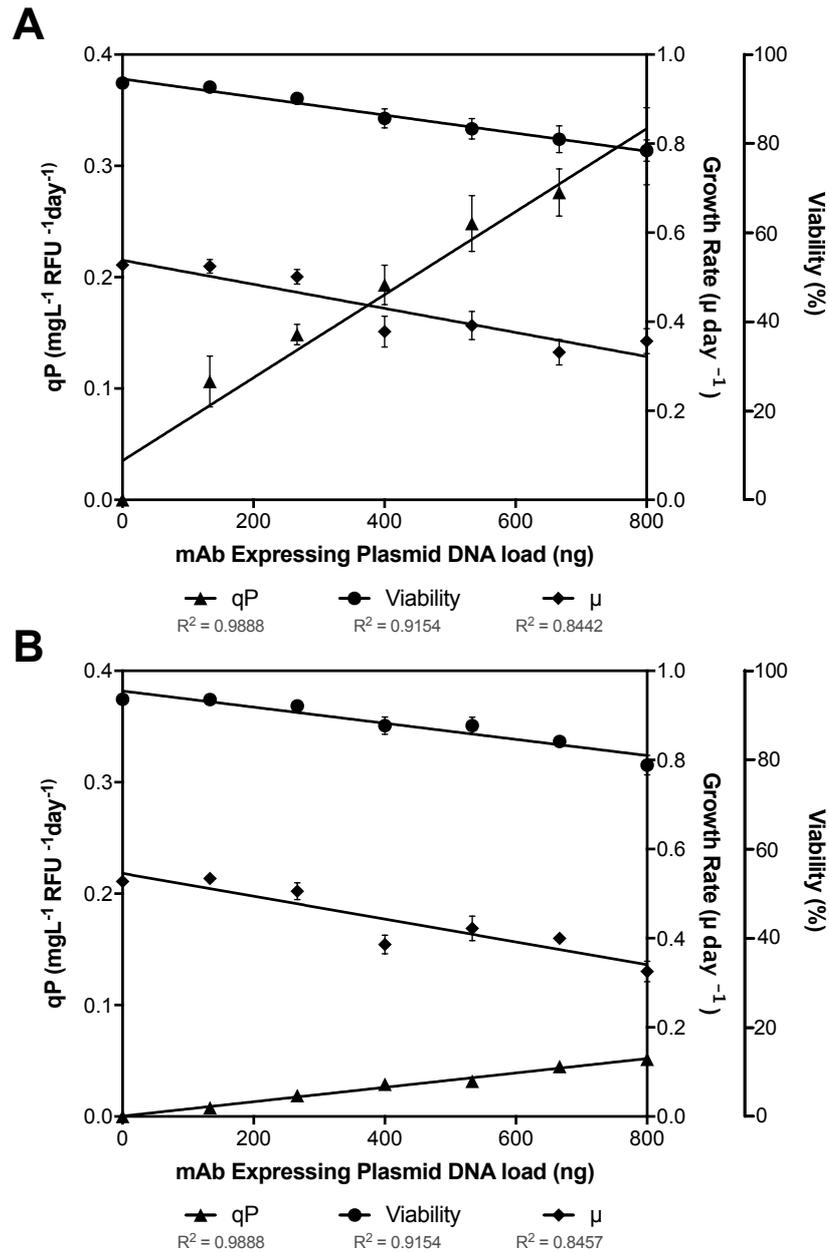


Figure 7.3: Effect of Varying Transfected mAb-Plasmid Type and Copy Number

Varying amounts of recombinant mAb plasmid were transfected, with total DNA kept constant (800 ng) using empty-‘Expression-Plasmid’. After 72 hours, qP, growth and viability were measured, from expression of the ETE-mAb (A) and DTE-mAb (B). The data represent the average and SEM across two biological replicates, each with three technical replicates.

7.3.3. Effect of Multiple Gene Expression

It was identified that to express the effector genes alongside a recombinant product gene there was a resultant additional biosynthetic burden on cells to express multiple genes simultaneously. A reduction in productivity output (titre and qP) was seen from co-transfections with altered amounts of eGFP plasmid (mimicking effector gene expression); shown in **Figure 7.4**. The additional biosynthetic burden of GFP expression does not affect cell growth (IVCD) within the 72-hour culture period, but titre and qP were both, in general, reduced slightly as a result of additional eGFP expression. At the highest GFP load tested, there was a reduction in productivity (for both titre and qP) of 25%, on average. This indicated that when effects were compared to control samples containing no effector genes, the percentage changes measured did not account for an additional biosynthetic burden resulting from the additional and altered gene expression from the co-expression effector gene(s). Percentage changes (in growth and productivity relative to control samples containing no effector genes) measured could therefore have been slightly dampened due to this concept. Each effector gene could potentially have different biosynthetic burden effect on cells and their productivity, making it hard to define and accommodate. It can be assumed, from this example, that an alteration does occur from additional gene expression, but it is not that considerable. Ultimately, only effector genes that cause significant effects, and improve the cell factory's production capability despite any additional biosynthetic burden, are of real long-term interest. A key result is that increased eGFP plasmid transfected resulted in increased GFP expression (MFI). This provided evidence to confirm the platform setup accomplished titratable and specified gene expression levels just by altering the volume (and therefore quantity) of plasmids transfected.

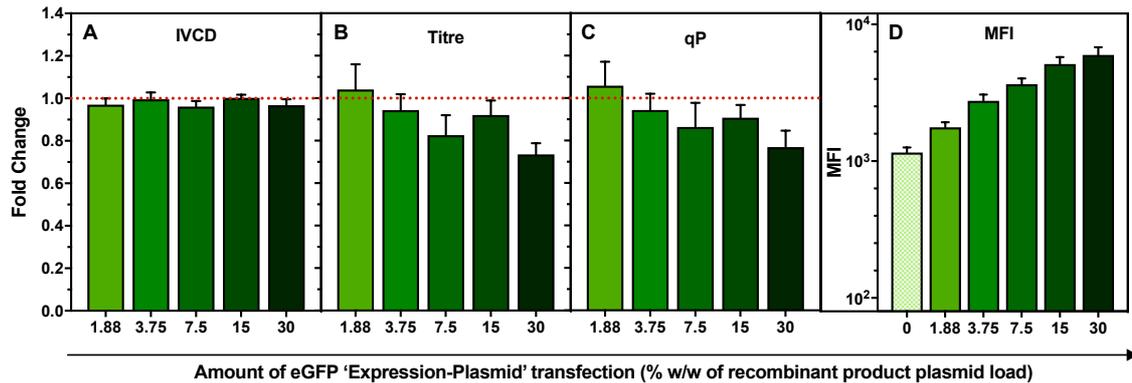


Figure 7.4: Alteration in eGFP-'Expression-Plasmid' Levels Demonstrated Effects Upon Recombinant mAb Co-Expression

DTE-mAb plasmid was co-transfected with 1.88-30% of eGFP plasmid, (w/w) relative to the amount of mAb-plasmid (533 ng); empty-'Expression-Plasmid' equalised total DNA load to 800 ng across all transfections. Control transfections, without eGFP plasmid, were used for data normalisation. The effects upon (A) IVCD, (B) titre, (C) qP and (D) MFI were measured after 72 hours, shown is the average and SEM from three biological replicates, each with two technical replicates.

7.3.4. Initial Effector Gene Screen Demonstrated the Importance of Transfect Gene Dose

An initial screen was undertaken with seven components from the effector gene library: ATF6c, BiP, CypB, ERO1L α , Largen, PDI and XBP1s. These genes were selected due to their previous success in improving DTE protein production within the DCJ laboratory at the University of Sheffield (ATF6c, BiP, CypB, PDI and XBP1s; Johari et al., 2015; Pybus et al., 2013), previous successful combinatorial effects/potential (ERO1L α ; Cain et al., 2013; Mohan and Lee, 2010), or for their novelty and functional potential (Largen; Yamamoto et al., 2014).

The initial screen of the seven genes formed part of a proof of concept of the platform setup, first testing the platform for individual gene functionality then combinatorial effects. A static culture mode was used, with effects on growth and productivity measured after 72 hours. Improved culture performance, mediated by different modes of actions (IVCD or qP) was of interest. The effects from four different levels of each gene (15, 7.5, 3.75 and 1.88% (w/w) of recombinant DTE-mAb plasmid load), as a percentage of control transfections containing no effector genes, are shown in **Figure 7.5**.

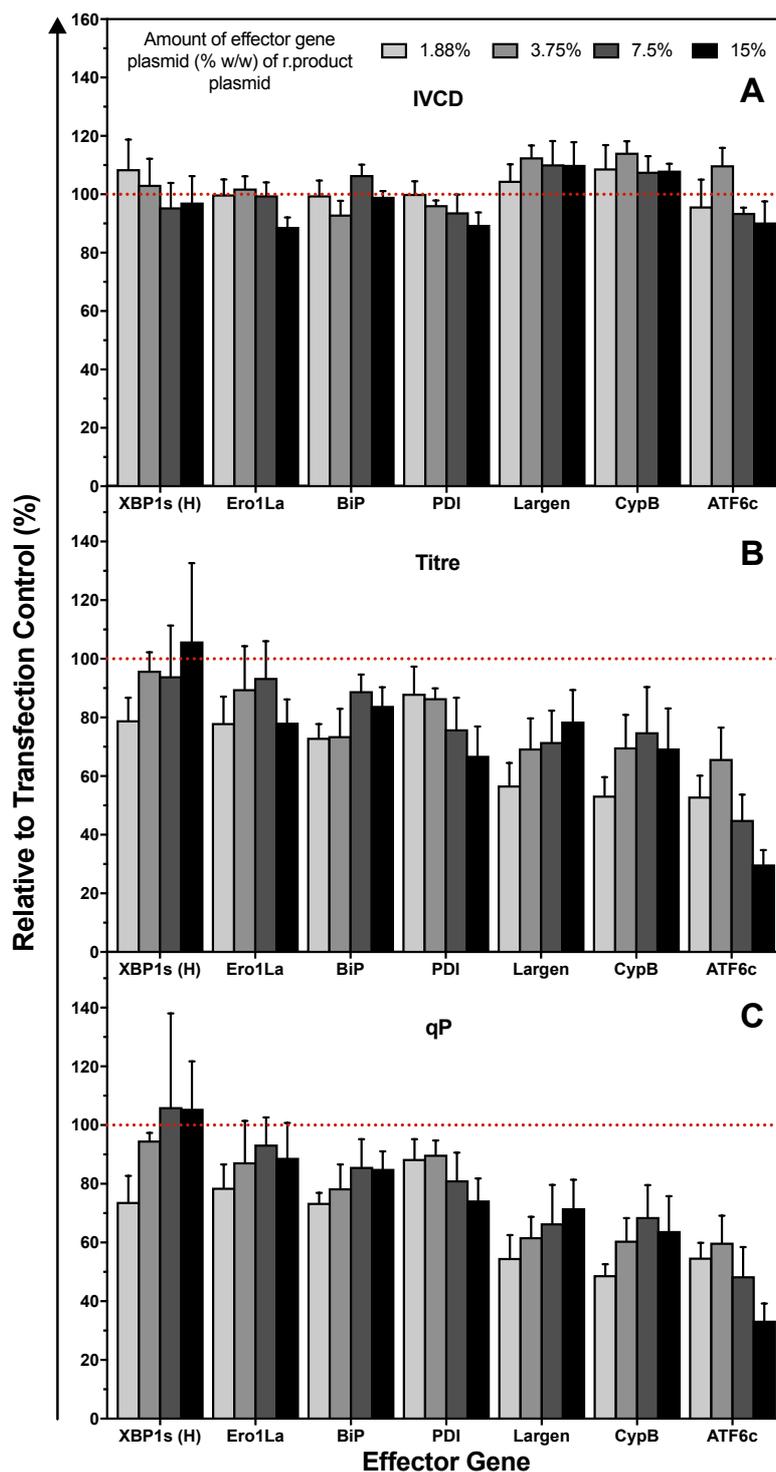


Figure 7.5: Static Screen: Effects from Co-Transfected Effector Genes at Different Plasmid Doses

Seven effector genes co-transfected at four levels: 15, 7.5, 3.75 and 1.88% (w/w) of DTE-mAb plasmid (533 ng); empty empty-‘Expression-Plasmid’ equalised plasmid load across transfections. (A) IVCD, (B) titre and (C) qP were calculated after 72 hours of static growth. Data is shown as percentage of control transfections (red dashed line) containing no effector gene expression. The data represent the average and SEM of three biological replicates, each with three technical replicates.

Differences and patterns derived from altered effector gene doses transfected were apparent. When very small amounts of effector gene were expressed, small functional effects on the CHO cell factory were apparent; low copy numbers and transfected gene loads are sufficient to perturb cellular systems. The trend derived from testing different levels of gene expression (for XBP1-s, ER α , BiP, Largen and CypB) demonstrated positive trends in functionality. Increased DNA loads of these genes had greater beneficial effect upon productivity.

Component specific effects were seen, with XBP1s (H) overexpression having the most promising positive effect upon production of the DTE-mAb. As the dose of XBP1s plasmid increased, the volumetric titre produced further increased. The improvement was mediated by increased specific productivity indicating the effect was derived from an increased cellular production capability. The trend from XBP1s overexpression was apparent despite the variability present; substantial variation of this nature was not ideal as it reduces the chance and confidence in trends/effects being measured. Some of this variation derived from the culture setup (as discussed in **Section 5.3.6**) is mitigated through the use of a shaking deep-well plate culture format (**Section 5.3.4**) and this format was used for subsequent experiments described in this chapter.

Expression of the UPR transactivator gene, ATF6c, resulted in considerable negative effects on recombinant protein production, with increased levels transfected worsening the negative effects. The outcome was primarily mediated by a reduction in qP, but was combined with a small reduction in accumulation of cell biomass with both IVCD and viability reduced after ATF6c transfection. After 72 hours, the viability was approximately 10% lower than all other transfections undertaken (data not shown) suggesting there was increased apoptosis, an unwanted downstream effect from ATF6c expression/activation.

Minimal changes in growth were apparent within the 72-hour, static culture period used here. This may reflect the limitations of the short duration, sub-optimal growth conditions. Overexpression of Largen and CypB genes resulted in approximately 10% improvement in biomass accumulation (IVCD), across all gene doses. This effect, if extrapolated, would increase the population of producing cells later on in culture; this would need to be balanced with the negative impact on specific productivity seen. For several genes (XBP1s,

ERo1L α , BiP and PDI) their expression resulted in a negative impact on IVCD, with greater effects as the transfected gene level increased. This trend could have been due to the additional biosynthetic burden introduced from increased effector gene expression or due to the effect of the expressed protein itself. Either way, the negative impact was not desirable for constitutive gene expression, with the gene(s) being expressed throughout culture, this negative impact on cell growth would compromise the accumulation of cellular biomass.

7.3.5. Combinatorial Expression of Effector Genes

7.3.5.1. Two-Level Factorial Design

The progression to screening the overexpression of multiple genetic components was demonstrated with four genes from the initial singular gene screen. ERo1L α /XBP1s both demonstrated enhanced production mediated by increased specific production and CypB/Largen demonstrated improved cell growth (IVCD) capability. The effect (on IVCD, volumetric titre and specific productivity responses) from simultaneous overexpression was measured after 72-hours of static culture, and presented as a percentage relative to control transfections that contained no effector gene expression. A two-level factorial DoE design was used to establish whether there were any interactions present. All four genes were each tested at a high (+1) and low (-1) level, with no overexpression (0 ng) used for the low-test levels (to help infer if an effect was due to the presence of a particular gene or not) and a high gene level 7.5% (40 ng) (w/w) of DTE-mAb plasmid load. This DNA load for all genes in the singular gene screens, had the most beneficial effect on recombinant protein production, without being too detrimental to growth (particularly the case for ERo1L α). The four genes, known as factors, tested at two levels in a full factorial design, resulted in 16 test combinations (outlined in **Appendix Table C.6**). The results from the three measured outputs are shown in **Figure 7.6**.

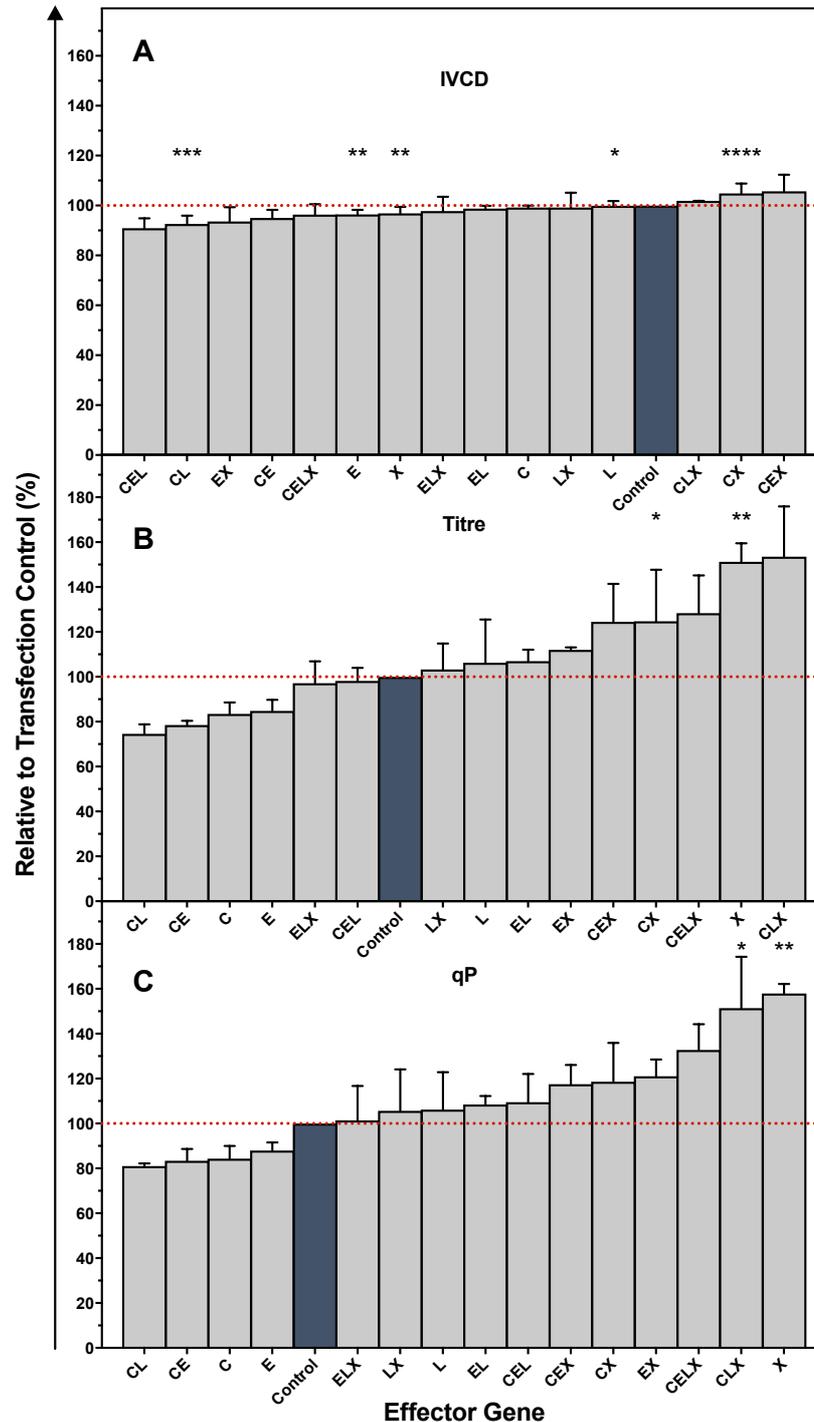


Figure 7.6: The Effect from Multiple Effector Component Co-Transfections

Two-level factorial DoE design investigated multi-component co-transfections. Four different effector genes CypB, ERO1 α , Lagen and XBP1s (labelled in the figure as C, E, L and X, respectively) were transfected in combinations at two different levels: 0% and 7.5% (w/w) of DTE-mAb plasmid load (533 ng). The effect upon (A) IVCD, (B) titre and (C) qP after 72 hours of static culture is shown relative to the transfection control (no effector gene expression). The data represent the average and SEM of two biological experiments, each with three technical replicates. Statistically significant influential factors, analysed using ANOVA, are shown by asterisks (**** P<0.0001, ***P<0.001, **P<0.01, *P<0.05).

Using DoE software (Design-Expert® 10.0.6.0) half-normal probability plots were generated for each response output (IVCD, titre and qP; **Appendix Figure C.1**). These were used to identify the most influential factors that affected the response outputs. Models describing the relationship between the outputs and the influential factors were generated and analysed for significance (using ANOVAs). The models are described by the following equations, with the positive/negative influence of each factor on the output shown by the positive and negative coefficients, respectively, (A, B, C and D refer to CypB, ERo1L α , Largen and XBP1s, respectively).

$$IVCD = 98.06 + 0.24A - 1.24B - 0.89C + 1.46D - 1.99AC + 2.42AD \quad (7.1)$$

$$Titre = 107.97 + 0.26A + 0.56C + 16.38D + 4.85AC + 8.17AD - 4.34CD + 7.07ACD \quad (7.2)$$

$$qP = 110.46 - 0.64A + 1.61C + 15.34D + 7.25AC + 4.92AD - 4.59CD + 7.75ACD \quad (7.3)$$

All models were statistically significant and significant factors within each were highlighted, this analysis is summarised in **Table 7.3**. For some models, predictability was better than for others; likely to be due the inclusion of fewer insignificant terms within the model, otherwise required for model hierarchy. The use of a DoE factorial design approach was a useful and effective tool to test for interactions between components. The Design-Expert® software provided a straight-forward indication, with statistical significance of influential factors and highly impacting interactions; proving very useful to screen for successful synergistic gene combinations.

From the coefficients for each factor in the equations and the statistical significance, the most influential factors affecting the outputs were identified. For IVCD, several factors and combinations significantly negatively affected cell growth (ERo1L α , Largen and CypB/Largen). The most significant beneficial factor was an interaction between CypB and XBP1s. However, as shown in **Figure 7.6**, no major differences in IVCD resulted from any combination of effector gene overexpressed when compared to control transfections.

Table 7.3: Analysis Summary of Outputs from the Two-Level Factorial Design Used to Investigate Combinations of Effector Genes

Factors A, B, C and D correspond to effector genes CypB, ER α , Largen and XBP1s, respectively. Normal plots of residuals demonstrated normality of residuals for all models (Appendix Figure C.1).

Response	Factor	Sum of Squares	P Value (<i>significant factors, P < 0.05, in bold</i>)	Model Predictability (Pred/Adj R ²)	Adeq Precision
IVCD	Model	229.40	<0.0001	0.79/0.89	17.73
	A	0.94	0.4957		
	B	24.71	0.0054		
	C	12.70	0.0282		
	D	34.01	0.0021		
	AC	63.54	0.0002		
	AD	93.49	<0.0001		
Titre	Model	6847.29	0.0170	0.28/0.66	6.17
	A	1.07	0.9419		
	C	5.07	0.8740		
	D	4925.03	0.0014		
	AC	376.73	0.1956		
	AD	1068.77	0.0446		
	CD	301.04	0.2423		
	ACD	799.59	0.0736		
qP	Model	6340.73	0.0151	0.30/0.67	6.36
	A	6.56	0.8483		
	C	41.23	0.6336		
	D	3767.41	0.0015		
	AC	841.42	0.0556		
	AD	386.90	0.1675		
	CD	336.60	0.1946		
	ACD	960.61	0.0437		

Greater effects were apparent for titre and specific production responses. As seen in **Figure 7.6 B and C**, many combinations of effector genes improved production beyond that of control transfections. From the statistical analysis of the model and the coefficients in Equation (7.2), the most significant factor impacting the titre was overexpression of XBP1s. Additionally, the combination of CypB and XBP1s was significant in positively affecting the titre and there were several model terms with positive or negative impact on titre but did not reach statistical significance.

A similar scenario for qP was seen (Equation (7.3)). Again, XBP1s was the most significant and influential factor contributing to an increase in qP. Any transfection containing XBP1-s overexpression, on average, improved qP beyond the control transfections (**Figure 7.6**). In addition, there was a significant positive effect on qP derived from an interaction between overexpression of CypB, Largen and XBP1s. It should also be noted that there were several insignificant terms influencing qP, but to a lesser extent. In particular, almost significant ($p=0.0556$) was an interaction between Largen and XBP1s co-overexpression, which had negative effect on both qP and titre.

These data suggest that XBP1s overexpression is a consistent effector gene which improved production levels. There were potentially beneficial interactions when XBP1s overexpression was combined with other genes - CypB and, to a lesser extent, Largen. This combination was investigated further, as described the next section, using a DoE-RSM design. No substantial changes in IVCD were made for any overexpression combination, but beneficially no gene had a major negative effect on IVCD.

7.3.5.2. Central Composite Response Surface Model Design

As effector genes are likely to have enhanced or added benefits when overexpressed in combinations with other effector gene engineering strategies, an approach aimed at optimising solutions was required. To demonstrate how to investigate and optimise combinations of interacting factors a DoE-RSM design was implemented. Following on from the DoE two-level factorial design, simultaneous overexpression of CypB (A), Largen (B) and XBP1-s (C) at different levels was used as some interaction between these genes was indicated. A three-factor (A, B, C), two-level (high and low) rotatable central-composite design was used. Use of a design of this type should increase model predictive power, help the entire design space to be navigated and allow curvature of interactions to be predicated, if present.

The levels of overexpression used were based upon the findings from the single gene overexpression studies. Each effector was overexpressed in every transfection with a range of levels used (summarised in **Table 7.4**). A high and low level of each gene formed the boundaries of the design space, with axial

factors (very low and very high levels) used to help generate quadratic models with curvature predictions. A mid-point level, repeated six times, provided an estimate used to account for pure error within the models generated. A DoE approach enabled experimental combinations to be analysed from fewer experimental runs than an OFAT approach. Here, 14 different combinations, including 6 repeated mid-points, made up the experimental runs undertaken; transfections were undertaken independently on three discrete occasions and averaged to account for any biological or transfection variation.

Table 7.4: Levels of Factors Used in Central Composite Design Combining CypB, Largen and XBP1s Effector Genes

	Low	High	Mid	-Axial	+Axial
A: CypB (ng)	14	32	23	7.86	38.14
B: Largen (ng)	25	70	47.5	9.66	85.34
C: XBP1s (ng)	20	53.33	36.7	8.64	64.7

The effect of combinatorial overexpression on IVCD, volumetric titre and specific productivity responses, was measured after 72-hours of static culture, with all effects represented as a percentage of control transfections that contained no effector gene expression (**Figure 7.7**). The response outputs were analysed in Design-Expert® software and models were generated for each to describe any relationships present. They were analysed for variance and significance using ANOVAs. No data transformation was recommended by the software for any model. For IVCD, no significant effects were measured and no factors impacted on IVCD substantially as shown in **Figure 7.7** (all combinations reduced IVCD marginally). There was an insignificant lack of fit and an adequate signal to noise ratio (>4), which meant the design space could be navigated, but as there were no significant terms and the model was insignificant, IVCD was not analysed any further here.

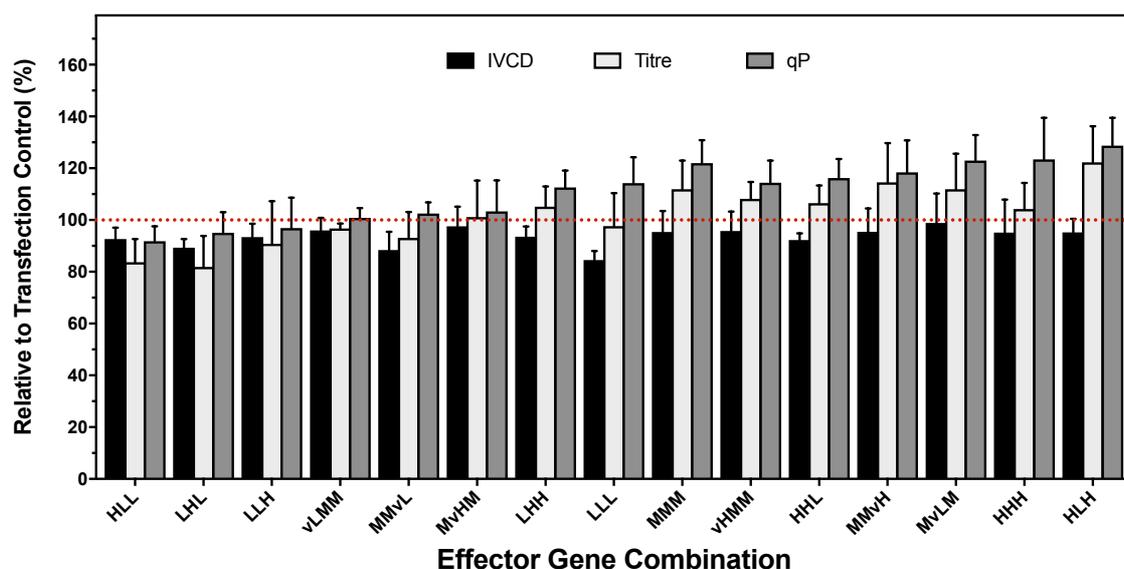


Figure 7.7: The Effect from Simultaneous Co-Transfections of Multiple Effector Components at Differing Levels

A DoE-RSM design investigated interactions from multiple component co-transfections. Effector genes CypB, Largen, and XBP1s, were co-transfected together over a range of levels: L=low, M=medium, H=high, vL=low axial and vH=high axial levels (notated in the figure in the order of CypB-XBP1s-Largen). The effect upon IVCD, titre and qP after 72 hours of static culture is shown relative to the transfection control (cells transfected with no effector genes). Shown is the average and SEM of three independent experiments.

The Design-Expert® software provided an indication of the order of model to use, with a linear model recommended and used to describe titre; and a two-factor interaction (2FI) model for qP response. There were found to be two insignificant model terms (AB (CypB/Largen) and BC (Largen/XBP1s) for qP. When these interactions (from overexpression of Largen with the other effector genes) were included, they affected the fit of the model to the data and removal of these (model reduction) improved the model's significance and predictability (improved predicted versus adjusted R²). The equations used to describe the two response output models are below (Equations (7.4) and (7.5)). Coded factors are used to identify the relative impact of factors, by comparing the factor coefficients (positive or negative coefficients indicate a positive or a negative factor impacts, respectively; A, B, C refer to CypB, Largen and XBP1s, respectively).

$$\text{Titre} = 102.90 + 2.14A - 3.37B + 8.80C \quad (7.4)$$

$$qP = 110.92 + 1.87A - 4.11B + 8.04C + 5.49AC \quad (7.5)$$

The statistical analysis (summarised in **Table 7.5**) undertaken on the models outlined above, demonstrated that the titre and qP models were statistically significant. It also provided insight into which factors significantly impacted the responses. Both models were highly significant, fit the data and variance well, had insignificant 'lack-of-fit', had predicted and adjusted R² values within agreement of each other (within <0.2 difference) and had adequate signal to noise to sufficiently model the design space ('adeq precision' >4).

Table 7.5: Summary of Response Model Analysis of Outputs from the Central Composite Design

Factors A, B and C correspond to effector genes CypB, Largen and XBP1s, respectively. Normal plots of residuals demonstrated normality of residuals for all models (**Appendix Figure C.2**).

Response	Factor	Sum of Squares	P Value (<i>significant factors, P < 0.05, in bold</i>)	Model Predictability (Pred/Adj R ²)	Adeq Precision
Titre	Model	1274.79	0.0060	0.33/0.44	7.87
	A	62.31	0.3615		
	B	155.51	0.1572		
	C	1056.97	0.0014		
qP	Model	1401.51	0.0005	0.44/0.64	11.63
	A	47.64	0.2729		
	B	230.89	0.0242		
	C	882.15	0.0002		
	AC	240.82	0.0218		
IVCD	Model	79.52	0.1446	-0.09/0.15	4.41

For each significant model, the responses measured within the design space were represented as response surface three-dimensional plots (**Figure 7.8**), providing visual representation of interactions and how each factor changes the response output across the design space. From statistical analysis of titre responses (**Figure 7.8 A, B, C**), and the factor coefficients in Equation (7.4), overexpression of XBP1s positively and significantly impacted titres. It was the major contributing factor to improving titre beyond that of control transfections. Although not statistically significant, Largen overexpression negatively impacted titre, whereas CypB introduced small improvements. There were no significant interactions between factors.

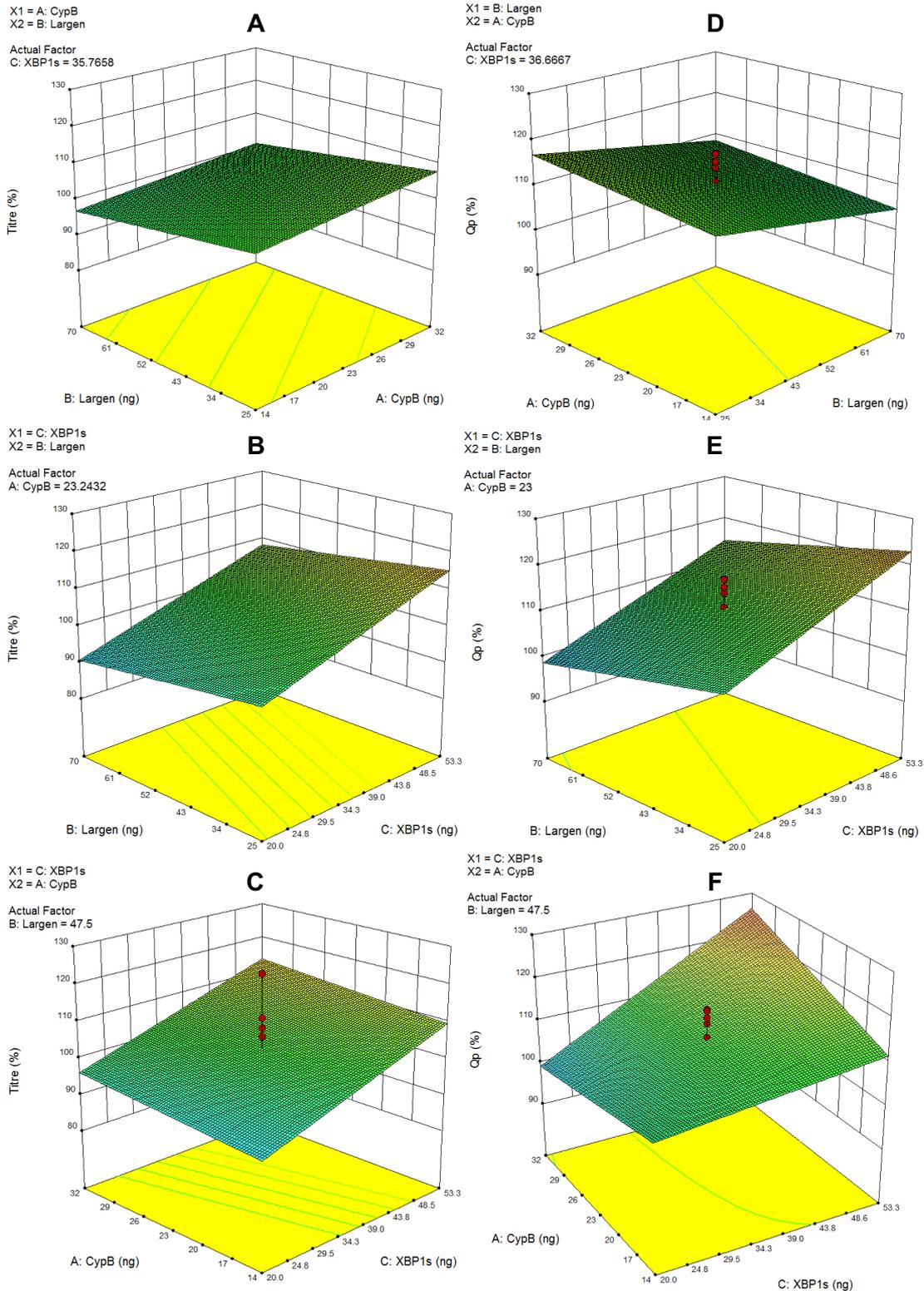


Figure 7.8: Three-Dimensional Surface Graphs Representing Titre and qP Responses from Co-Transfections with CypB, Lagen and XBP1s

Three-dimensional response surface plots illustrate the modelled responses of titre and qP after 72 hours of culture. Titre/qP are shown as a function of co-transfected (A)/(D) CypB and Lagen, with XBP1s fixed at 36.7 ng (B)/(E) Lagen and XBP1s, with CypB fixed at 23 ng (C)/(F) CypB and XBP1s, with Lagen fixed at 47.5 ng.

Statistical analysis of qP responses (**Figure 7.8 D, E, F**) and the factor coefficients in Equation (7.5) demonstrated that the most significant impact on qP was from the overexpression of XBP1s ($P=0.0002$). It substantially and positively improved specific productivity beyond that of control transfections. Overexpression of Lagen alone also significantly affected the specific production response but in a negative manner. As shown in **Figure 7.8 D, E** there was an undesirable downward trend in qP from increasing Lagen overexpression. Co-overexpression of CypB and XBP1s had a significant positive interaction on qP response, demonstrated in **Figure 7.8 F**. The positive impact of XBP1s was enhanced further when higher levels of CypB were co-overexpressed. Reduced Lagen overexpression improved the qP and titre responses across all CypB/XBP1s combinations within the range of the design space (data not shown).

To summarise, from the DoE-RSM combining CypB, Lagen and XBP1s overexpression, Lagen overexpression introduced undesirable effects to production when combined with CypB and XBP1s overexpression. Conversely, XBP1s overexpression was beneficial and created the most significant improvements in production (titre and qP). There was a positive interaction from CypB and XBP1-s co-overexpression which improved specific production levels. Even with the presence of Lagen at low levels (25 ng), CypB and XBP1-s, overexpression (at the highest levels test) improved qP by ~29% (and titre by ~23%). Co-overexpression of these components is therefore of interest to improve the manufacturing capability of the test DTE-mAb, but the interaction should be investigated further. Using the numerical optimisation function, the Design-Expert® software could generate theoretical predictions to try to achieve a desired specific output objective, here - to increase production output beyond that of control transfections, to maximise titre (and qP) without being too detrimental to cell growth. A range of optimal solutions within the design space were suggested: for example, co-transfection of 43, 15 and 85 ng of CypB, Lagen and XBP1s to improve qP by ~80% and titre by ~38%, while not affecting growth. If Lagen overexpression was removed an optimal loading of 58 ng of CypB and 113 ng of XBP1s was suggested to improve IVCD marginally (7% increase), and titre and qP considerably (by 58% and 160%, respectively). These

recommendations would need to be tested experimentally to confirm that these increases were possible. The range of solutions demonstrated that XBP1s and CypB would have had more impact and benefit to improving productivity if used beyond their highest level within the design space here (>53.3 and >32 ng, respectively). A second DoE screen is recommended, with a higher design space and the absence of Largin overexpression.

7.3.6. Rapid Auditioning of the Effector Gene Library

7.3.6.1. Use of Deep-Well Shaking Culture Mode Improved Post-Transfection Culture

The functionality of the effector genes expressed so far were limited, and only marginal effects were measured. This was potentially attributable to the sub-optimal post-transfection culture mode used, which hampered post-transfection culture and recovery. Employing a more optimal culture mode and extending culture duration slightly was anticipated to improve the developed transfection platform and therefore the screening process. As outlined in **Section 5.3.4**, the use of a deep-well shaking culture mode could introduce several benefits, including increased culture volume (allowing for parallel response measurements), increased culture performance (for example, improved gas transfer) and extended culture longevity. The transient cell line CHO-T2 could be comparably transfected using the transfection platform and growth in deep-well plates enabled implementation of the shaking culture setup and therefore facilitated improvements in the screening process.

The CHO-T2 cell line was previously engineered to express EBNA-1 and GS elements which enhanced and prolonged expression of transfected plasmids containing OriP DNA elements (Daramola et al., 2014). To make use of this functionality, to help maintain and improve plasmid expression, the MedImmune antibody expression plasmids were modified to contain an OriP genetic element. The OriP element was cut from an existing MedImmune plasmid using SapI restriction sites and cloned into the existing plasmids encoding for the mAb. This resulted in the plasmid being ~2 kb larger (approximately a 20% increase). The

recombinant product plasmid used herein contained the DTE-mAb with the OriP element (referred to as DTE-mAb-OriP).

Transient transfections with DTE-mAB-OriP were undertaken, then cultures were grown in the semi-optimised shaking mode in deep, 96-well plates. Daily VCD and viability measurements were taken from separate culture wells and cells were fed using an optimised fed-batch regime to boost post-transfection recovery, growth and longevity. The time points where media was supplemented are highlighted in **Figure 7.9**, which shows the growth profile of the transiently producing CHO-T2 cells grown in the semi-optimised shaking, deep-well culture mode. Day five was used as the sampling point for screening experiments, where the cells were in the exponential (growth) phase of culture ($VCD = \sim 3 \times 10^6$ cells mL^{-1}) and had recovered from the electroporation procedure. The extended culture duration helped to amplify effects compared to the day three static screen, but a fairly rapid process was maintained to enable screening throughput. Gene expression was enhanced (average volumetric titre measured on day five post-transfection was 1.29 mg L^{-1}), greater than from static 72-hour culture setup (average titre of 0.22 mg L^{-1}). Additionally, longer duration cultures in subsequent, more comprehensive experiments, could be investigated.

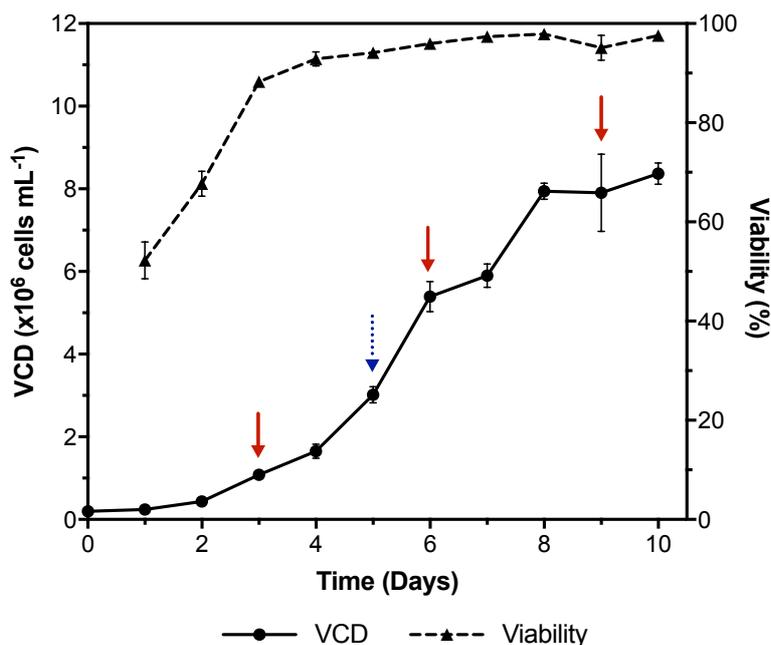


Figure 7.9: Shaking, Fed-Batch Culture Mode Using Deep, 96-Well Plates

CHO-T2 cells, transiently transfected with DTE-mAb-OriP plasmid, were combined and seeded into multiple wells in deep, 96-well plates. VCD and viability were measured daily, from separate wells, using the Vi-Cell XR. Titre was measured on day five (blue arrow). Red arrows indicate feed points, where 10% of the culture volume was replaced with a 1:1 ratio of Efficient Feed A+B. The data points represent the average and SEM of two independently performed experiments, each with three technical replicates.

7.3.6.2. High-Throughput Effector Component Screen

The deep-well shaking culture mode and CHO-T2 cell line were used to screen the remaining 28 components in the effector target library (outlined previously in **Table 7.2**). The 28 different genes were each overexpressed in the high-throughput transient transfection platform at four levels (15, 7.5, 3.75, 1.88% (w/w) relative to the loading of the DTE-mAb-OriP plasmid co-transfected). This demonstrated that large numbers of engineering strategies could be auditioned rapidly and simultaneously in-parallel; with seven effector genes, each altered at four different levels, compared per transfection plate with multiple transfection plates setup consecutively. The effects upon IVCD, titre and specific productivity were measured after five days and, as before, each output response was normalised, as a percentage relative change, to internal control transfections containing no effector gene expression present within each transfection plate. Post-transfection viability was not assessed; therefore, it was not possible to infer whether decreased IVCD was due to reduced cell viability.

Overall, individual gene/protein and dose effects were observed for the majority of genes. Positive effects on titre were generally mediated by improvements in specific production levels with very marginal changes to growth apparent throughout, potentially still due to the relatively short culture duration, which included an initial period of post-transfection recovery. Small reductions in growth were deemed acceptable, these should not affect growth too considerably during longer, fed-batch production culture mode. The effects from the overexpression of the 28 genes, categorised into three different functional modules, are outlined in **Figure 7.10** (ER protein processing) **Figure 7.11** (Golgi complex processing/vesicular transport) and **Figure 7.12** (transcription factors). Statistical analysis using one-way ANOVA with a multiple comparison Dunnett's test, calculated using GraphPad Prism, compared all the genetic effectors at all levels transfected to the control transfection containing no engineering strategy. The results/significance for all comparisons made in the Dunnett's test are shown in the three figures and in **Appendix Table C.7**.

Effects of ER Protein Processing Components (Figure 7.10):

Overexpression of effector components categorised as associated with ER protein processing had very little impact upon culture growth. Overexpression of Canx (calnexin – an ER lectin chaperone) and, to a lesser extent, CRELD2 (a protein with a diverse range of functions including roles in protein folding, interacting with other chaperones, and the ER stress response) created small improvements to cell growth, with this effect present across all gene loads of CANX tested. There was a substantial detrimental effect on growth (~20% reduction) from the overexpression of FKBP11 (a PPIase enzyme which catalyses folding of proline-containing proteins). The upregulation of FKBP11 may have made processing taking place within the ER unequal, affecting normal cellular functioning and proliferation.

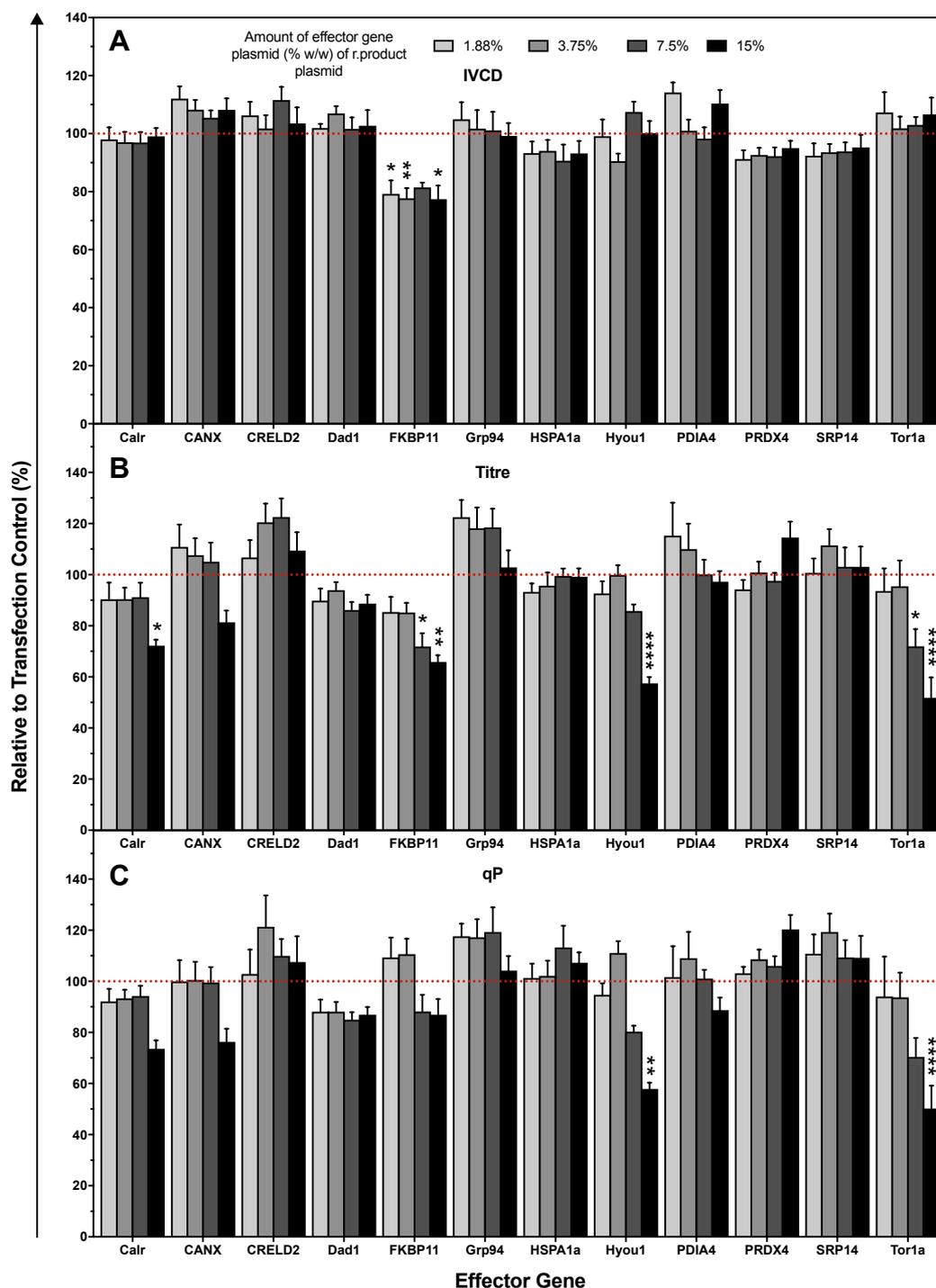


Figure 7.10: Responses from Co-Transfection of ER Protein Processing Effector Genes at Different Doses

Twelve ER protein processing related effector genes were co-transfected at four different levels: 15, 7.5, 3.75 and 1.88% (w/w) of DTE-mAb-OriP plasmid (533 ng); empty-‘Expression-Plasmid’ equalised DNA load across transfections. Culture IVCD (A) titre (B) and qP (C) were calculated after five days. Data is shown as a percentage change relative to control transfections (red-dashed line) containing no effector genes. The data represent the average and SEM responses from three biological replicates, each with three technical replicates. Statistically significant differences from control transfections are shown by asterisks (**** P<0.0001, ***P<0.001, **P<0.01, *P<0.05).

Larger changes to specific productivity were seen and this was the primary mode which increased volumetric titre. The ER chaperone Grp94 (functioning in relation to folding and stabilising of proteins within the ER) improved specific productivity by approximately 20%, with lower gene doses more beneficial (7.5-1.88%). Overexpression of higher amounts (15%) of PRDX4 (enzyme associated with re-oxidising PDI in the ER) also improved specific productivity levels by ~20%, but a slight penalty on IVCD was also seen. The overexpression of CRELD2 along with improving growth very slightly, improved specific productivity (particularly from transfected loads of 3.75-7.5%). This resulted in increases in titre of up to 22.5%. For several genes, there was a negative correlation to productivity resulting from an increased overexpression level. Increased overexpression of FKBP11, Hyou1 (ER chaperone), PDIA4 (catalyses disulphide bonds) and Tor1a (ER chaperone) appeared to disrupt the balance of processing and secretory pathways of proteins through the ER to a greater degree, and reduced the secretory output as a result.

Effects of Golgi Complex Processing and Vesicular Transport Components

(Figure 7.11):

Screening the collection of novel engineering targets, with functions relating to Golgi complex and vesicular transport of proteins through the secretory pathway, demonstrated that overexpression of several genes resulted in reduced cell growth compared to that of the control culture. In particular, reductions were seen from the overexpression of a collection of SNARE/tether proteins across all gene doses tested (with average changes of Bet1: <20%, Bet3: <30%, P115: <15% and Stx5a: <12%). These proteins have roles in COPII vesicle transport, shuttling proteins away from the ER and through the Golgi complex. However, these genes were all found to have positive effects upon specific production of the DTE-mAb. Bet 1 (Golgi complex SNARE protein associated with vesicle docking and COPII vesicle fusion) increased qP by up 20%; Bet 3 (a tether component, functioning in association with COPII vesicles and the Golgi complex membrane) increased qP by up to 35%; P115 (another tether protein which has a role in initial COPII vesicle docking in the ER) increased qP by up to 22.6%; and Stx5a (another SNARE proteins with a similar functions to Bet1) increased qP across all gene loads to a similar degree, with an average increase of 17%.

The effector gene MCFD2, (which has a role in cargo removal and transport from the ER to the Golgi complex) improved IVCD, by more than 10%, across all gene levels tested, however, there was no trend between the gene doses tested. As there was no change to specific productivity, the increased cell growth mediated an increased titre, by 10-20%, on average. Generally, there were minimal beneficial/substantial improvements to the titres resulting from overexpression of Golgi complex/secretory components. Many genes boosted cell specific productivity, but also reduced IVCD, which resulted in unchanged titres overall. The most promising Golgi complex/secretory components to improve titre were MCFD2, Bet 3 and Stx5a.

Gene dose effects were apparent, illustrated by the reduced productivity seen in the increasingly darker bars in **Figure 7.11**. In many transfections, there were greater negative effects arising from increased effector gene load transfected/expressed. There appear to be negative, or no responses, from ERGIC-53/CERT overexpression. It should be re-noted that the additional expression of an effector gene was likely to increase the biosynthetic burden to cells, reducing their specific productivity, and therefore the titre of recombinant product generated, when compared to control transfections containing no effector gene expression (demonstrated with GFP expression in **Figure 7.4**). ERGIC-53/CERT may have not been functioning correctly, or not able to function sufficiently at the level transfected; the response seen may have been just due to the increased copy number of each gene being expressed, without further investigation this could not be established.

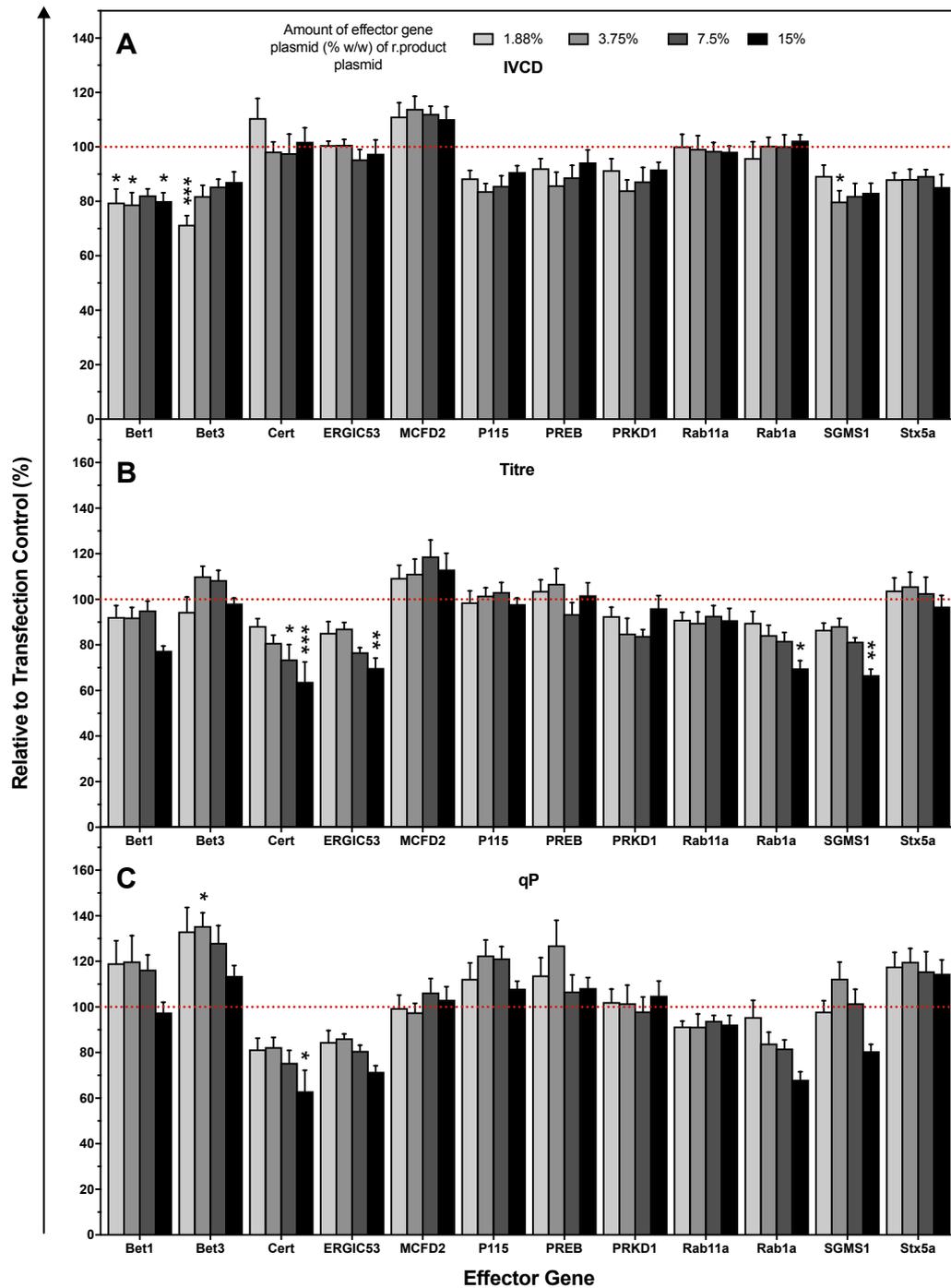


Figure 7.11: Responses from Co-Transfection of Golgi Complex Processing and Vesicular Transport Effector Genes at Different Doses

Twelve Golgi complex processing and vesicular transport related effector genes were co-transfected at four different levels: 15, 7.5, 3.75 and 1.88% (w/w) of DTE-mAb-OriP plasmid (533 ng); empty-‘Expression-Plasmid’ equalised DNA load across transfections. Culture IVCD (A) titre (B) and qP (C) were calculated after five days. Data is shown as a percentage change relative to control transfections (red-dashed line) containing no effector genes. The data represent the average and SEM responses from three biological replicates, each with three technical replicates. Statistically significant differences from control transfections are shown by asterisks (**** P<0.0001, ***P<0.001, **P<0.01, *P<0.05).

Effects of Transcription Factor Components (Figure 7.12):

A more promising approach to improve the production of the DTE-mAb was to target and overexpress transcription factors, with very significant improvements in production resulting from five days overexpression of CHO derived XBP1s (XBP1s (C)). Overexpression of transcription factor CREB3L2 (which has a role in regulating the secretory capacity of cells, including upregulating many secretory machinery components when activated) consistently increased cell growth across all gene doses (average increase in IVCD of 15%). There was little change to specific production levels, but the increased biomass mediated increased titres by an average of 16.8%, beyond the control transfections. Overexpression of transcription factor TFE3 (functions in the upregulation of many Golgi complex components when activated), caused adverse effects to cell growth, with an approximate 20% reduction in IVCD over five days, across the different loads tested. As viability was not assessed, it could not be ascertained if cell death had been induced as a result of TFE3 overexpression. With no substantial change to cellular specific productivity, apart from at higher levels, overall volumetric titres were lower for TFE3 overexpression than for control transfection samples. YY1 (multifunctional transcription factor), previously overexpressed and shown to have significant potential benefit to CHO cell production levels (Tastanova et al., 2016), did not introduce any major changes to cellular specific productivity or growth at the levels tested.

The overexpression of XBP1s (C) resulted in the largest improvement in production compared to the control transfections containing no engineering strategies, across the whole screen. An increase of over two-fold (up to 129%) in titre resulted from five-day cultures with XBP1s (C) overexpression. This increase was found to be mediated by increased cellular specific productivity levels, with qP increases of up to 158%. Advantageously, overexpression of XBP1s (C) did not substantially affect IVCD; there were small reductions, of up to 10%, during the five-day culture period. The benefit on qP substantially outweighed this small negative effect on IVCD.

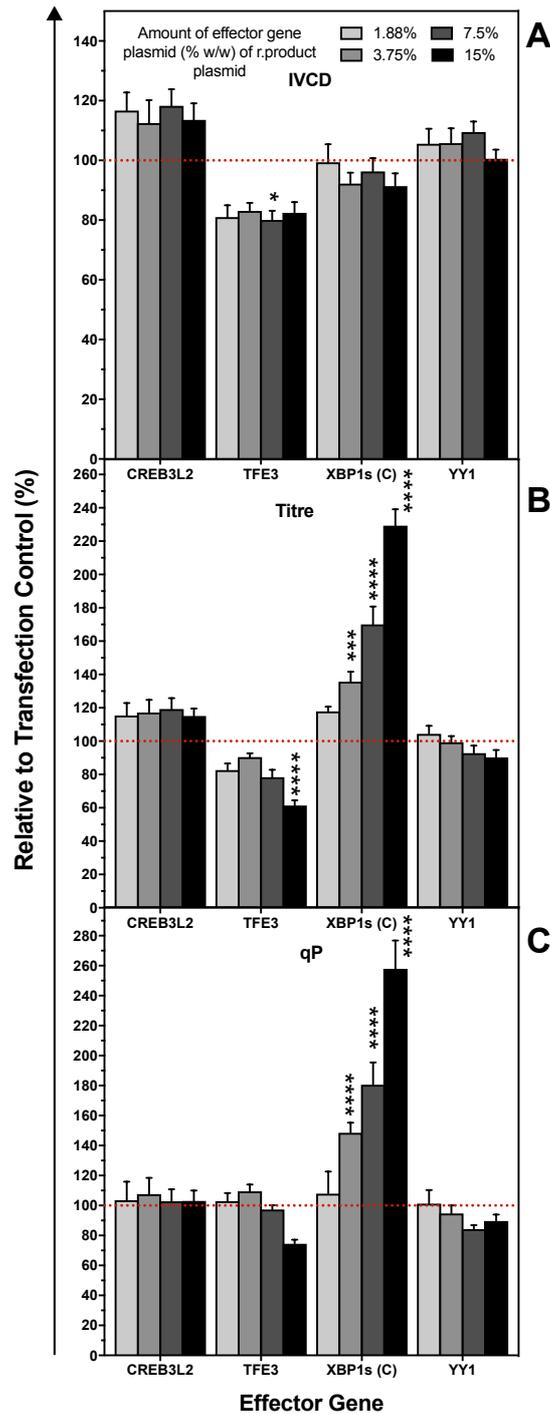


Figure 7.12 Responses from Co-Transfection of Transcription Factor Effector Genes at Different Doses

Four transcription factor effector genes were co-transfected at four different levels: 15, 7.5, 3.75 and 1.88% (w/w) of DTE-mAb-OriP plasmid (533 ng); empty-‘Expression-Plasmid’ equalised DNA load across transfections. Culture IVCD (A) titre (B) and qP (C) were calculated after five days. Data is shown as a percentage change relative to control transfections (red-dashed line) containing no effector genes. The data represent the average and SEM responses from three biological replicates, each with three technical replicates. Statistically significant differences from control transfections are shown by asterisks (**** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$).

A dose dependent effect on production was clear from the different levels of XBP1s (C) transfected. There was a very strong correlation ($R^2=0.9979$ and 0.9868 for titre and qP, respectively; shown in **Figure 7.13**) between the amount of effector gene dose transfected and both the specific productivity and the titre response percentage changes. Higher gene doses (3.75-15%) were found to have statistically significant improvements in production responses, using one-way ANOVA with a multiple comparison Dunnett's test, calculated using GraphPad Prism. The clear linear relationship between the dose of XBP1s (C) plasmid transfected and the increased production response, indicates that even higher amounts of plasmid load are likely to further enhance the improvement. However, a saturation point, where further gene copies would be detrimental to the cell, is likely to occur at a critical gene dose.

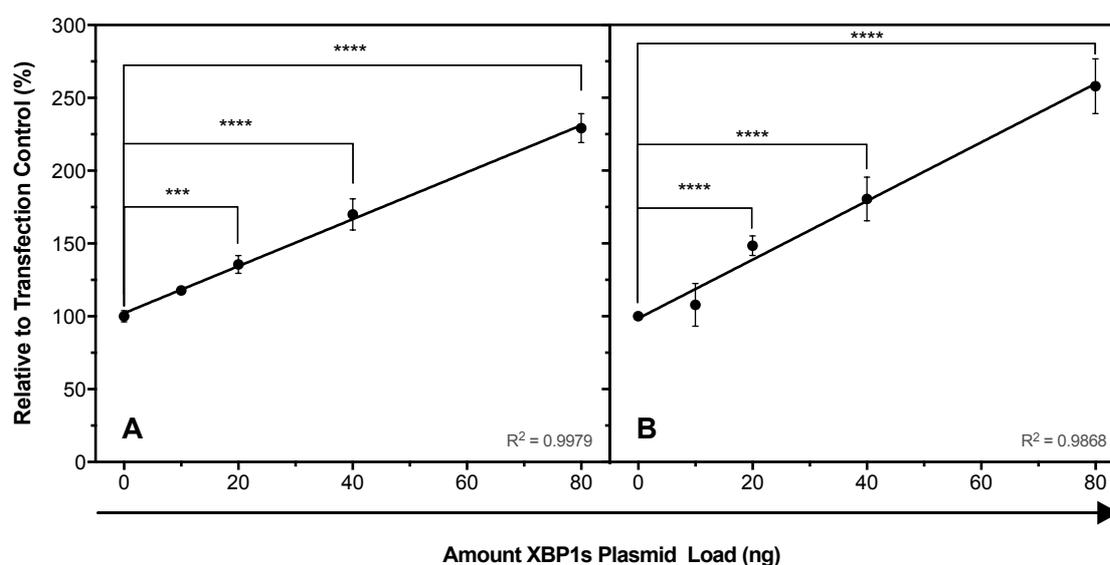


Figure 7.13: Clear Dose-Dependent Response from XBP1s (C) Overexpression

Four different plasmid doses of XBP1s (C) (80, 40, 20 and 10 ng, corresponding to 15, 7.5, 3.75 and 1.88% (w/w) of DTE-mAb-OriP plasmid (533 ng)) were transfected into CHO-T2 cells; empty-'Expression-Plasmid' equalised the DNA load for all transfections. The percentage change, relative to control transfections (100%) containing no effector genes, calculated after five days, for titre (**A**) and qP (**B**) are shown. The data represent the average and SEM responses from three biological replicates, each with three technical replicates. Statistically significant differences in production from control transfections are shown with asterisks (**** $P<0.0001$, *** $P<0.001$).

The *C. griseus* XBP1s protein was expressed in this second screening process, as opposed to the human version (XBP1s (H)) expressed in the static screens described earlier. As the XBP1s transcription factor interacts with CHO

components within the cell to regulate a wide collection of ER components, a native CHO version of the protein was thought to be preferable for its functionality and its potential to enhance the effect introduced from its overexpression. The amino acid sequences for the two XBP1s proteins are compared in **Appendix Figure C.3**, they share 88.9% amino acid sequence similarity. Overexpression of the two XBP1s sequences/proteins was compared for one gene dose, 7.5% (**Figure 7.14**). In this example, there were very similar effects from overexpression of the two different proteins. This perhaps emphasises the successful capabilities of this protein to improve expression of the DTE-mAb used here. It also suggests that the species source does not always influence the effect a protein has within the cell, perhaps related to the conservation of critical functional regions of the protein. This is converse to that previously seen from human/CHO YY1 overexpression in CHO cells (Tastanova et al., 2016). Both human (Becker et al., 2008; Cain et al., 2013; Pybus et al., 2013) and mouse (Hansen et al., 2015; Ku et al., 2007) XBP1s proteins have previously been shown to have beneficial effects in CHO cells. Here, it is shown that the CHO XBP1s protein is a similarly effective target when overexpressed.

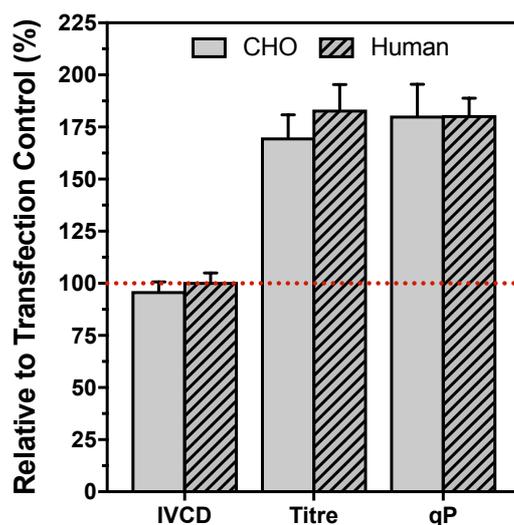


Figure 7.14: Overexpression of CHO and Human XBP1s Demonstrated Very Similar Responses

Overexpression of human and CHO XBP1s proteins at 7.5% (w/w) of DTE-mAb-OriP plasmid loading (533 ng) revealed very similar effects to IVCD, titre and qP after five days. Data is shown as a percentage change relative to control transfections (red-dashed line) containing no effector genes. The data represent the average and SEM responses from three biological replicates, each with three technical replicates.

7.4. Discussion

Within this chapter, a proof-of-concept of the developed high-throughput transient transfection platform has been demonstrated. Many simultaneous co-transfections were undertaken to audition a library of effector accessory genes for their functional effect on the expression of a model DTE-mAb.

Firstly, a library of genes was developed based upon information derived from a variety of sources. The sequences were synthesised and cloned into the same backbone plasmid for standardised expression. Seven genes were initially expressed in a three-day static screening setup, with different levels of genes comparatively expressed simply by altering the plasmid copy numbers transfected. Two combinatorial screens followed, which demonstrated screening approaches using DoE methodologies. The DoE approach aimed to identify significant impacting factors and optimise any interactions between effector genes. A larger screen with the remainder of the effector gene library followed. The implementation of a shaking, deep-well plate culture format was beneficial, when testing for component functionality. This is exemplified by overexpression XBP1s (H), with greater and less varied responses gained after five days of

shaking culture compared to three days of static culture which boosted qP response changes from 6% up to 81%. As the shaking culture mode enabled longer culture durations, this lengthened the production period and improved the VCDs, titres and functional effector capabilities. Population viabilities improved from 88% to 95% between days three and five and cells had more time to recover from the electroporation process. The use of a culture setup of this nature is recommended for any future platform implementation. The remainder of the seven initial gene targets were not overexpressed within the shaking, deep-well culture setup. The elevated effect from XBP1s (H) overexpression demonstrates that the initial static screen should be repeated to further identify their effect upon DTE-mAb expression.

The screening results demonstrate and emphasise the importance and need to test different copy numbers of effector genes. Either through transfecting the plasmid at a variety of loads, as used here, or by adjusting expression using different strengths of synthetic promoters. This idea has also recently been reported as an important concept to consider in an up to date review (Hansen et al., 2017). Trends from altering the dose of the gene transfected were observed for many genes. The normal balance of the ER, for example, was clearly disturbed to differing degrees as a result of different levels of effector gene overexpression.

7.4.1. Model DTE-mAb

The examples outlined within this chapter used a DTE-mAb provided by MedImmune as the model recombinant protein to primarily demonstrate the use of the platform as a screening tool, but to also try to identify factors to improve its expression. By altering the copy numbers of plasmid transfected, different levels of expression were compared to each other, and to those of a model ETE-mAb. Both products demonstrated very similar trends, of small, linear reductions in growth rate and viability correlating with a linear increase in specific production. The overall qP for the DTE-mAb was significantly lower than for the ETE-mAb, indicating the DTE-mAb was produced at a significantly slower rate. There were linear responses, from transfecting more plasmid copies, seen for both products. No saturation level was reached, unlike that of the production of a DTE fusion

protein, where after a certain plasmid loading, more gene copies no longer increased qP, and instead introduced a detrimental effect on growth rate (Johari et al., 2015). This trend indicated that it was unlikely that expression of this DTE-mAb was initiating cellular UPR/ER stress pathways. If there was a significant ER bottleneck, these responses would likely be initiated, and a greater reduction in viability/ μ would have resulted.

Many engineering strategies (genes and doses) were rapidly auditioned using the developed transient transfection platform. For the DTE-mAb tested here, the experiments consistently suggested that overexpression of XBP1s can significantly improve the product's expression and yield. Plasmid loads of greater than 15% (w/w) of recombinant mAb plasmid load are recommended. The use of the highest loads tested here (15%) resulted in titre improvements of 129%, relative to when no engineering strategy was used. A one gene engineering strategy is preferential, it is the most straight-forward and simplest to implement. As no interactions with XBP1s overexpression with other genes in the deep-well system were investigated at this stage, it is hard to indicate if a multiple strategy would be of significantly added benefit for this DTE-mAb. This is something that needs to be investigated further.

As a consequence of the effector gene screens and the study into altering the number of mAb plasmid copies transfected, it is suggested the lower yields for the DTE-mAb may incorporate a limitation related to an upstream cellular process such as transcription, translation, in addition or oppose to an ER folding and assembly or secretion; there is a slower production rate of the product compared to an ETE-mAb. The heavy and light chain gene ratio may have been sub-optimal (here equal copies were transfected, though interference from the use of a multigene plasmid may have diverged this ratio). The ratio should be investigated and optimised if used for subsequent product tests. (e.g Pybus et al., 2013). The overexpression of the XBP1s gene did relieve the production problems to some degree, significantly more than any other component screened. XBP1s is a major part of the UPR, activated naturally in response to ER stress (**Section 2.4.2**), it binds to regulatory sequences of ER components to increase their transcription in the nucleus (Lee et al., 2003). It plays an essential role in the natural immune system in the differentiation of B cells into antibody-producing plasma cells

(Iwakoshi et al., 2003; Shaffer et al., 2004). XBP1s expression has been shown to physically increase the size of the ER compartment; it upregulates a wide assortment of genes (e.g. many heat shock proteins: p58^{IPK}, ERdj4, HEDJ and PDIs: PDI5) related to secretory pathways; increases protein synthesis levels, numbers of ribosomes, and efficiency of translocation into the ER; and lowers levels of protein and polypeptide degradation (Glimcher, 2010; Gulis et al., 2014; Shaffer et al., 2004; Tigges and Fussenegger, 2006). It would be important to confirm that overexpression of XBP1s within this system is resulting in the anticipated intracellular effects and its mechanism of operation is as expected. The many beneficial attributes of XBP1s widely improve protein synthesis, processing and secretion. For this DTE-mAb, they appear to improve the cell's capability and speed to synthesise and process the nascent polypeptides, leading to an increased and less limited production rate.

7.4.2. Overexpression of XBP1s

Overexpression of transcription factors, like XBP1s, demonstrate the benefits of engineering a wide number of proteins. Transcription factors can alter the expression of many genes simultaneously, in a concerted and relative manner; a functionally relevant balance of up/down-regulation is probable. Transcription factors are likely to be generic effectors, if expressed at the right level, improving cell factory capabilities generally, unless there is a major bottleneck elsewhere that remains (Delic et al., 2014; Hansen et al., 2017). Using the evidence here, and the many previous successful examples, XBP1s should be tested for its effect in all future effector gene screens.

Overexpression of XBP1s demonstrates that beneficial solutions could be rapidly identified using the platform. Further optimisation, including increasing the plasmid load transfected, are likely to enhance the beneficial effect and should be investigated. Overexpression of XBP1s correlated negatively with a slight reduction in cell growth (IVCD). This is unfavourable for longer duration cultures, where maximum biomass accumulation is desirable to increase the antibody-producing material. This correlation that an increased qP can strongly and negatively correlate with reduced IVCD has been demonstrated previously (Johari et al., 2015). For engineering strategies which improve qP, the balance

with cell growth should be considered. If growth is affected, it could be potentially balanced/adjusted by altering the effector plasmid loading/functional level expressed or by co-expressing a gene that would improve growth simultaneously. For example, in the scenario presented here, MCFD2, CREB3L2 or Canx, which all improved IVCD slightly, without lowering qP, could be co-expressed alongside XBP1s and may relieve the negative effect of XBP1s on cell growth. Alternatively, an inducible system could be used to switch on XBP1s overexpression later on in culture once biomass has accumulated. For example, a Tet-On promoter system could be utilised, the promoter would only initiate transcription of the XBP1s gene in the presence of doxycycline/tetracycline, which could be added to cultures at the beginning of the stationary production period (Gulis et al., 2014; Misaghi et al., 2014).

7.4.3. Effector Gene Library

The rational approaches to generate the effector gene library used logical reasoning to consolidate the many potential gene targets that could have been tested. Based upon the genes that showed the most functional effects in the screens here (positive effectors: Bet1, Bet3, Canx, CREB3L2, CRELD2, Grp94, MCFD2, Stx5a and XBP1s and negative effectors: FKBP11, Hyou1, TFE3 and Tor1a), the plasma cell transcriptomic data provided the most beneficial information. Eight out of the nine genes causing positive effects were upregulated in plasma cells compared to B cells. The use of CHO omic datasets, which highlighted genes associated with (higher) producing CHO cells were also of benefit. The association with a producing-phenotype did not always introduce an advantageous effect. For example, overexpression of Hyou1 and FKBP11, both upregulated in plasma cells, had detrimental effects upon production of the DTE-mAb across the levels tested. Association with an antibody-producing phenotype appeared to be a potentially better source of information than the use of previous overexpression data and/or function alone. It must be emphasised that the effects seen were likely to be product- and potentially platform-specific (Johari et al., 2015; Pybus et al., 2013). Different effects may be measured in a different context, for example for a different level of recombinant product expression, cell line, duration of culture or in a SGE system rather than TGE system. The different

effector genes are likely to have a different effect upon the production of a different recombinant proteins having different cellular limitations.

Many of the effector genes within the library were likely to have positive interactions with each other (some were picked with this consideration) and so would have enhanced functionality as a consequence. Suggested beneficial combinations include: MCFD2 and ERGIC-53 (which function in a 1:1 complex with each other, with a role in protein export from the ER), Calr and Canx (which function together within the ER glycosylation pathway), XBP1s and CREB3L2 (which could boost qP and IVCD simultaneously by being different functional mediators) and the SNARE/tether proteins (Bet1, Bet3, P115 and Stx5a), which are likely to function together, but as they reduce cell growth this would need to be considered with their overexpression. A DoE approach was demonstrated within the chapter, outlining and recommending how combinatorial effectors could be screened and investigated. The sub-optimal static culture setup was used for the proof-of-concept of the DoE methodology, but XBP1s was still demonstrated as a beneficially significant effector gene. Perhaps due to its significant effect compared to all other effectors, its impact dominated the combinatorial studies. There was some significant impact upon production from its co-expression with CypB (for titres from the factorial design and qP from the RSM design). Higher loadings of XBP1s were suggested to be optimal, and should be tested in combination with a lower loading of CypB in an attempt to identify further improvement. The use of DoE approaches to investigate combinatorial effects was straight-forward: they reduced the number of runs (transfections) required; the interface was user-friendly; they provided statistical model-based analysis of variation, significant impactors and interactions, and can provide theoretical optimal combinations to test further.

7.5. Summary

The processes, screens and outcomes from this chapter begin to demonstrate the application and benefit of implementing a high-throughput transient transfection approach to try and identify ways of improving recombinant protein production and cell factory performance. A library of effector components was established using evidence from genetic association with high-producing

phenotypes. Functions within the ER and Golgi complex, related to protein folding, assembly, processing and secretion of the cell, were primary targets. The components were rapidly over-expressed, in a short duration screen and in-parallel, to test their functionality with a model DTE-mAb. The level of effector gene transfected was shown to have an effect upon its response; gene dosage should be considered when overexpressing genetic components. A DoE methodology can be implemented to investigate for and optimise the stoichiometry of significant interactions derived from co-expressing multiple effectors simultaneously. Here, a one-gene solution was found to be most optimal for the DTE-mAb. The sequences encoding XBP1s (both human and CHO), were consistently shown to alter the production of the DTE-mAb most significantly and effectively compared to the rest of the gene library, with increases in qP of up to 158% after five days of culture. Further improvements may be derived by higher levels of expression of XBP1s and co-transfection of XBP1s with other components, such as CypB or CREB3L2 to mediate in-parallel improvements to cell growth. The potential for the platform to be used for other relevant functionalities and future screening methodologies is discussed in the next chapter, along with some clear suggestions and thoughts on future perspectives and experiments.

8. Conclusions and Further Work

8.1. Summary Discussion and Conclusions

In the introductory chapters of this thesis a review into the literature described biopharmaceutical production in detail and focussed on cell engineering to improve the production capabilities of the CHO cell host. There were many promising examples demonstrating that cell engineering can be an effective strategy to improve recombinant protein titres, but there were many inconsistencies with mixed results between studies using diverse products, variable amounts of expression, and different host cells. The results in some of these studies may have been different, had differing experimental conditions or variables been employed.

In order to efficiently exploit cell engineering, a process to screen many of the targets of potential benefit in a standardised manner, to identify likely product specific solutions and to investigate effects of different levels of engineering (dose effects) was desirable. Additionally, with any complex interacting system, such that that taking place within cells to manufacture a recombinant protein, modulation/co-ordination of multiple components was likely to be advantageous. The manufacture of DTE proteins was of primary interest for a screen of this nature as they have increased requirement and potential for improvement in yield. However, identification of effectors that generally improve recombinant protein production are also of significance. The thesis advanced to describing the development of a high throughput platform, with the purpose of identification of engineering strategies to improve the CHO cell factory production of an industrially-relevant DTE-mAb.

To make rapid assessments of gene expression, a TGE system was developed as described in chapter five. A micro-scale electroporation setup with a high efficiency (over 94% of the population expressing the transfected gene) was developed. Plate-based electroporation based on Nucleofection technology

enabled hundreds of standardised transactions to be undertaken in-parallel. The micro-scale volumes involved meant lower levels of reagent and manual resource were required per transaction and many transfections could feasibly be set up simultaneously. When compared to an alternative chemical plate-based transfection approach, the electroporation process was superior for a short-duration culture setup, with plasmids introduced and expressed at higher immediate levels.

There was found to be a balance required between the DNA load and the number of cells transfected; adequate DNA levels were needed to ensure measurable titres, after a short-duration, of a DTE product expression and to provide sufficient capacity to transfect multiple plasmids at once. When the number of cells being transfected were sufficiently increased to account for low-volume related losses, post-transfection viability was much improved. The implementation of a deep-well plate shaking culture format also improved post-transfection viability. This culture format was beneficial for post-transaction recovery, increased culture duration, and improved cell growth; use of a culture format like this is recommended for future uses of the platform. The larger culture volumes from the use of deep-well cultures reduced variation in the platform setup, with only one culture well per transfection being required. Response outputs could be measured from the same culture, reducing the additional error otherwise introduced. As a micro-scale, high-throughput setup, variation is to be expected, particularly due to the low culture volumes being handled throughout. The number of replicates set up could have been increased, above the use of triplicates, to reduce variation further, but this would have reduced throughput, one of the main design criteria for the platform. From the results in chapter seven, the use of three technical and three biological replicates with each run normalised to an internal control, was sufficient for dose effects and significant changes to be identified within a screen. The primary output of chapter five was a micro-scale plate-based transient transfection platform which enabled the work in the next two chapters be developed.

Establishment of whether multiple genes could be transfected into cells at specific relative stoichiometry using the platform, to facilitate multigene engineering investigations in the future, was the focus in chapter six. The expression of

multiple genes simultaneously, and at specific relative stoichiometry, was demonstrated using the developed transfection platform. Fluorescent protein reporter expression enabled analysis of expression using flow cytometry, beneficial for its quick, high-throughput sampling and single cell analysis. The ratio and variation of expression of the three fluorescent protein genes across individual cells was compared to the ratio at which the genes were transfected. A limitation of the experimental setup, is that is very difficult, with a standard (basic) flow cytometer, with limited excitation lasers and emission filters, for independent measurements of multiple fluorescent proteins to be made simultaneously, due to their broad and overlapping fluorescence spectra. For this work, independent measurements could be made, but to enable this all proteins were measured at lower than optimal voltages to remove overlaps. This impacted the ability to ascertain transfection efficiency within these experiments and introduced false negative populations which then could not be assessed for their expression of the fluorescent protein genes. With more time and resource, the combination of fluorescent proteins may have been improved, meaning the required voltage reduction to reduce overlaps may have not been required.

Two approaches to deliver multiple genes to cells were compared: use of separate single gene plasmids; and a multigene plasmid containing all genes in-tandem. It was clear that further engineering, to introduce modulation within a multigene plasmid, was required to overcome the interference present between the genes on the plasmid. Despite an uneven expression ratio, comparisons between cells demonstrated transfection with a multigene plasmid was, as expected, less variable than that of single co-transfected plasmids. Variation was still relatively high (average CV, across all multigene gene expression, was 20.7%) from multigene expression. Variation in expression was likely in part to be due to the nature of the biological system, with cells being measured in different states within the cell division cycle and, genes having different transcriptional rates and proteins having different half-lives/turnovers within the cell. Some variation would have been introduced by the random process of plasmid uptake, and flow cytometry detection and analysis may have still been somewhat uneven between the three proteins, despite the improvement strategies explained in chapter six. These factors cannot easily be altered, being

inherent within the experimental setup used. Despite the heightened variation, the use of separate plasmids is still preferable. It is a logically easier and a more flexible approach to test different combinations of genes, with less plasmid cloning requirements and genes being tested easily interchanged by swapping the combination of plasmids transfected. When separate plasmids were co-transfected, the ratio of expression within individual cells adhered, on average, to the ratio to which the genes were transfected (e.g. when genes were transfected at a 1:1:1 ratio the expression ratio was 1.038:1.021:1.043). Approximately 80% of cells expressed the three genes with a ratio of 0.5:1:1.5. In spite of variation, a specific relative ratio of three genes was shown to be achievable within the electroporation setup by altering the co-transfection ratio of the separate plasmids carrying the genes.

Chapter seven describes implementation of the platform to rapidly express a variety of effector genes. There are increasing resources available (omic datasets) which can help in the decision making of what cellular targets to engineer. Omic data help to highlight genes that have an association with a (high) producing phenotype. However, there is still insufficient functional knowledge and understanding of many cellular processes, particularly within CHO cells, for the effects of target genes to be predictable. Therefore, currently there is a requirement to screen genes for potential effect and benefit; with a platform of the nature described here being an advantageous tool for this application.

The transfection platform was used to rapidly screen several effector genes; with 28 genes tested simultaneously and each gene titrated into cells at four different concentrations. This gene titration demonstrated that testing a range of gene expression is important for assessing the response generated. The functional amount of each gene should be investigated in an initial singular gene screen to establish if a gene has a functional effect which is of interest, and, if so, what levels of gene expression should be used for further investigation. A DTE-mAb, shown to have poor production titres, was used as an example recombinant protein to test within the platform, to demonstrate the platform's use as a screening tool. A titration of recombinant-product plasmid load was included within the screening process. This, along with the effector gene screens suggested that expression of the DTE-protein had slower levels of protein

production but did not indicate a significant ER bottleneck nor induce a substantial ER stress response within cells. No further effects on growth or viability compared to that seen for an ETE-mAb were apparent when the copy number transfected was increased, and no ER-specific effector gene had a significant improvement upon titre. The overexpression of effector genes did appear to be perturbing the normal cellular growth and production processes, with some small changes apparent but these were not substantially improving production of this particular DTE-mAb.

XBP1s overexpression improved cell specific production of the DTE-mAb and resulted in significant improvements in volumetric titres and demonstrated that a solution using the platform could be identified. Based on previous literature, it is proposed that overexpression of the XBP1s transcription factor increased protein synthesis and the general capacity of cells to produce and process more product, thus demonstrating the advantage of multigene engineering. Though it would be difficult to implement such widespread effects by up-regulation of a lot of individual effector gene targets. Overexpression of single genes may well introduce improvements, particularly if known and substantial bottlenecks are present as demonstrated previously (Johari et al., 2015), but engineering multiple factors is likely to relieve more bottlenecks and therefore be of more benefit.

The screen of the gene library demonstrated a few novel components of potential benefit. Of particular interest were components of the secretory pathway – proteins of the SNARE and tether families. These novel proteins increased cellular specific productivity and are thought likely to have additive, or even synergistic, effects when combined due to their related functionality. Cell growth was reduced as a result of their overexpression, this would need consideration – potentially balanced by using an in-parallel method to improve cell growth.

8.2. Alternative Platform Uses

The platform setup facilitated straight-forward exchange of components within a particular transfection screen. There are many other functional uses for the developed platform, in addition to screening of synthetic effector genes for their effect upon product expression. Using the transfection platform, synthetic and novel parts can be rapidly assessed, for example for their function, differences or

reproducibility. Novel cell lines and recombinant products could be screened in a high-throughput manner to begin to ascertain their manufacturing potential. Three examples follow, demonstrating alternative uses for the developed platform that have already been undertaken.

There is general development and progression in the biopharmaceutical industry towards having well-characterised, controllable biopharmaceutical expression through developing 'next-generation' synthetic and predictable cellular factories (Brown and James, 2016). To work towards achieving this concept, 'toolboxes' or libraries of characterised synthetic parts are required, but will need to be designed, synthesised and evaluated first. The described high-throughput transfection system can be utilised within this development path as an approach to rapidly and simultaneously test the effect of numerous, synthetic genetic parts. Within the DCJ laboratory (University of Sheffield) there is ongoing development of synthetic promoters, useful for facilitating and altering precise protein transcription, important to achieve multiple-gene stoichiometry.

Many in-parallel transfections of different synthetic promoter variants, prepared for multigene plasmid development, have been compared and validated for their relative promoter strength and behaviour with different proteins (work undertaken by Yash Patel). In **Figure 8.1A**, is an example using fluorescent protein expression (here expression TagBFP) to infer promoter strength. There is scope for hundreds of synthetic parts to be tested rapidly, and in a standardised manner, for their function in the CHO cell factory to be quickly ascertained.

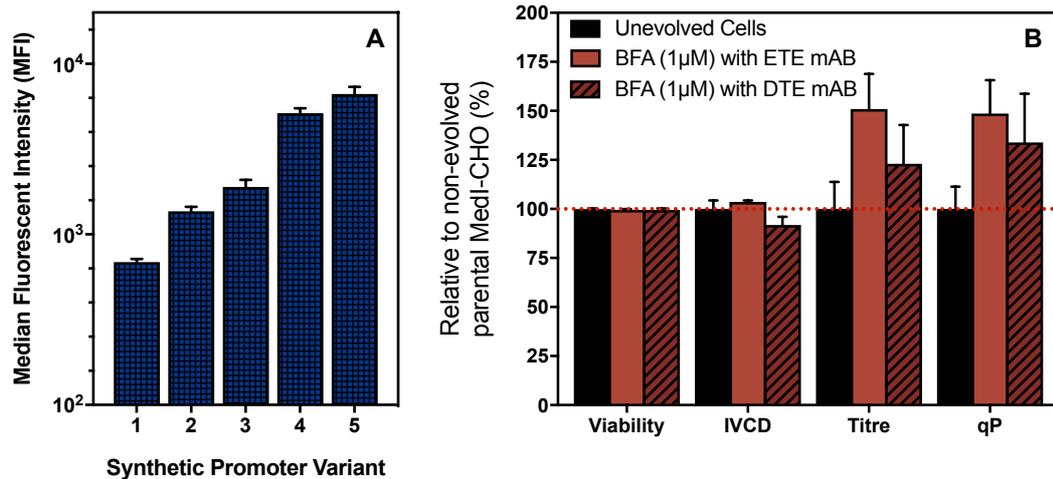


Figure 8.1: The Platform Can Be Used to Investigate the Functionality of Synthetic Parts and the Manufacturing Performance of Novel Cell Lines

(A) Five different synthetic promoter variants, of differing strengths, were tested for their functionality, indicated by TagBFP reporter expression measured at 24 hours. (B) A MEDI-CHO cell line, previously evolved to grow in brefeldin A (BFA), was transiently transfected separately with ETE-mAb and DTE-mAb plasmids. Culture viability, IVCD, titre and qP are shown relative to a non-evolved parental line and were measured after 72 hours. All data shown is the average and SEM from three biological and three technical replicates. The BFA cell line was provided by MedImmune/Nicholas Barber. Yash Patel provided the data for the synthetic promoters.

Evolution studies, undertaken in a collaboration between MedImmune and Nicholas Barber, have resulted in a MEDI-CHO derived cell line evolved to withstand and grow in the presence of 1 μ M brefeldin A (BFA), a fungal-derived lactone antiviral compound. BFA has been shown to trigger ER stress through inhibiting protein transport between the ER and Golgi complex, by preventing COP-I coat proteins binding to Golgi complex membranes (Dinter and Berger, 1998; Helms and Rothman, 1992). If cells were adapted to overcome this inhibitory effect from BFA, it is hypothesised they would have an increased secretory capacity and production capability. The transient transfection platform was implemented to initially test this hypothesis by investigating the novel cell line's ability to transiently express recombinant proteins. The early proof-of-concept study demonstrated the cell line had an enhanced ability to produce recombinant products, particularly for an ETE-mAb, when compared to a non-evolved parental MEDI-CHO host (shown in **Figure 8.1B**). This finding initiated larger follow-on experimentation including longer-term SGE studies. The

example demonstrates the use of the platform use to screen for manufacturing performance of novel cell lines.

Finally, the platform can be used to test and compare the CHO cell factory's ability to produce different recombinant products and sequence variants, or even unusual products and their derivatives; through straight-forwardly exchanging the plasmid being transfected. As part of a proof-of-concept study, undertaken with the University of Sheffield Molecular Biology and Biotechnology department and Dr Adam Brown within the DCJ laboratory, the platform was used to test the CHO cell's ability to produce a bacterial lysozyme, found to be troublesome to manufacture in bacteria, called Cellosyl. The protein sequence was engineered to prevent unwanted glycosylation, then tested using the developed transient platform to assess the ability of CHO cells to make and secrete the product. The subsequent knowledge that the product could be successfully produced and secreted in CHO cells, enabled larger scale and downstream experimentation to be initiated. The platform has the potential for rapid and simultaneous testing of potentially hundreds of products (including unusual products) and their derivatives, for their ability to be expressed with the CHO cell factory. Material and information can be rapidly generated for successive experimentation to be designed and undertaken.

8.3. Future Work

For the DTE-mAb used here, there are several recommendations to attempt to further improve its expression. The first seven effector genes, from the library developed within the work here, should be investigated in the longer duration, deep-well plate shaking culture format to further investigate if they have any functional benefit, like that found for XBP1s. The linear trend in improvement from XBP1s co-overexpression indicated that increasing the gene dose would likely be of further benefit. Effects later on in culture need to be investigated to ensure the small reduction in growth from XBP1s overexpression does not substantially slow down growth or the maximum IVCD reached. Co-overexpression of CREB3L2 with XBP1s, may enhance cell growth and improve production capability further. The impact of XBP1s overexpression on the ETE-mAb should be investigated, to help indicate if the solution is product-specific or a potentially

generic solution, for this cell line or for all products produced within this platform. As the problem with the DTE-mAb did not appear to be a major folding and assembly limitation, expression may be improved through testing different sequence variants, different signal peptides or alterations in the ratio of HC to LC transfected. All may speed up the otherwise slow production of the protein. If nothing improves titre a “diagnostic transient assay” could be used, like that used by Johari et al., 2015, which found UPR activation was induced during expression of a DTE fusion protein. This tool aims to diagnose where a gene/protein’s limitations are within the cell. It is not a high-throughput approach, so not suitable for widespread use, but could be useful to speed up finding solutions for particular problematic proteins.

An important and immediate piece of future work should centre around further testing the capability of the platform to screen functionality of multiple genetic engineering strategies to find optimal multigene engineering solutions. To achieve this, alternative recombinant molecules should be incorporated, preferably other industrial proteins such as the fusion protein used by Johari et al., 2015, which was found to produce very poor titres. For any new molecule screened, a gene titration should be undertaken to help identify potential cellular limitations, to see how it affects cell growth and viability as well as titre. The primary aim of testing different effector genes is to identify a set, or sets, of beneficial effector genes. These would then allow the DoE combinatorial approach to be examined fully as a suitable strategy to identify and optimise multigene combinations and their relative stoichiometry; something that was difficult to confirm when a single gene was having the significant effect.

For any optimal combination of effector components, the solution should be further tested including over a longer duration and in a scaled-up mode. An important criterion is to try and establish how well effects compare and how predictable the short-term transient production screen is. This could initially just cover longer-term TGE studies to investigate effects during late-phase culture. Long-term TGE could be useful to aid in the manufacture of sufficient product to use as part of pre-clinical studies. As clonal SGE is the industry and regulatory standard system for recombinant protein manufacture, it is important to investigate this. The optimal way of introducing multiple product and effector

genes simultaneously to ensure they are all integrated into the genome in parallel, with the need for only one selectable marker gene, and therefore expressed together, is to use a singular multigene construct with all components present in tandem, like those described by Guye et al., 2013 and Kriz et al., 2010 and used in chapter six. SGE of multigene plasmids would result in more efficient and predictable outcomes if combined with a targeted gene integration approach (e.g. Cre/Lox) to incorporate the genes into highly active region of heterochromatin. Precise relative stoichiometry of the components will be essential, the multigene construct requires further engineering to achieve this, and to overcome the current positional transcriptional interference (as seen in chapter six). In the literature the use of large insulator sequences, significantly increasing plasmid size, are used to overcome interference (Guye et al., 2013; Hasegawa and Nakatsuji, 2002; Yahata et al., 2007). Modularity may be improved with the use of shorter synthetic terminator sequences to prevent transcriptional interference (Curran et al., 2015; Gasanov et al., 2015). The optimal means of achieving stoichiometry would be to harness synthetic promoter technology (Brown et al., 2014; Brown and James, 2016). Specific promoters could titrate the relative expression level of each gene, including to accommodate for any inherent positional differences; inducible synthetic promoters could provide control of gene expression during culture (Brown and James, 2016; Gulis et al., 2014; Misaghi et al., 2014); and promoters could be designed to reduce cell epigenetic silencing of an integrated multigene construct, improving cell line stability (Brown and James, 2016).

Different functional modules of cellular processes, for example translational or glycosylation pathways, could be added into the screening effector gene library in the future. Product quality is an important feature of recombinant protein production that should be considered where possible. Compatible assays, in terms of throughput (96-well plate-based) and volume (micro-scale) to investigate product quality should be developed and incorporated to provide an indication of product quality within the screening process. Throughout the screening processes it is also important for a learning process to be built in, the knowledge from responses measured should be fed back into subsequent design and screening processes.

In the future, a high-throughput genetic screening platform could be beneficially implemented early-on in biopharmaceutical manufacturing pipelines. Instead of primarily screening for intrinsic clonal heterogeneity to search for an optimal phenotype, many other synthetic elements could also be screened to search for enhanced conditions that go beyond that possible from only screening within nature's boundaries. This type of screening process only becomes possible as the cost of DNA synthesis significantly reduces as it becomes more widely utilised. In 2005, DNA synthesis cost US \$1 per bp; this had reduced ten-fold by 2016, to just US \$0.10 per bp (Petroni, 2016). Technological development is further reducing costs, with as little as US \$0.03 per base now publicised, (Gen9, 2016). These low costs would feasibly enable hundreds of construct variants to be synthesised. A transfection screening platform, will be crucial to test their functionality *in vivo* and screen for optimal elements. In the future the development of a cell-free expression system could be incorporated, with functionality tested in systems containing lower complexity otherwise present in a whole cell system (Hodgman and Jewett, 2012; Thoring et al., 2016). This may introduce a significantly heightened high-throughput analysis of components.

Many synthetic variants could be designed *in silico* to tune different aspects of the host cell factory; these could include different synthetic sequence variants, promoters, effector genes, signal sequences or UTRs, and tested at different levels. Potentially the addition of functional chemical effectors and genetic knockdown could also be incorporated into the screening process. All have the potential to improve recombinant protein production and combined together significantly enhance the CHO cell factory. An important function of a screening process of this nature should be to learn from the responses measured using *in silico* modelling analysis, significant effects and interactions should be fed back into the design process to direct future synthesis and screening processes. Consequently, the speed and benefit of CHO cell engineering could dramatically improve. Ultimately product-specific "designer" cell factories would be possible and employed for specific recombinant protein production. There would no longer be "difficult-to-express" proteins, "designer" engineering solutions for these proteins would be readily achievable.

References

- Aebi M. 2013. N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta - Mol. Cell Res.* **1833**:2430–2437.
- Aggarwal S. 2011. What's fueling the biotech engine--2010 to 2011. *Nat. Biotechnol.* **29**:1083–9.
- Agrawal V, Yu B, Pagila R, Yang B, Simonsen C, Beske O. 2013. A High-Yielding, CHO-K1–Based Transient Transfection System. *Bioprocess Int.* **11**:28–35.
- Alder NN, Shen Y, Brodsky JL, Hendershot LM, Johnson AE. 2005. The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *J. Cell Biol.* **168**:389–399.
- Andersen DC, Krummen L. 2002. Recombinant protein expression for therapeutic applications. *Curr. Opin. Biotechnol.* **13**:117–23.
- Andrianantoandro E, Basu S, Karig DK, Weiss R. 2006. Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**:1–14.
- Appenzeller-Herzog C, Riemer J, Zito E, Chin K-T, Ron D, Spiess M, Ellgaard L. 2010. Disulphide production by Ero1 α -PDI relay is rapid and effectively regulated. *EMBO J.* **29**:3318–3329.
- Baldi L, Hacker DL, Adam M, Wurm FM. 2007. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnol. Lett.* **29**:677–84.
- Bandaranayake AD, Almo SC. 2014. Recent advances in mammalian protein production. *FEBS Lett.* **588**:253–260.
- Barnes LM, Bentley CM, Dickson AJ. 2000. Advances in animal cell recombinant protein production: GS-NS0 expression system. *Cytotechnology* **32**:109–23.
- Barron N, Kumar N, Sanchez N, Doolan P, Clarke C, Meleady P, O'Sullivan F, Clynes M. 2011a. Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. *J. Biotechnol.* **151**:204–211.
- Barron N, Sanchez N, Kelly P, Clynes M. 2011b. MicroRNAs: Tiny targets for engineering CHO cell phenotypes? *Biotechnol. Lett.* **33**:11–21.
- Beck A. 2011. Biosimilar, biobetter and next generation therapeutic antibodies. *MAbs* **3**:107–110.
- Becker E, Florin L, Pfizenmaier K, Kaufmann H. 2008. An XBP-1 dependent bottle-neck in production of IgG subtype antibodies in chemically defined serum-free Chinese hamster ovary (CHO) fed-batch processes. *J. Biotechnol.* **135**:217–223.
- Becker E, Florin L, Pfizenmaier K, Kaufmann H. 2010. Evaluation of a combinatorial cell engineering approach to overcome apoptotic effects in XBP-1(s) expressing cells. *J. Biotechnol.* **146**:198–206.
- Becker J, Hackl M, Rupp O, Jakobi T, Schneider J, Szczepanowski R, Bekel T, Borth N, Goesmann A, Grillari J, Kaltschmidt C, Noll T, Pühler A, Tauch A, Brinkroff K. 2011. Unraveling the Chinese hamster ovary cell line transcriptome by next-generation sequencing. *J. Biotechnol.* **156**:227–235.
- Benner SA, Sismour AM. 2005. Synthetic biology. *Nat. Rev. Genet.* **6**:533–543.
- Berting A, Farcet MR, Kreil TR. 2010. Virus susceptibility of Chinese hamster ovary (CHO) cells and detection of viral contaminations by adventitious agent testing. *Biotechnol. Bioeng.* **106**:598–607.
- Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat. Cell Biol.* **2**:326–332.
- Birch JR, Racher AJ. 2006. Antibody production. *Adv. Drug Deliv. Rev.* **58**:671–685.
- Borth N, Mattanovich D, Kunert R, Katinger H. 2005. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. *Biotechnol. Prog.* **21**:106–11.
- Braakman I, Bulleid NJ. 2011. Protein folding and modification in the Mammalian endoplasmic reticulum. *Annu. Rev. Biochem.* **80**:71–99.

- Brinkrolf K, Rupp O, Laux H, Kollin F, Ernst W, Linke B, Kofler R, Romand S, Hesse F, Budach WE, Galosy S, Müller D, Noll T, Wienberg J, Jostock T, Leonard M, Grillari J, Tauch A, Goesmann A, Helk B, Mott JE, Pühler A, Borth N. 2013. Chinese hamster genome sequenced from sorted chromosomes. *Nat. Biotechnol.* **31**:694–5.
- Brown AJ, James DC. 2016. Precision control of recombinant gene transcription for CHO cell synthetic biology. *Biotechnol. Adv.* **34**:492–503.
- Brown AJ, Sweeney B, Mainwaring DO, James DC. 2014. Synthetic promoters for CHO cell engineering. *Biotechnol. Bioeng.* **111**:1638–47.
- Brown HC, Gangadharan B, Doering CB. 2011. Enhanced biosynthesis of coagulation factor VIII through diminished engagement of the unfolded protein response. *J. Biol. Chem.* **286**:24451–24457.
- Butler M, Meneses-Acosta A. 2012. Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Appl. Microbiol. Biotechnol.* **96**:885–94.
- Cain K, Peters S, Hailu H, Sweeney B, Stephens P, Heads J, Sarkar K, Ventom A, Page C, Dickson A. 2013. A CHO cell line engineered to express XBP1 and ERO1- α has increased levels of transient protein expression. *Biotechnol. Prog.* **29**:697–706.
- Calfon M, Zeng H, Urano F, Till JH, Hubbard SR, Harding HP, Clark SG, Ron D. 2002. IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA. *Nature* **415**:92–96.
- Camire J, Kim D, Kwon S. 2017. Enhanced production of recombinant proteins by a small molecule protein synthesis enhancer in combination with an antioxidant in recombinant Chinese hamster ovary cells. *Bioprocess Biosyst. Eng.* **40**:1049–1056.
- Campos-Da-Paz M, Costa CS, Quilici LS, Simoes IDC, Kyaw CM, Maranhao AQ, Brigido MM. 2008. Production of recombinant human factor VIII in different cell lines and the effect of human XBP1 co-expression. *Mol. Biotechnol.* **39**:155–158.
- Carlage T, Kshirsagar R, Zang L, Janakiraman V, Hincapie M, Lyubarskaya Y, Weiskopf A, Hancock WS. 2012. Analysis of dynamic changes in the proteome of a Bcl-XL overexpressing Chinese hamster ovary cell culture during exponential and stationary phases. *Biotechnol. Prog.* **28**:814–823.
- Chakrabarti A, Chen AW, Varner JD. 2011. A review of the mammalian unfolded protein response. *Biotechnol. Bioeng.* **108**:2777–93.
- Chiang GG, Sisk WP. 2005. Bcl-xL mediates increased production of humanized monoclonal antibodies in chinese hamster ovary cells. *Biotechnol. Bioeng.* **91**:779–792.
- Chung J, Lim S, Hong Y, Hwang S, Lee G. 2004. Effect of Doxycycline-Regulated Calnexin and Calreticulin Expression on Specific Thrombopoietin Productivity of Recombinant Chinese Hamster Ovary Cells. *Biotechnol. Bioeng.* **85**:839–546.
- Coco-Martin JM, Harmsen MM. 2008. A review of therapeutic protein expression by mammalian cells. *Bioprocess Int.* **6**:28–33.
- Codamo J, Hou JJC, Hughes BS, Gray PP, Munro TP. 2011. Efficient mAb production in CHO cells incorporating PEI-mediated transfection, mild hypothermia and the co-expression of XBP-1. *J. Chem. Technol. Biotechnol.* **86**:923–934.
- Cohen RN, van der Aa M, Macaraeg N, Lee AP, Szoka FC. 2009. Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *J. Control. Release* **135**:166–174.
- Cortez L, Sim V. 2014. The therapeutic potential of chemical chaperones in protein folding diseases. *Prion* **8**:197–202.
- Cost GJ, Freyvert Y, Vafiadis A, Santiago Y, Miller JC, Rebar E, Collingwood TN, Snowden A, Gregory PD. 2010. BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol. Bioeng.* **105**:330–340.
- Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY. 1995. Understanding, improving and using green fluorescent proteins. *Trends Biochem. Sci.* **20**:448–55.
- Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. 2015. Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. *ACS Synth. Biol.* **4**:824–

832.

- Daramola O, Stevenson J, Dean G, Hatton D, Pettman G, Holmes W, Field R. 2014. A high-yielding CHO transient system: Coexpression of genes encoding EBNA-1 and GS enhances transient protein expression. *Biotechnol. Prog.* **30**:132–141.
- Datta P, Linhardt RJ, Sharfstein ST. 2013. An 'omics approach towards CHO cell engineering. *Biotechnol. Bioeng.* **110**:1255–1271.
- Davies SL, O'Callaghan PM, Mcleod J, Pybus LP, Sung YH, Rance J, Wilkinson SJ, Racher AJ, Young RJ, James DC. 2011. Impact of gene vector design on the control of recombinant monoclonal antibody production by chinese hamster ovary cells. *Biotechnol. Prog.* **27**:1689–1699.
- Davies SL, Lovelady CS, Grainger RK, Racher AJ, Young RJ, James DC. 2013. Functional heterogeneity and heritability in CHO cell populations. *Biotechnol. Bioeng.* **110**:260–74.
- Davis R, Schooley K, Rasmussen B, Thomas J, Reddy P. 2000. Effect of PDI overexpression on recombinant protein secretion in CHO cells. *Biotechnol. Prog.* **16**:736–43.
- Delic M, Göngrich R, Diethard M, Gasser B. 2014. Engineering of Protein Folding and Secretion—Strategies to Overcome Bottlenecks for Efficient Production of Recombinant Proteins. *Antioxid. Redox Signal.* **21**:414–437.
- Demain AL, Vaishnav P. 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27**:297–306.
- Dinnis DM, James DC. 2005. Engineering mammalian cell factories for improved recombinant monoclonal antibody production: lessons from nature? *Biotechnol. Bioeng.* **91**:180–9.
- Dinnis DM, Stansfield SH, Schlatter S, Smales CM, Alete D, Birch JR, Racher AJ, Marshall CT, Nielsen LJ, James DC. 2006. Functional Proteomic Analysis of GS-NS0 Murine Myeloma Cell Lines With Varying Recombinant Monoclonal Antibody Production Rate. *Biotechnol. Bioeng.* **94**:830–841.
- Dinter A, Berger EG. 1998. Golgi-disturbing agents. *Histochem. Cell Biol.* **109**:571–590.
- Doolan P, Melville M, Gammell P, Sinacore M, Meleady P, McCarthy K, Francullo L, Leonard M, Charlebois T, Clynes M. 2008. Transcriptional profiling of gene expression changes in a PACE-transfected CHO DUKX cell line secreting high levels of rhBMP-2. *Mol. Biotechnol.* **39**:187–199.
- Dorner A, Wasley L, Kaufman R. 1989. Increased synthesis of secreted proteins induces expression of glucose-regulated proteins in butyrate-treated Chinese hamster ovary cells. *J. Biol. Chem.* **264**:20602–20607.
- Douglas KL. 2008. Toward development of artificial viruses for gene therapy: A comparative evaluation of viral and non-viral transfection. *Biotechnol. Prog.* **24**:871–883.
- Dreesen I a J, Fussenegger M. 2011. Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of chinese hamster ovary cells. *Biotechnol. Bioeng.* **108**:853–66.
- Dumont J, Ewart D, Mei B, Estes S, Kshirsagar R. 2016. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol.* **36**:1110–1122.
- Durocher Y, Loignon M. 2011. Process, Vectors and Engineered Cell Lines for Enhanced Large-Scale Transfection. US Patent 2011/0039339 A1.
- Ellgaard L, Frickel E-M. 2003. Calnexin, calreticulin, and ERp57: teammates in glycoprotein folding. *Cell Biochem. Biophys.* **39**:223–247.
- Ellgaard L, Helenius A. 2003. Quality control in the endoplasmic reticulum. *Nat. Rev. cell Biol.* **4**:181–191.
- Ellgaard L, Ruddock LW. 2005. The human protein disulphide isomerase family: substrate interactions and functional properties. *EMBO Rep.* **6**:28–32.
- Ellis T, Adie T, Baldwin G. 2011. DNA assembly for synthetic biology: from parts to pathways and beyond. *Integr. Biol.* **3**:109–118.
- Estes B, Hsu YR, Tam LT, Sheng J, Stevens J, Haldankar R. 2015. Uncovering methods for the prevention of protein aggregation and improvement of product quality in a transient

- expression system. *Biotechnol. Prog.* **31**:258–267.
- Evrogen. 2008. Blue fluorescent protein TagBFP. *Blue Fluoresc. protein TagBFP Inf. Sheet*. <http://evrogen.com/protein-descriptions/TagBFP-description.pdf>.
- Fan L, Kadura I, Krebs LE, Hatfield CC, Shaw MM, Frye CC. 2012. Improving the efficiency of CHO cell line generation using glutamine synthetase gene knockout cells. *Biotechnol. Bioeng.* **109**:1007–15.
- Fann CH, Guarna MM, Kilburn DG, Piret JM. 1999. Relationship between recombinant activated protein C secretion rates and mRNA levels in baby hamster kidney cells. *Biotechnol. Bioeng.* **63**:464–472.
- Ferrer-Mirallès N, Domingo-Espín J, Corchero J, Vázquez E, Villaverde A. 2009. Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* **8**:17.
- Fischer S, Buck T, Wagner A, Ehrhart C, Giancaterino J, Mang S, Schad M, Mathias S, Aschrafi A, Handrick R, Otte K. 2014. A functional high-content miRNA screen identifies miR-30 family to boost recombinant protein production in CHO cells. *Biotechnol. J.* **9**:1279–1292.
- Fischer S, Handrick R, Otte K. 2015. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnol. Adv.* **33**:1878–1896.
- Florin L, Pegel A, Becker E, Hausser A, Olayioye MA, Kaufmann H. 2009. Heterologous expression of the lipid transfer protein CERT increases therapeutic protein productivity of mammalian cells. *J. Biotechnol.* **141**:84–90.
- Le Fourn V, Girod P-A, Buceta M, Regamey A, Mermoud N. 2014. CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. *Metab. Eng.* **21**:91–102.
- Fox RM, Hanlon CD, Andrew DJ. 2010. The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. *J. Cell Biol.* **191**:479–492.
- Fox SR, Tan HK, Tan MC, Wong SCNC, Yap MGS, Wang DIC. 2005. A detailed understanding of the enhanced hypothermic productivity of interferon- γ by Chinese-hamster ovary cells. *Biotechnol. Appl. Biochem.* **41**:255.
- Frand AR, Kaiser CA. 1998. The ERO1 Gene of Yeast Is Required for Oxidation of Protein Dithiols in the Endoplasmic Reticulum. *Mol. Cell* **1**:161–170.
- Fratantoni JC, Dzekunov S, Singh V, Liu LN. 2003. A non-viral gene delivery system designed for clinical use. *Cytotherapy* **5**:208–211.
- Fukuda M. 2008. Membrane traffic in the secretory pathway: Regulation of secretory vesicle traffic by Rab small GTPases. *Cell. Mol. Life Sci.* **65**:2801–2813.
- Gasanov NB, Toshchakov S V., Georgiev PG, Maksimenko OG. 2015. The use of transcription terminators to generate transgenic lines of chinese hamster ovary cells (CHO) with stable and high level of reporter gene expression. *Acta Naturae* **7**:74–80.
- Geisse S. 2009. Reflections on more than 10 years of TGE approaches. *Protein Expr. Purif.* **64**:99–107.
- Geisse S, Fux C. 2009. Recombinant Protein Production by Transient Gene Transfer into Mammalian Cells. *Methods Enzymol.* 1st ed. Elsevier Inc. Vol. 463 223–238 p.
- Gen9. 2016. Gen9 Announces Next Generation of the BioFab® DNA Synthesis Platform. *Sci.* <http://www.the-scientist.com/?articles.view/articleNo/45760/title/Gen9-Announces-Next-Generation-of-the-BioFab--DNA-Synthesis-Platform-/>.
- Gerngross TU. 2004. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat. Biotechnol.* **22**:1409–14.
- Glimcher LH. 2010. XBP1: the last two decades. *Ann. Rheum. Dis.* **69**:67–71.
- Grav LM, Lee JS, Gerling S, Kallehauge TB, Hansen AH, Kol S, Lee GM, Pedersen LE, Kildegaard HF. 2015. One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. J.* **10**:1446–1456.
- Gulis G, Simi KC, de Toledo R, Maranhao A, Brigido M. 2014. Optimization of heterologous protein production in Chinese hamster ovary cells under overexpression of spliced form of human X-box binding protein. *BMC Biotechnol.* **14**:26.
- Guye P, Li Y, Wroblewska L, Duportet X, Weiss R. 2013. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Res.* **41**:3–8.

- Hacker DL, Derow E, Wurm FM. 2005. The CELO adenovirus Gam1 protein enhances transient and stable recombinant protein expression in Chinese hamster ovary cells. *J. Biotechnol.* **117**:21–29.
- Hackl M, Borth N, Grillari J. 2012. MiRNAs - pathway engineering of CHO cell factories that avoids translational burdening. *Trends Biotechnol.* **30**:405–406.
- Hansen HG, Nilsson CN, Lund AM, Kol S, Grav LM, Lundqvist M, Rockberg J, Lee GM, Andersen MR, Kildegaard HF. 2015. Versatile microscale screening platform for improving recombinant protein productivity in Chinese hamster ovary cells. *Sci. Rep.* **5**:18016.
- Hansen HG, Pristovšek N, Kildegaard HF, Lee GM. 2017. Improving the secretory capacity of Chinese hamster ovary cells by ectopic expression of effector genes: Lessons learned and future directions. *Biotechnol. Adv.* **35**:64–76.
- Harding HP, Zhang Y, Ron D. 1999. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* **397**:271–274.
- Haredy AM, Nishizawa A, Honda K, Ohya T, Ohtake H, Omasa T. 2013. Improved antibody production in Chinese hamster ovary cells by ATF4 overexpression. *Cytotechnology* **65**:993–1002.
- Harreither E, Hackl M, Pichler J, Shridhar S, Auer N, Łabaj PP, Scheideler M, Karbiener M, Grillari J, Kreil DP, Borth N. 2015. Microarray profiling of preselected CHO host cell subclones identifies gene expression patterns associated with increased production capacity. *Biotechnol. J.* **10**:1625–1638.
- Hartley CL, Edwards S, Mullan L, Bell PA, Fresquet M, Boot-Handford RP, Briggs MD. 2013. Armet/Manf and Creld2 are components of a specialized ER stress response provoked by inappropriate formation of disulphide bonds: Implications for genetic skeletal diseases. *Hum. Mol. Genet.* **22**:5262–5275.
- Hasegawa K, Nakatsuji N. 2002. Insulators prevent transcriptional interference between two promoters in a double gene construct for transgenesis. *FEBS Lett.* **520**:47–52.
- Hawley TS, Hawley RG, Telford WG. 2017. Fluorescent Proteins for Flow Cytometry. *Curr. Protoc. Cytom.*:9.12.1-9.12.20.
- Hayes NVL, Smales CM, Klappa P. 2010. Protein disulfide isomerase does not control recombinant IgG4 productivity in mammalian cell lines. *Biotechnol. Bioeng.* **105**:770–779.
- Heim R, Cubitt AB, Tsien RY. 1995. Improved green fluorescence. *Nature* **373**:663–664.
- Helms JB, Rothman JE. 1992. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature* **360**:352–354.
- Hernández Bort J, Hackl M, Höflmayer H, Jadhav V, Harreither E, Kumar N, Ernst W, Grillari J, Borth N. 2011. Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. *Biotechnol. J.* **7**:500–515.
- Hetz C. 2012. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nat. Rev. Mol. Cell Biol.* **13**:89–102.
- Hodgman E, Jewett M. 2012. Cell-Free Synthetic Biology: Thinking Outside the Cell. *Metab Eng.* **14**:261–269.
- Hong WJ, Lev S. 2014. Tethering the assembly of SNARE complexes. *Trends Cell Biol.* **24**:35–43.
- Hu WS, Zhou W, Seth G, Ozturk S, Zhang C. 2012. Cell Culture Bioprocess Engineering 1st ed. Wei-Shou Hu 1-336 p.
- Huang Y-M, Hu W, Rustandi E, Chang K, Yusuf-Makagiansar H, Ryll T. 2010. Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol. Prog.* **26**:1400–10.
- Hussain H, Maldonado-Agurto R, Dickson AJ. 2014. The endoplasmic reticulum and unfolded protein response in the control of mammalian recombinant protein production. *Biotechnol. Lett.* **36**:1581–1593.
- Hwang SJ, Jeon CJ, Cho SM, Lee GM, Yoon SK. 2011. Effect of chemical chaperone addition on production and aggregation of recombinant flag-tagged COMP-angiopoietin 1 in Chinese hamster ovary cells. *Biotechnol. Prog.* **27**:587–591.

- Hwang SO, Chung JY, Lee GM. 2003. Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. *Biotechnol. Prog.* **19**:179–84.
- Invitrogen. 2011. Lipofectamine® LTX and PLUS™ Reagents. http://norstromlab.wikispaces.com/file/view/lipofectamine_LTX_and_PLUS_man.pdf.
- Invitrogen. 2013. Lipofectamine® LTX & PLUS™ Reagent Protocol. https://tools.lifetechnologies.com/content/sfs/manuals/LipofectamineLTX_PLUS_Reag_protocol.pdf.
- Ishaque A, Thrift J, Murphy JE, Konstantinov K. 2007. Over-Expression of Hsp70 in BHK-21 Cells Engineered to Produce Recombinant Factor VIII Promotes Resistance to Apoptosis and Enhances Secretion. *Biotechnol. Bioeng.* **97**:144–155.
- Iwakoshi NN, Lee A-H, Vallabhajosyula P, Otipoby KL, Rajewsky K, Glimcher LH. 2003. Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1. *Nat. Immunol.* **4**:321–329. <http://www.nature.com/doi/10.1038/ni907>.
- Jaffé SRP, Strutton B, Levarski Z, Pandhal J, Wright PC. 2014. Escherichia coli as a glycoprotein production host: Recent developments and challenges. *Curr. Opin. Biotechnol.* **30**:205–210.
- Jahn R, Scheller RH. 2006. SNAREs — engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.* **7**:631–643.
- Jayapal K, Wlaschin K, Hu W, Yap M. 2007. Recombinant protein therapeutics from CHO cells - 20 years and counting. *Chem. Eng. Prog.* **103**:40–47.
- Jemal A, Bray F, Centerm M, Ferlay J, Ward E, Forman D. 2011. Global Cancer Statistics: 2011. *CA Cancer J Clin* **61**:69–90.
- De Jesus M, Wurm FM. 2011. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *Eur. J. Pharm. Biopharm.* **78**:184–8.
- Johari YB, Estes SD, Alves CS, Sinacore MS, James DC. 2015. Integrated cell and process engineering for improved transient production of a “difficult-to-express” fusion protein by CHO cells. *Biotechnol. Bioeng.* **112**:2527–2542.
- Johnson IS. 1983. Human Insulin from Recombinant DNA Technology. *Science (80-)*. **219**:632–637.
- Johnston SL. 2007. Biologic therapies: what and when? *J. Clin. Pathol.* **60**:8–17.
- Jordan M, Wurm F. 2004. Transfection of adherent and suspended cells by calcium phosphate. *Methods* **33**:136–43.
- Jordan M, Wurm FM. 2003. Co-transfer multiple plasmids/viruses as an attractive method to introduce several genes in mammalian cells. *New Compr. Biochem.* **38**:337–348.
- Jossé L, Mark Smales C, Tuite MF. 2010. Transient expression of human TorsinA enhances secretion of two functionally distinct proteins in cultured chinese hamster ovary (CHO) cells. *Biotechnol. Bioeng.* **105**:556–566.
- Kadesch T, Berg P. 1986. Effects of the position of the simian virus 40 enhancer on expression of multiple transcription units in a single plasmid. *Mol. Cell. Biol.* **6**:2593–601.
- Kelley B. 2009. Industrialization of mAb production technology: the bioprocessing industry at a crossroads. *MAbs* **1**:443–52.
- Kichler A, Leborgne C, Danos O. 2005. Dilution of reporter gene with stuffer DNA does not alter the transfection efficiency of polyethylenimines. *J. Gene Med.* **7**:1459–1467.
- Kildegaard HF, Baycin-Hizal D, Lewis NE, Betenbaugh MJ. 2013. The emerging CHO systems biology era: Harnessing the 'omics revolution for biotechnology. *Curr. Opin. Biotechnol.* **24**:1102–1107.
- Kim JY, Kim YG, Lee GM. 2012. CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. *Appl. Microbiol. Biotechnol.* **93**:917–930.
- Kim NS, Lee GM. 2000. Overexpression of bcl-2 inhibits sodium butyrate-induced apoptosis in Chinese hamster ovary cells resulting in enhanced humanized antibody production. *Biotechnol. Bioeng.* **71**:184–193.
- Kim NS, Lee GM. 2002. Inhibition of sodium butyrate-induced apoptosis in recombinant Chinese hamster ovary cells by constitutively expressing antisense RNA of caspase-3. *Biotechnol.*

- Bioeng.* **78**:217–228.
- Kim TK, Eberwine JH. 2010. Mammalian cell transfection: the present and the future. *Anal. Bioanal. Chem.* **397**:3173–8.
- Kitney R, Freemont P. 2012. Synthetic biology - the state of play. *FEBS Lett.* **586**:2029–36.
- Kohler G, Milstein C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**:495–497.
- Kontermann RE, Brinkmann U. 2015. Bispecific antibodies. *Drug Discov. Today* **20**:838–847.
- Kou T-C, Fan L, Zhou Y, Ye Z-Y, Liu X-P, Zhao L, Tan W-S. 2011. Detailed understanding of enhanced specific productivity in Chinese hamster ovary cells at low culture temperature. *J. Biosci. Bioeng.* **111**:465–469.
- Krämer O, Klausing S, Noll T. 2010. Methods in mammalian cell line engineering: From random mutagenesis to sequence-specific approaches. *Appl. Microbiol. Biotechnol.* **88**:425–436.
- Kretzmer G. 2002. Industrial processes with animal cells. *Appl. Microbiol. Biotechnol.* **59**:135–42.
- Kriz A, Schmid K, Baumgartner N, Ziegler U, Berger I, Ballmer-Hofer K, Berger P. 2010. A plasmid-based multigene expression system for mammalian cells. *Nat. Commun.* **1**:120.
- Ku S, Ng D, Yap M, Chao S-H. 2007. Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. *Biotechnol. Bioeng.* **99**:155–64.
- Kunaparaju R, Liao M, Sunstrom N-A. 2005. Epi-CHO, an episomal expression system for recombinant protein production in CHO cells. *Biotechnol. Bioeng.* **91**:670–677.
- Kwok R. 2010. Five hard truths for synthetic biology. *Nature* **463**:288–290.
- Kwon R-J, Kim SK, Lee S-I, Hwang S-J, Lee GM, Kim J-S, Seol W. 2006. Artificial transcription factors increase production of recombinant antibodies in Chinese hamster ovary cells. *Biotechnol. Lett.* **28**:9–15.
- Lai T, Yang Y, Ng SK. 2013. Advances in mammalian cell line development technologies for recombinant protein production. *Pharmaceuticals* **6**:579–603.
- Lalonde M-E, Durocher Y. 2017. Therapeutic glycoprotein production in mammalian cells. *J. Biotechnol.* **251**:128–140.
- Lang S, Benedix J, Fedeles S V., Schorr S, Schirra C, Schauble N, Jalal C, Greiner M, Hassdenteufel S, Tatzelt J, Kreutzer B, Edelmann L, Krause E, Rettig J, Somlo S, Zimmermann R, Dudek J. 2012. Different effects of Sec61, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *J. Cell Sci.* **125**:1958–1969.
- Lee A-H, Iwakoshi NN, Glimcher LH. 2003. XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. *Mol. Cell. Biol.* **23**:7448–59.
- Lee JS, Grav LM, Lewis NE, Kildegaard HF. 2015. CRISPR/Cas9-mediated genome engineering of CHO cell factories: Application and perspectives. *Biotechnol. J.* **10**:979–994.
- Lewis AM, Abu-Absi NR, Borys MC, Li ZJ. 2016. The use of 'Omics technology to rationally improve industrial mammalian cell line performance. *Biotechnol. Bioeng.* **113**:26–38.
- Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO. 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. *Nat. Biotechnol.* **31**:759–67.
- Li F, Vijayasankaran N, Shen A (Yijuan), Kiss R, Amanullah A. 2010. Cell culture processes for monoclonal antibody production. *MAbs* **2**:466–479.
- Lim SF, Chuan KH, Liu S, Loh SOH, Chung BYF, Ong CC, Song Z. 2006. RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab. Eng.* **8**:509–522.
- Lim Y, Wong NSC, Lee YY, Ku SCY, Wong DCF, Yap MGS. 2010. Engineering mammalian cells in bioprocessing - current achievements and future perspectives. *Biotechnol. Appl. Biochem.* **55**:175–89.
- Lin-hong L, Shivakumar R, Feller S, Allen C, Weiss JM, Dzekunov S, Singh V, Holaday J, Fratantoni J, Liu LN. 2002. Highly efficient, large volume flow electroporation. *Technol.*

- Cancer Res. Treat.* **1**:341–50.
- Lipman NS, Jackson LR, Trudel LJ, Weis-Garcia F. 2005. Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources. *Inst. Lab. Anim. Res. J.* **46**:258–268.
- Lodish H, Berk A, Kaiser C, Krieger M, Scott M, Bretscher A, Ploegh H, Matsudaira P. 2008. Molecular Cell Biology. Ed. Katherine Ahr Sixth Edit. W.H Freeman and Company 11, 12, 534, 581 p.
- Lonza. 2009. Amaxa™ 96-well Shuttle™ Protocol for CHO-S (Invitrogen). http://bio.lonza.com/fileadmin/groups/marketing/Downloads/Protocols/Generated/Optimized_Protocol_190.pdf.
- Lonza Amaxa. Nucleofector™ Technology. *Nucleofector™ Technol. Flyer*. http://bio.lonza.com/fileadmin/groups/FAQs/public/Technology_Flyer.pdf.
- Majors BS, Arden N, Oyler GA, Chiang GG, Pederson NE, Betenbaugh MJ. 2008a. E2F-1 overexpression increases viable cell density in batch cultures of Chinese hamster ovary cells. *J. Biotechnol.* **138**:103–106.
- Majors BS, Betenbaugh MJ, Pederson NE, Chiang GG. 2008b. Enhancement of transient gene expression and culture viability using Chinese hamster ovary cells overexpressing Bcl-x(L). *Biotechnol. Bioeng.* **101**:567–78.
- Majors BS, Betenbaugh MJ, Pederson NE, Chiang GG. 2009. Mcl-1 overexpression leads to higher viabilities and increased production of humanized monoclonal antibody in Chinese hamster ovary cells. *Biotechnol. Prog.* **25**:1161–8.
- Mansouri M, Berger P. 2014. Strategies for multigene expression in eukaryotic cells. *Plasmid* **75**:12–17.
- Mariotto AB, Robin Yabroff K, Shao Y, Feuer EJ, Brown ML. 2011. Projections of the cost of cancer care in the United States: 2010–2020. *J. Natl. Cancer Inst.* **103**:117–128.
- Masterton R, Smales CM. 2014. The impact of process temperature on mammalian cell lines and the implications for the production of recombinant proteins in CHO cells **2**:49–61.
- Materna SC, Marwan W. 2005. Estimating the number of plasmids taken up by a eukaryotic cell during transfection and evidence that antisense RNA abolishes gene expression in *Physarum polycephalum*. *FEMS Microbiol. Lett.* **243**:29–35.
- Mehier-Humbert S, Guy RH. 2005. Physical methods for gene transfer: Improving the kinetics of gene delivery into cells. *Adv. Drug Deliv. Rev.* **57**:733–753.
- Misaghi S, Chang J, Snedecor B. 2014. It's time to regulate: Coping with product-induced nongenetic clonal instability in CHO cell lines via regulated protein expression. *Biotechnol. Prog.* **30**:1432–1440.
- Mohan C, Kim YG, Koo J, Lee GM. 2008. Assessment of cell engineering strategies for improved therapeutic protein production in CHO cells. *Biotechnol J* **3**:624–630.
- Mohan C, Lee GM. 2010. Effect of inducible co-overexpression of protein disulfide isomerase and endoplasmic reticulum oxidoreductase on the specific antibody productivity of recombinant Chinese hamster ovary cells. *Biotechnol. Bioeng.* **107**:337–46.
- Mohan C, Park SH, Chung JY, Lee GM. 2007. Effect of doxycycline-regulated protein disulfide isomerase expression on the specific productivity of recombinant CHO cells: thrombopoietin and antibody. *Biotechnol. Bioeng.* **98**:611–5.
- Mozley OL, Thompson BC, Fernandez-Martell A, James DC. 2014. A mechanistic dissection of polyethylenimine mediated transfection of CHO cells: To enhance the efficiency of recombinant DNA utilization. *Biotechnol. Prog.* **30**:1161–1170.
- Nelson AL. 2010. Antibody fragments: hope and hype. *Landes Biosci.* **2**:77–83.
- Neumann E, Schaefer-Ridder M, Wang Y, Hofschneider PH. 1982. Gene transfer into mouse lymphoma cells by electroporation in high electric fields. *EMBO J.* **1**:841–845.
- Nishimiya D. 2014. Proteins improving recombinant antibody production in mammalian cells. *Appl. Microbiol. Biotechnol.* **98**:1031–42.
- Nishimiya D, Mano T, Miyadai K, Yoshida H, Takahashi T. 2013. Overexpression of CHOP alone and in combination with chaperones is effective in improving antibody production in

- mammalian cells. *Appl. Microbiol. Biotechnol.* **97**:2531–2539.
- Nissom PM, Sanny A, Kok YJ, Hiang YT, Chuah SH, Shing TK, Lee YY, Wong KTK, Hu W-S, Sim MYG, Philp R. 2006. Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. *Mol. Biotechnol.* **34**:125–140.
- Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. 2015. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* **15**:160–171.
- Nyathi Y, Wilkinson BM, Pool MR. 2013. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim. Biophys. Acta - Mol. Cell Res.* **1833**:2392–2402.
- O’Callaghan PM, McLeod J, Pybus LP, Lovelady CS, Wilkinson SJ, Racher AJ, Porter A, James DC. 2010. Cell line-specific control of recombinant monoclonal antibody production by CHO cells. *Biotechnol. Bioeng.* **106**:938–51.
- Oh-hashii K, Koga H, Ikeda S, Shimada K, Hirata Y, Kiuchi K. 2009. CRELD2 is a novel endoplasmic reticulum stress-inducible gene. *Biochem. Biophys. Res. Commun.* **387**:504–510.
- Ohya T, Hayashi T, Kiyama E, Nishii H, Miki H, Kobayashi K, Honda K, Omasa T, Ohtake H. 2008. Improved Production of Recombinant Human Antithrombin III in Chinese Hamster Ovary Cells by ATF4 Overexpression. *Biotechnol. Bioeng.* **100**:317–324.
- Omasa T, Takami T, Ohya T, Kiyama E, Hayashi T, Nishii H, Miki H, Kobayashi K, Honda K, Ohtake H. 2008. Overexpression of GADD34 enhances production of recombinant human antithrombin III in Chinese hamster ovary cells. *J. Biosci. Bioeng.* **106**:568–73.
- Oyadomari S, Mori M. 2004. Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death Differ.* **11**:381–389.
- Pandhal J, Wright PC. 2010. N-Linked glycoengineering for human therapeutic proteins in bacteria. *Biotechnol. Lett.* **32**:1189–1198.
- Pascoe DE, Arnott D, Papoutsakis ET, Miller WM, Andersen DC. 2007. Proteome Analysis of Antibody-Producing CHO Cell Lines With Different Metabolic Profiles. *Biotechnol. Bioeng.* **98**:391–410.
- Peng R-W, Fussenegger M. 2009. Molecular engineering of exocytic vesicle traffic enhances the productivity of Chinese hamster ovary cells. *Biotechnol. Bioeng.* **102**:1170–81.
- Peng R, Abellan E, Fussenegger M. 2011. Differential Effect of Exocytic SNAREs on the Production of Recombinant Proteins in Mammalian Cells. *Biotechnol. Bioeng.* **108**:611–620.
- Perlmutter DH. 2002. Chemical chaperones: A pharmacological strategy for disorders of protein folding and trafficking. *Pediatr. Res.* **52**:832–836.
- Petrone J. 2016. DNA writers attract investors. *Nat. Biotechnol.* **34**:363–364.
- Pichler J, Galosy S, Mott J, Borth N. 2011. Selection of CHO host cell subclones with increased specific antibody production rates by repeated cycles of transient transfection and cell sorting. *Biotechnol. Bioeng.* **108**:386–394.
- Pollard MG, Travers K evenJ, Weissman JS. 1998. Ero1p: a novel and ubiquitous protein with an essential role in oxidative protein folding in the endoplasmic reticulum. *Mol. Cell* **1**:171–82.
- Priola JJ, Calzadilla N, Baumann M, Borth N, Tate CG, Betenbaugh MJ. 2016. High-throughput screening and selection of mammalian cells for enhanced protein production. *Biotechnol. J.* **11**:853–865.
- Proudfoot NJ. 1986. Transcriptional interference and termination between duplicated α -globin gene constructs suggests a novel mechanism for gene regulation. *Nature* **322**:562–565.
- Puck TT, Kao FT. 1967. Genetics of somatic mammalian cells, V. Treatment with 5-bromodeoxyuridine and visible light for isolation of nutritionally deficient mutants. *Proc. Natl. Acad. Sci. U. S. A.* **58**:1227–1234.
- Pybus LP, Dean G, West NR, Smith A, Daramola O, Field R, Wilkinson SJ, James DC. 2013. Model-directed engineering of “difficult-to-express” monoclonal antibody production by Chinese hamster ovary cells. *Biotechnol. Bioeng.* **111**:372–385.
- Rader RA. 2008. (Re)defining biopharmaceutical. *Nat. Biotechnol.* **26**:743–751.
- Rahimpour A, Vaziri B, Moazzami R, Nematollahi L, Barkhordari F, Kokabee L, Adeli A, Mahboudi F. 2013. Engineering the Cellular Protein Secretory Pathway for Enhancement of

- Recombinant Tissue Plasminogen Activator Expression in Chinese Hamster Ovary Cells: Effects of Cert and XBP1s Genes. *J. Microbiol. Biotechnol.* **23**:1116–1122.
- Rajan RS, Tsumoto K, Tokunaga M, Tokunaga H, Kita Y, Arakawa T. 2011. Chemical and pharmacological chaperones: application for recombinant protein production and protein folding diseases. *Curr. Med. Chem.* **18**:1–15.
- Rajendra Y, Kiseljak D, Baldi L, Hacker DL, Wurm FM. 2011. A simple high-yielding process for transient gene expression in CHO cells. *J. Biotechnol.* **153**:22–26.
- Rajendra Y, Kiseljak D, Manoli S, Baldi L, Hacker DL, Wurm FM. 2012. Role of non-specific DNA in reducing coding DNA requirement for transient gene expression with CHO and HEK-293E cells. *Biotechnol. Bioeng.* **109**:2271–2278.
- Rita Costa A, Elisa Rodrigues M, Henriques M, Azeredo J, Oliveira R. 2010. Guidelines to cell engineering for monoclonal antibody production. *Eur. J. Pharm. Biopharm.* **74**:127–38.
- Roth SD, Schüttrumpf J, Milanov P, Abriss D, Ungerer C, Quade-Lyssy P, Simpson JC, Pepperkok R, Seifried E, Tonn T. 2012. Chemical Chaperones Improve Protein Secretion and Rescue Mutant Factor VIII in Mice with Hemophilia A. *PLoS One* **7**.
- Rupp O, Becker J, Brinkrolf K, Timmermann C, Borth N, Pühler A, Noll T, Goesmann A. 2014. Construction of a public CHO cell line transcript database using versatile bioinformatics analysis pipelines. *PLoS One* **9**.
- Sasaki K, Yoshida H. 2015. Organelle autoregulation - Stress responses in the ER, Golgi, mitochondria and lysosome. *J. Biochem.* **157**:185–195.
- Sato Y, Kojima R, Okumura M, Hagiwara M, Masui S, Maegawa K, Saiki M, Horibe T, Suzuki M, Inaba K. 2013. Synergistic cooperation of PDI family members in peroxiredoxin 4-driven oxidative protein folding. *Sci. Rep.* **3**:2456.
- Schmidpeter PAM, Schmid FX. 2015. Prolyl isomerization and its catalysis in protein folding and protein function. *J. Mol. Biol.* **427**:1609–1631.
- Schröder M. 2008. Engineering eukaryotic protein factories. *Biotechnol. Lett.* **30**:187–96.
- Schroeder H, Cavacini L. 2010. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**:41–52.
- Serrano L. 2007. Synthetic biology: promises and challenges. *Mol. Syst. Biol.* **3**:1–5.
- Sethuraman N, Stadheim T a. 2006. Challenges in therapeutic glycoprotein production. *Curr. Opin. Biotechnol.* **17**:341–6.
- Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, Yu X, Yang L, Tan BK, Rosenwald A, Hurt EM, Petroulakis E, Sonenberg N, Yewdell JW, Calame K, Glimcher LH, Staudt LM. 2004. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21**:81–93.
- Shaner NC, Steinbach PA, Tsien RY. 2005. A guide to choosing fluorescent proteins. *Nat. Methods* **2**:905–909.
- Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference – a crash course. *Trends Genet.* **21**:339–345.
- Shen J, Chen X, Hendershot L, Prywes R. 2002. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of golgi localization signals. *Dev. Cell* **3**:99–111.
- Shukla AA, Gottschalk U. 2013. Single-use disposable technologies for biopharmaceutical manufacturing. *Trends Biotechnol.* **31**:147–154.
- Shukla AA, Thömmes J. 2010. Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends Biotechnol.* **28**:253–261.
- Siegel RL, Miller KD, Jemal A. 2016. Cancer statistics. *CA Cancer J Clin* **66**:7–30.
- Smales CM, Dinnis DM, Stansfield SH, Alete D, Sage EA, Birch JR, Racher AJ, Marshall CT, James DC. 2004. Comparative proteomic analysis of GS-NSO murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnol. Bioeng.* **88**:474–488.
- Sou SN, Sellick C, Lee K, Mason A, Kyriakopoulos S, Polizzi KM, Kontoravdi C. 2015. How does

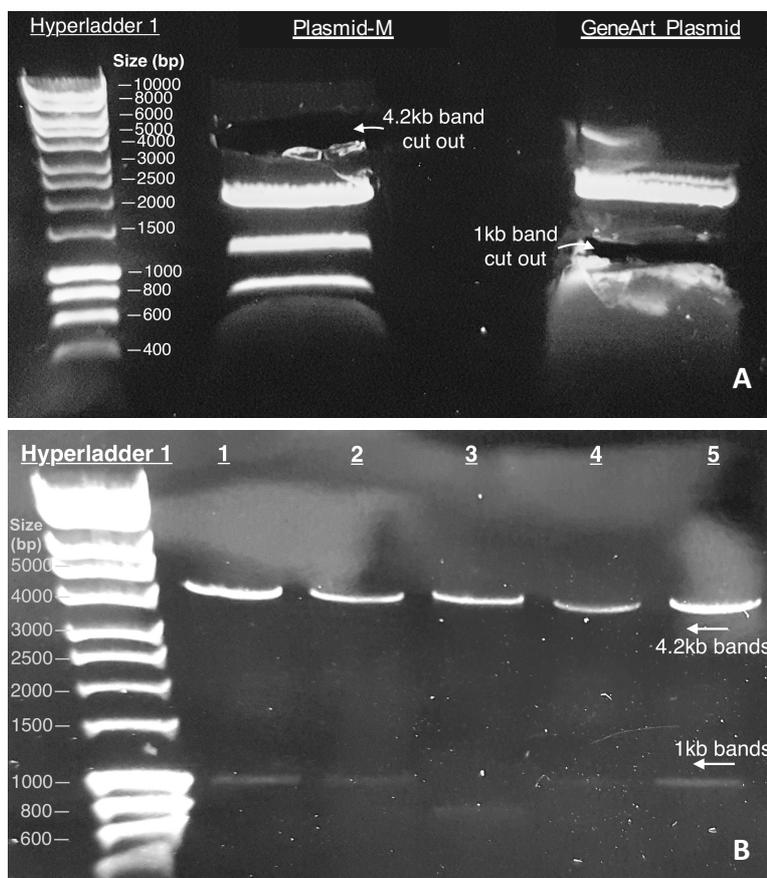
- mild hypothermia affect monoclonal antibody glycosylation? *Biotechnol. Bioeng.* **112**:1165–1176.
- Steger K, Brady J, Wang W, Duskin M, Donato K, Peshwa M. 2015. CHO-S antibody titers >1 gram/liter using flow electroporation-mediated transient gene expression followed by rapid migration to high-yield stable cell lines. *J. Biomol. Screen.* **20**:545–51.
- Strack RL, Hein B, Bhattacharyya D, Hell SW, Keenan RJ, Glick BS. 2009. A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* **48**:8279–8281.
- Subach OM, Gundorov IS, Yoshimura M, Subach F V., Zhang J, Grünwald D, Souslova EA, Chudakov DM, Verkhusha V V. 2008. Conversion of Red Fluorescent Protein into a Bright Blue Probe. *Chem. Biol.* **15**:1116–1124.
- Sung YH, Lee GM. 2005. Enhanced human thrombopoietin production by sodium butyrate addition to serum-free suspension culture of Bcl-2-overexpressing CHO cells. *Biotechnol. Prog.* **21**:50–57.
- Sung YH, Lee JS, Park SH, Koo J, Lee GM. 2007. Influence of co-down-regulation of caspase-3 and caspase-7 by siRNAs on sodium butyrate-induced apoptotic cell death of Chinese hamster ovary cells producing thrombopoietin. *Metab. Eng.* **9**:452–464.
- Szul T, Sztul E. 2011. COPII and COPI Traffic at the ER-Golgi Interface. *Physiology* **26**:348–364.
- Tabar MS, Hesaraki M, Esfandiari F, Samani FS, Vakilian H, Baharvand H. 2015. Evaluating electroporation and lipofectamine approaches for transient and stable transgene expressions in human fibroblasts and embryonic stem cells. *Cell J.* **17**:438–450.
- Tan JGL, Lee YY, Wang T, Yap MGS, Tan TW, Ng SK. 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. *Biotechnol. J.* **10**:790–800.
- Taniguchi M, Nadanaka S, Tanakura S, Sawaguchi S, Midori S, Kawai Y, Yamaguchi S, Shimada Y, Nakamura Y, Matsumura Y, Fujita N, Araki N, Yamamoto M, Oku M, Wakabayashi S, Kitagawa H, Yoshida H. 2015. TFE3 Is a bHLH-ZIP-type Transcription Factor that Regulates the Mammalian Golgi Stress Response. *Cell Struct. Funct.* **40**:13–30.
- Tastanova A, Schulz A, Folcher M, Tolstrup A, Puklowski A, Kaufmann H, Fussenegger M. 2016. Overexpression of YY1 increases the protein production in mammalian cells. *J. Biotechnol.* **219**:72–85.
- Telford WG, Hawley T, Subach F, Verkhusha V, Hawley RG. 2012. Flow cytometry of fluorescent proteins. *Methods* **57**:318–330.
- Thompson BC, Segarra CRJ, Mozley OL, Daramola O, Field R, Levison PR, James DC. 2012. Cell line specific control of polyethylenimine-mediated transient transfection optimized with “Design of experiments” methodology. *Biotechnol. Prog.* **28**:179–187.
- Thompson B, Clifford J, Jenns M, Smith A, Field R, Nayyar K, James DC. 2017. High-throughput quantitation of Fc-containing recombinant proteins in cell culture supernatant by fluorescence polarization spectroscopy. *Anal. Biochem.* **534**:49–55.
- Thoring L, Wüstenhagen DA, Borowiak M, Stech M, Sonnabend A, Kubick S. 2016. Cell-free systems based on CHO cell lysates: Optimization strategies, synthesis of “difficult-to-express” proteins and future perspectives. *PLoS One* **11**:1–21.
- Tigges M, Fussenegger M. 2006. Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. *Metab. Eng.* **8**:264–72.
- Tjio JH, Puck TT. 1958. Genetics of Somatic Mammalian Cells 2. Chromosomal Constitution of Cells in Tissue Culture. *J. Exp. Med.* **108**:259–268.
- Underhill MF, Smales CM, Naylor LH, Birch JR, James DC. 2007. Transient Gene Expression Levels from Multigene Expression Vectors. *Biotechnol. Prog.* **23**:435–443.
- Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**:97–130.
- Walsh G. 2002. Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. *Eur. J. Pharm. Sci.* **15**:135–138.
- Walsh G. 2010. Biopharmaceutical benchmarks 2010. *Nat. Biotechnol.* **28**:917–24.

- Walsh G. 2014. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32**:992–1000.
- Walsh G, Jefferis R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **24**:1241–52.
- Walter P, Ron D. 2012. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science (80-)*. **334**:1081–1086.
- Wang Z, Ma X, Fan L, Rhee WJ, Park TH, Zhao L, Tan WS. 2012. Understanding the mechanistic roles of 30Kc6 gene in apoptosis and specific productivity in antibody-producing Chinese hamster ovary cells. *Appl. Microbiol. Biotechnol.* **94**:1243–1253.
- Wedemeyer WJ, Welker E, Scheraga H a. 2002. Proline Cis-Trans Isomerization and Protein Folding. *Biochemistry* **41**:14637–14644.
- Whitford WG. 2006. Fed-Batch Mammalian Cell Culture in Bioproduction. *Bioprocess Int.* **4**:30–40.
- Wu S-C. 2009. RNA interference technology to improve recombinant protein production in Chinese hamster ovary cells. *Biotechnol. Adv.* **27**:417–22.
- Wulhfard S, Baldi L, Hacker D, Wurm F. 2010. Valproic acid enhances recombinant mRNA and protein levels in transiently transfected Chinese hamster ovary cells. *J. Biotechnol.* **148**:128–132.
- Wurm F, Bernard A. 1999. Large-scale transient expression in mammalian cells for recombinant protein production. *Curr. Opin. Biotechnol.* **10**:156–9.
- Wurm F. 2013. CHO Quasispecies—Implications for Manufacturing Processes. *Processes* **1**:296–311.
- Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.* **22**:1393–8.
- Xiao S, Shiloach J, Betenbaugh MJ. 2014. Engineering cells to improve protein expression. *Curr. Opin. Struct. Biol.* **26**:32–38.
- Xie L, Wang DIC. 2006. Fed-batch cultivation of animal cells using different medium design concepts and feeding strategies. 1994. *Biotechnol. Bioeng.* **95**:270–84.
- Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol* **29**:735–741.
- Yahata K, Kishine H, Sone T, Sasaki Y, Hotta J, Chesnut JD, Okabe M, Imamoto F. 2005. Multi-gene Gateway clone design for expression of multiple heterologous genes in living cells: Conditional gene expression at near physiological levels. *J. Biotechnol.* **118**:123–134.
- Yahata K, Maeshima K, Sone T, Ando T, Okabe M, Imamoto N, Imamoto F. 2007. cHS4 Insulator-mediated Alleviation of Promoter Interference during Cell-based Expression of Tandemly Associated Transgenes. *J. Mol. Biol.* **374**:580–590.
- Yamamoto K, Gandin V, Sasaki M, McCracken S, Li W, Silvester JL, Elia AJ, Wang F, Wakutani Y, Alexandrova R, Oo YD, Mullen PJ, Inoue S, Itsumi M, Lapin V, Haight J, Wakeham A, Shahinian A, Ikura M, Topisirovic I, Sonenberg N, Mak TW. 2014. Largen: A molecular regulator of mammalian cell size control. *Mol. Cell* **53**:904–915.
- Yamamoto K, Sato T, Matsui T, Sato M, Okada T, Yoshida H, Harada A, Mori K. 2007. Transcriptional Induction of Mammalian ER Quality Control Proteins Is Mediated by Single or Combined Action of ATF6 α and XBP1. *Dev. Cell* **13**:365–376.
- Yamane-Ohnuki N, Kinoshita S, Inoue-Urakubo M, Kusunoki M, Iida S, Nakano R, Wakitani M, Niwa R, Sakurada M, Uchida K, Shitara K, Satoh M. 2004. Establishment of FUT8 knockout Chinese hamster ovary cells: An ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol. Bioeng.* **87**:614–622.
- Ye J, Kober V, Tellers M, Naji Z, Salmon P, Markusen JF. 2009. High-level protein expression in scalable CHO transient transfection. *Biotechnol. Bioeng.* **103**:542–551.
- Yee JC, Gerdtsen ZP, Hu WS. 2009. Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. *Biotechnol. Bioeng.* **102**:246–

263.

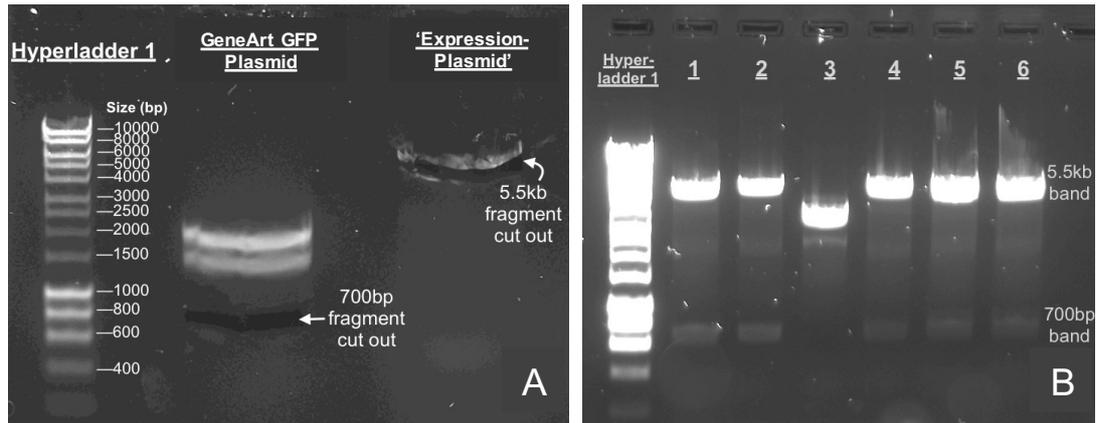
- Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* **107**:881–891.
- Zhang X, Garcia IF, Baldi L, Hacker DL, Wurm FM. 2010. Hyperosmolarity enhances transient recombinant protein yield in Chinese hamster ovary cells. *Biotechnol. Lett.* **32**:1587–1592.
- Zhu J. 2012. Mammalian cell protein expression for biopharmaceutical production. *Biotechnol. Adv.* **30**:1158–70.

Appendix A



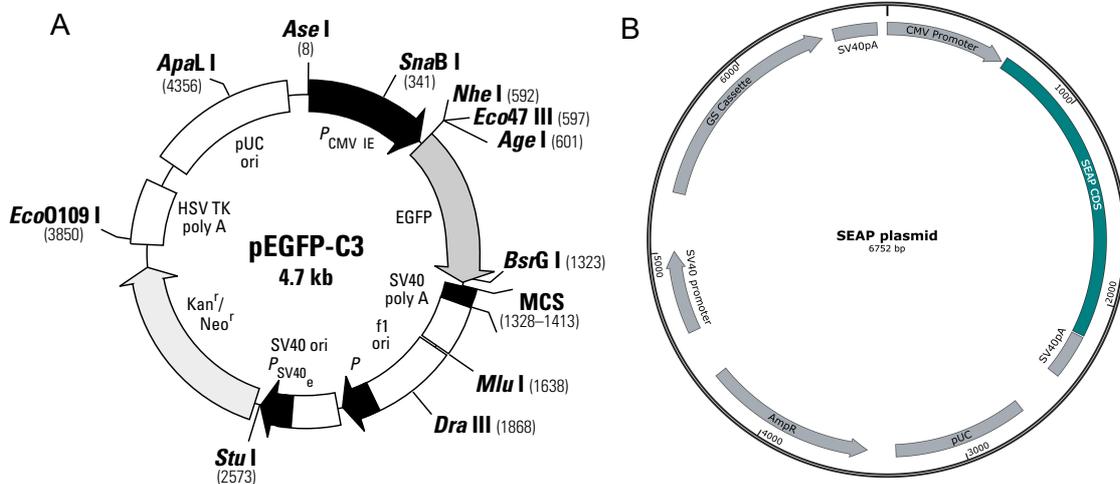
Appendix Figure A.1: Construction of 'Expression-Plasmid'

(A) Agarose gel separation of DNA fragments, from HindIII and KpnI digestion, of Plasmid-M and the synthesised GeneArt plasmid. The required fragments (4.2kb and 1kb, respectively) were cut out from the gel, purified and ligated together. (B) Positive colonies from the ligation reaction were amplified, purified using mini-prep procedure then digested with HindIII and KpnI and separated by gel electrophoresis to identify successful ligations. Identity of plasmid #5 ('Expression-Plasmid') confirmed with sequence verification. Hyperladder 1 present on all gels enabled approximate size determination.



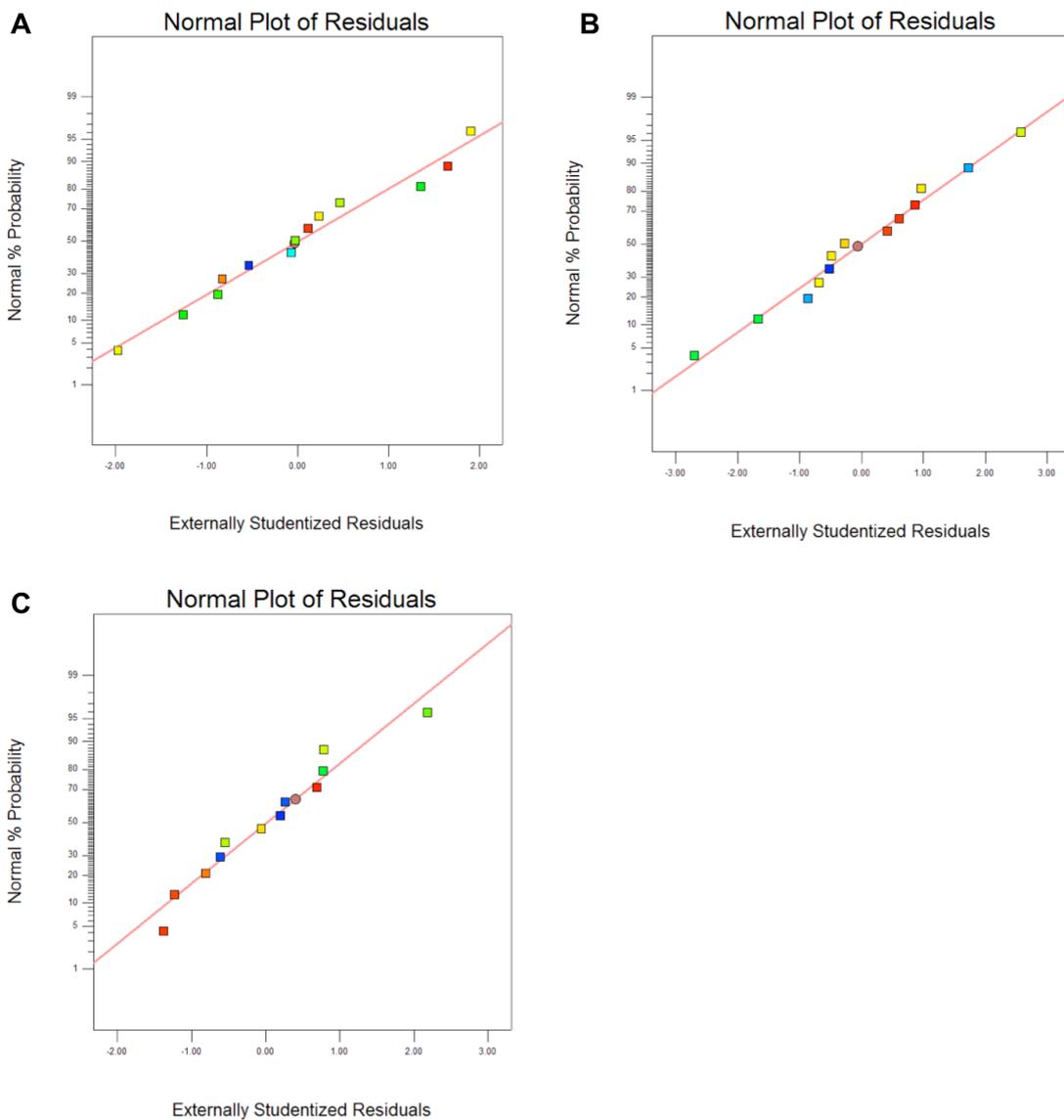
Appendix Figure A.2: Construction of eGFP 'Expression-Plasmid'

(A) Through digestion with *Agel* and *SbfI* and separation by gel electrophoresis, the insert fragment and the linearised expression-plasmid were isolated. (B) The two sequences with compatible ends were ligated together, with selected positive colonies checked for successful integration using a double digest with *Agel/SbfI*. The presence, in this case, of 700bp and ~5.5kb bands indicate successful integration.



Appendix Figure A.3: Plasmid Maps of eGFP-C3 and SEAP Plasmids

(A) pEGFP-C3 vector (GenBank Accession number: U57607) encoding eGFP under a CMV promoter. The plasmid was provided by Dr Andrew Peden (University of Sheffield). (B) Plasmid encoding SEAP under a CMV promoter, plasmid also encodes GS cassette. Plasmid provided by Dr Adam Brown (University of Sheffield)



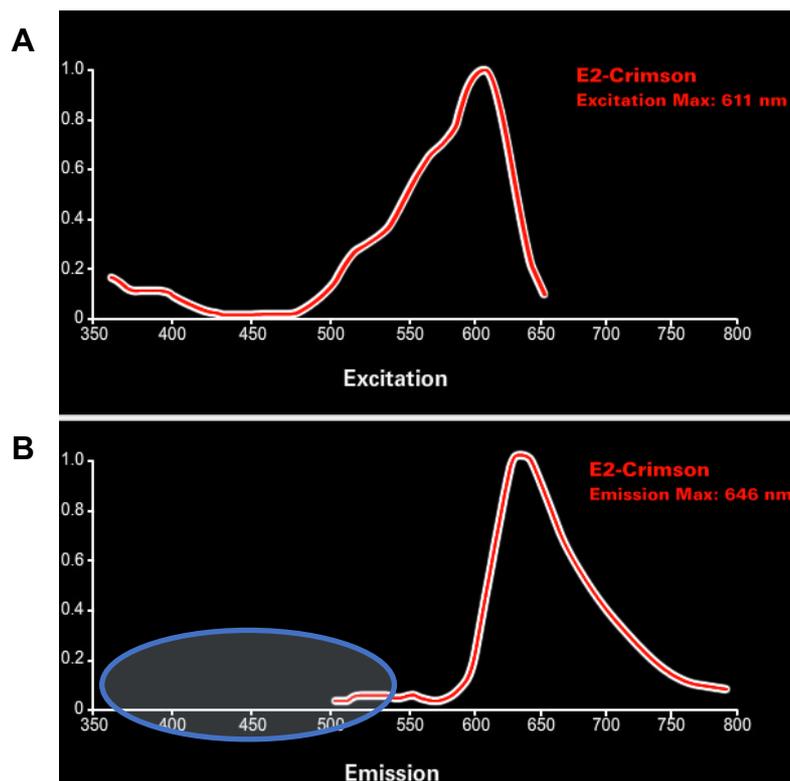
Appendix Figure A.4: Normal Plots of Residuals for DoE-RSM: Cell Number and DNA Load

Normality plots for: Viability (A) Titre (B) and qP (C).

Appendix B

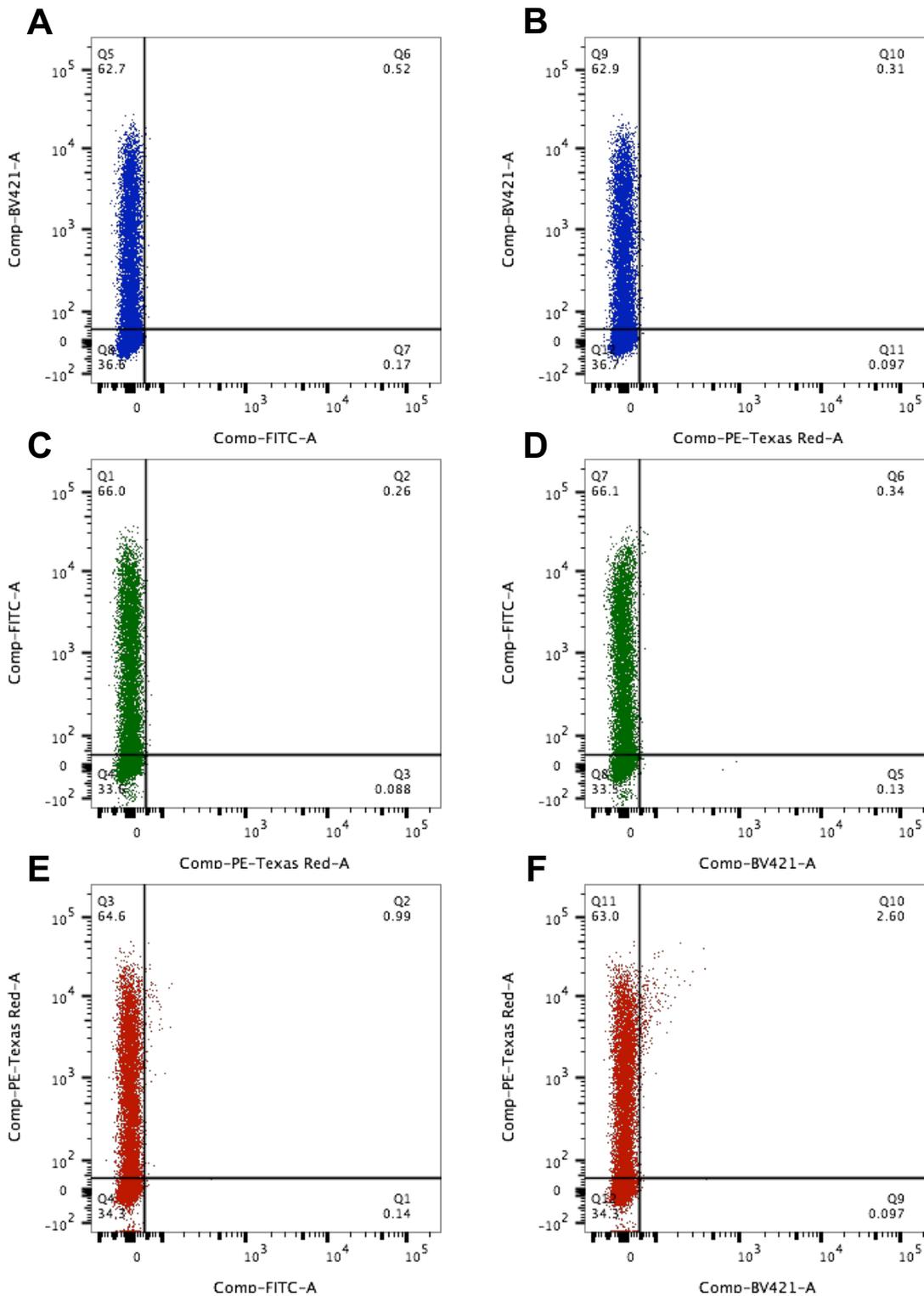
Appendix Table B.1: Nucleotide Sequences for Fluorescent Proteins

Fluorescent Protein	Nucleotide Sequence
eGFP	<p>atggtgagcaagggcgaggagctgtcaccggggtggtgccatcctggctgagctggacggcgacgt aaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgacct gaagttcatctgcaccaccggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacgg cgtgcagtgttcagccgtaccccaccacatgaagcagcagcacttctcaagtccgccatgccga aggctacgtccaggagcgcaccatcttcaaggacgacggcaactacaagaccgcgccgaggtg aagttcaggggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggc aacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagc agaagaacggcatcaaggtgaactcaagatccgccacaacatcgaggacggcagcgtgacgtcgc ccgaccactaccagcagaacaccccacatcggcgacggcccctgctgctgcccgacaaccactacct gagcaccagtcggccctgagcaagaccccaacgagaagcgcgatcacatggtcctgctggagttc gtgaccgccgcccggatcactctcgcatggacgagctgtacaagtga</p>
TagBFP	<p>atgagcgagctgattaaggagaacatgcacatgaagctgtacatggagggcaccgtggacaaccatc actcaagtgacatccgagggcgaaggcaagccctacgagggcaccagaccatgagaatcaag gtggtcgagggcgccctctccccttcgcttcgacatcctggctactagcttctctacggcagcaagac cttcatcaaccacaccagggtatccccgacttctcaagcagtcctccctgagggcttcatgaggag agagtcaccacatacgaagacggggcgctgctgaccgctaccaggacaccagcctccaggacgg ctgctcatctacaacgtcaagatcagaggggtgaacttcacatccaacggccctgtgatgcagaaga aaacactcggctgggagccctcaccgagacgctgtaccccgtgacggcggcctggaaggcagaa acgacatggccctgaagctcgtggcgaggccatctgatcgcaaacatcaagaccacatatagatcc aagaaacccgctaagaacctcaagatgctggcgttactatgtggactacagactggaaagaatcaa ggaggccaacaacgagacctacgtcgagcagcagaggtggcagtgccagatactgcgacctccc tagcaaaactggggcacaagcttaattaa</p>
mCherry	<p>atggtgtccaagggcgaagaggacaacatggccatcatcaagagttcatgcggtcaaggtgcacat ggaaggctcgtgaacggccacgagttcgagatcgagggcgagggcgaaggcagaccctacgagg gcaccagaccgccaagctgaaagtgaccaagggcgaccctgccctcgctgggacatcctgtc ccccagttatgtacggctccaaggcctacgtgaagcaccggccgatccccgactacctgaagct gagctccccgagggctcaagtgggagagagtgatgaactcgaggacggcggcggtgacagtg accaggacagctctctcaggacggcgagttcatctcaaagtgaagctgcggggcaccaactccc ttccgacggcccctgatgcagaaaaagacaatgggctgggaggcctcctccgagcggatgtaccctg aggatggcggcctgaagggcgagatcaagcagcggctgaagctgaaggatggcggcactacgac gccgaagtgaaaaccctacaaggccaagaaacccgtgacgtgctggcctacaacgtgaac atcaagctggacatcacctcccacaacgaggactacaccatcgtggaacagtagcagcgggcccag ggcagacactccaccggggcatggacgagctgtacaagtga</p>



Appendix Figure B.1: Excitation and Emission Spectra for E2-crimson

The excitation (A) and emission spectra (B) published for E2 crimson (images adapted from <http://learning.clontech.com/spectra.html>, Clontech, Takara Bio Europe). There appeared to be a minimal excitation by a 405nm laser, and no detection within a 450/50 bandwidth filter (highlighted by blue ellipse), the configuration used for TagBFP analysis.



Appendix Figure B.2: Flow cytometry Analysis of Single Fluorescent Protein Samples

Expression derived from transfections with each fluorescent protein separately: TagBFP (A, B), eGFP (C, D) and mCherry (E, F). Demonstrates independent detection of each fluorescent protein when measured with parameters used for detection of the other proteins.

Appendix C

Appendix Table C.1: Effector Gene Library Information: Protein Processing in the ER

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/</i> <u>Mouse</u>	References
BiP/ HSPA5/ GRP78	<ul style="list-style-type: none"> ER molecular chaperone Role in folding and assembly of proteins in the ER Role in UPR activation 	<ul style="list-style-type: none"> Previous overexpression studies, resulting in up to 1.5-fold improved productivity levels of mAbs/fusion proteins in CHO cells Widely associated with increased CHO production capacity Upregulated in plasma cells +2.5-fold (RNAseq) and +1.3-fold (Affymetrix) 	Human	3309	<i>P11021</i> <u>P20029</u>	Borth et al., 2005; Carlage et al., 2012; Doolan et al., 2008; Harreither et al., 2015; Johari et al., 2015; Pybus et al., 2013; Yee et al., 2009
Calr/ Calreticulin	<ul style="list-style-type: none"> ER molecular lectin chaperone Role in quality control of protein folding Holdase/foldase function, preventing aggregation, retaining incorrectly folded proteins for degradation 	<ul style="list-style-type: none"> Upregulated in plasma cells, +1.75-fold (RNAseq) Associated with increased CHO production capacity Previous overexpression (with calnexin) improved productivity of thrombopoietin up to 1.9-fold in CHO cells 	Human	811	<i>P27797</i> <u>P14211</u>	Carlage et al., 2012; Chung et al., 2004; Harreither et al., 2015
CANX/ Calnexin	<ul style="list-style-type: none"> ER molecular lectin chaperone Role in quality control of protein folding Holdase/foldase function, preventing aggregation, retaining incorrectly folded proteins for degradation 	<ul style="list-style-type: none"> Upregulated in plasma cells, +1.3-fold (RNAseq) and +2.1-fold (Affymetrix) Associated with increased CHO production capacity Previous overexpression (with calreticulin) improved productivity of thrombopoietin up to 1.9-fold in CHO cells 	Human	821	<i>P27824</i> <u>P35564</u>	Carlage et al., 2012; Chung et al., 2004; Doolan et al., 2008; Harreither et al., 2015
CRELD2	<ul style="list-style-type: none"> Diverse range of potential functions Role in regulating ER stress response, downstream of ATF6 Role in protein folding, interacts with chaperones, potential PDI function 	<ul style="list-style-type: none"> Upregulated in plasma cells, +4.9-fold (RNAseq) and +3.6-fold (Affymetrix) 	CHO	100689097	<i>Q6UXH1</i> <u>Q9CYA0</u>	Hartley et al., 2013; Ohashi et al., 2009

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: Human/ Mouse	References
CypB/ PPIB/ Cyclophilin B	<ul style="list-style-type: none"> ER PPIase, role in catalysing folding of proline-containing proteins 	<ul style="list-style-type: none"> Previous overexpression studies, resulting in up to 1.4-fold improved productivity levels of DTE proteins in CHO cells Upregulated in plasma cells +1.3 (RNAseq) and +1.3 (Affymetrix) Associated with increased CHO production capacity 	Human	5479	<u>P23284</u> <u>P24369</u>	Johari et al., 2015; Nissom et al., 2006; Pybus et al., 2013
Dad1/ OST2	<ul style="list-style-type: none"> Essential component of the N-oligosaccharyl transferase (OST) complex Required for efficient glycosylation within ER Presence may protect against apoptosis 	<ul style="list-style-type: none"> Upregulated in plasma cells, +1.2-fold (Affymetrix) Associated with increased CHO cell production capacity 	CHO	100767387	<u>P61803</u> <u>P61804</u>	Harreither et al., 2015
ERo1Lα/ ERO1A	<ul style="list-style-type: none"> Oxidoreductase- reoxidises PDI enzymes Role in disulphide bond formation in assembling proteins in ER 	<ul style="list-style-type: none"> Previous overexpression studies resulting in up to ~1.4-fold increase in antibody productivity levels in CHO cells, further improved when co-expressed with PDI/XBP1s Associated with increased CHO cell production capacity Upregulated in plasma cells +1.3 (RNAseq) and +3.1 (Affymetrix) 	Human	30001	<u>Q96HE7</u> <u>Q8R180</u>	Cain et al., 2013; Harreither et al., 2015; Mohan and Lee, 2010
FKBP11/ FK506 binding protein 11/ FKBP19	<ul style="list-style-type: none"> ER PPIase, role in catalysing folding of proline-containing proteins 	<ul style="list-style-type: none"> Upregulated in plasma cells, +4.2-fold (RNAseq) and +5.7-fold (Affymetrix) 	Human	51303	<u>Q9NYL4</u> <u>Q9D1M7</u>	
Grp94/ Hsp90b1/ Endoplasmic	<ul style="list-style-type: none"> ATPase activity Functions in ERAD Role in folding and stabilising other proteins Functions in the processing and transport of secreted proteins 	<ul style="list-style-type: none"> Upregulated in plasma cells +4.2 (RNAseq) and +2.7 (Affymetrix) Associated with increased CHO production capacity 	Human	7184	<u>P14625</u> <u>P08113</u>	Harreither et al., 2015

Gene Name(s)	• Function	• Rationale	Sequence	NCBI Gene ID	UniProt IDs: Human/ Mouse	References
HSPA1a/ Hsp701	<ul style="list-style-type: none"> ER molecular chaperone ATPase Associated with folding of proteins, quality control in ER Stabilises proteins, against aggregation Anti-apoptotic under ER stress 	<ul style="list-style-type: none"> Literature (function) Improved productivity of Factor VIII in BHK-21 cells by up to 50% Transcripts upregulated in producing CHO and MAK cells (omics evidence) 	CHO	100760510	<i>P0DMV8</i> <i>Q61696</i>	Ishaque et al., 2007; Yee et al., 2009
HYOU1/ Grp170/ ORP150	<ul style="list-style-type: none"> ER molecular chaperone role, associated with protein folding and secretion Cytoprotective role, reducing hypoxia-induced apoptosis 	<ul style="list-style-type: none"> Upregulated in plasma cells +2.9 (RNAseq) and +1.9 (Affymetrix) Multiple associations with increased CHO production capacity 	Human	10525	<i>Q9Y4L1</i> <i>Q9JKR6</i>	Carlage et al., 2012; Doolan et al., 2008; Harreither et al., 2015; Nissom et al., 2006
PDI/ P4HB	<ul style="list-style-type: none"> Catalyses the formation and rearrangement of disulphide bonds within protein assembly within the ER Role in inhibiting aggregation of misfolded proteins 	<ul style="list-style-type: none"> Associated with increased CHO production capacity Upregulated in plasma cells +2.8 (RNAseq) and +1.5 (Affymetrix) Previous overexpression studies resulting in <1.4-fold increase in antibody productivity levels in CHO cells 	Human	5034	<i>P07237</i> <i>P09103</i>	Borth et al., 2005; Carlage et al., 2012; Harreither et al., 2015; Johari et al., 2015; Mohan and Lee, 2010
PDIA4/ ERp72/ ERp70	<ul style="list-style-type: none"> Member of the ER PDI family of enzymes Catalyses disulphide bond formation, part of protein folding 	<ul style="list-style-type: none"> Upregulated in plasma cells, +3.2-fold (RNAseq) and +3.6-fold (Affymetrix) Associated with increased CHO cell production capacity 	Human	9601	<i>P13667</i> <i>P08003</i>	Carlage et al., 2012; Doolan et al., 2008; Harreither et al., 2015; Yee et al., 2009
PRDX4/ PRX4	<ul style="list-style-type: none"> Antioxidant enzyme (peroxidase), reduces hydrogen peroxide PDI oxidation enzyme (works in cooperation with Ero1Lα) Role in activating transcription factor NF-kappaB 	<ul style="list-style-type: none"> Upregulated in plasma cells, +5.1-fold (RNAseq) and +3.5-fold (Affymetrix) Literature (function) 	Human	10549	<i>Q13162</i> <i>Q08807</i>	Sato et al., 2013
SRP14	<ul style="list-style-type: none"> Nascent protein transfer into ER Part of the signal recognition particle which targets secretory proteins into the ER Role in elongation arrest, binds to RNA 	<ul style="list-style-type: none"> Previous overexpression improved productivity of DTE IgG by 7-fold increase and ETE IgG by 30% in CHO cells Associated with increased CHO cell production capacity 	CHO	100761457	<i>P37108</i> <i>P16254</i>	Le Fourn et al., 2014; Harreither et al., 2015

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/</i> <i>Mouse</i>	References
Tor1a/ Torsin-1A	<ul style="list-style-type: none"> • ATPase • ER molecular chaperone • Heat-shock protein-like • Role in quality control of protein folding, processing and stability in ER • Clearance of misfolded proteins, reducing aggregation ('disaggregase') 	<ul style="list-style-type: none"> • Previous overexpression improved productivity of secreted luciferase in CHO cells by up to 2.5-fold and IgG4 up to 1.3-fold 	Human	1861	<u><i>O14656</i></u> <u><i>Q9ER39</i></u>	Jossé et al., 2010

Appendix Table C.2: Effector Gene Library Information: Golgi Complex/Vesicular Secretion Processes

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/</i> <i>Mouse</i>	References
Bet1	<ul style="list-style-type: none"> • Golgi-associated SNARE protein, role in ER-derived vesicle docking on Golgi complex membranes • Facilitates fusion of COPII vesicles to each other 	<ul style="list-style-type: none"> • Upregulated in plasma cells, +2.8-fold (RNAseq) and +2.5-fold (Affymetrix) • Associated with increased CHO cell production capacity • 1.58-fold increase in transcripts in a producing CHO cell line • Literature (function) 	CHO	100761331	<u><i>O15155</i></u> <u><i>Q35623</i></u>	Harreither et al., 2015
Bet3/ TRAPPC3	<ul style="list-style-type: none"> • Early stage secretory pathway (ER to Golgi complex) • A subunit of the TRAPPI tether complex on the Golgi complex, associated with tethering of COPII vesicles 	<ul style="list-style-type: none"> • Literature due to its function 	CHO	100758455	<u><i>O43617</i></u> <u><i>Q55013</i></u>	
CERT/ Col4a3bp (active, de-phosphorylated version – CERT S132A)	<ul style="list-style-type: none"> • Golgi complex lipid synthesis • Traffics ceramide, a precursor for lipid components diacylglycerol and sphingomyelin, from ER to the Golgi complex 	<ul style="list-style-type: none"> • Previous overexpression improved productivity of IgG in CHO cells by up to 26% and t-PA up to 35% in two separate studies 	CHO	100689038	<u><i>Q9Y5P4</i></u> <u><i>Q9EQG9</i></u>	Florin et al., 2009; Rahimpour et al., 2013

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/Mouse</i>	References
ERGIC53/ LMAN1	<ul style="list-style-type: none"> Transmembrane mannose-specific lectin, recognises glycoproteins Forms a specific cargo-receptor complex 1:1 with MCFD2 Role in transport, sorting and recycling of proteins from ER to Golgi complex via ERGIC 	<ul style="list-style-type: none"> Upregulated in plasma cells, +4.3-fold (RNAseq) and +6.3-fold (Affymetrix) Associated with increased CHO cell production capacity 	CHO	100754529	<i>P49257</i> <i>Q9D0F3</i>	Harreither et al., 2015
MCFD2	<ul style="list-style-type: none"> Luminal protein Forms a specific cargo-receptor complex 1:1 with ERGIC53 Role in transport of proteins from ER to Golgi complex via ERGIC 	<ul style="list-style-type: none"> Upregulated in plasma cells, +2.35-fold (RNAseq) and +2.1-fold (Affymetrix) 	CHO	100764995	<i>Q8NI22</i> <i>Q8K5B2</i>	
p115/ USO1/ TAP/ General vesicular transport factor	<ul style="list-style-type: none"> Tether protein, required for vesicle transport within the Golgi complex, associated with SNARE complexes Role in initial COPII vesicle docking/fusion processes between Golgi complex membranes 	<ul style="list-style-type: none"> Upregulated in plasma cells, +2.6-fold (RNAseq) and +1.8-fold (Affymetrix) Literature (function) 	CHO	100756021	<i>O60763</i> <i>Q9W3N6</i>	
PREB/ Sec12	<ul style="list-style-type: none"> A GEF (Guanine nucleotide exchange factor) Recruitment role in COPII transport/export from the ER 	<ul style="list-style-type: none"> Upregulated in plasma cells, +1.9-fold (RNAseq) Associated with increased CHO cell production capacity 	CHO	100764138	<i>Q9HCU5</i> <i>Q9WUQ2</i>	Harreither et al., 2015
PRKD1	<ul style="list-style-type: none"> Serine/threonine-protein kinase Regulates fission (membrane budding) of vesicles from the Golgi complex destined for the plasma membrane 	<ul style="list-style-type: none"> Literature (function) Upregulated later on in CHO cell culture (production phase) 	CHO	100762119	<i>Q15139</i> <i>Q62101</i>	
Rab1a	<ul style="list-style-type: none"> Small GTPase Essential for vesicle tethering/binding at Golgi complex membranes Mediates recruitment of p115 tether to COPII vesicle 	<ul style="list-style-type: none"> Upregulated in plasma cells, +1.8-fold (RNAseq) 	CHO	100752891	<i>P62820</i> <i>P62821</i>	
Rab11a	<ul style="list-style-type: none"> Small GTPase Key regulator in intracellular membrane trafficking (post-Golgi complex vesicular trafficking) Involved in vesicle generation, endosome recycling and endo- and exocytosis 	<ul style="list-style-type: none"> Literature (function of interest, role in vesicle trafficking of secretory proteins) 	CHO	100767114	<i>P62491</i> <i>P62492</i>	

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/Mouse</i>	References
SGMS1/ sphingomyelin synthase 1	<ul style="list-style-type: none"> • Role in lipid synthesis at the Golgi complex membrane 	<ul style="list-style-type: none"> • Literature (function) 	CHO	100766719	Q86VZ5 Q8VCQ6	
Stx5a/ Syntaxin 5	<ul style="list-style-type: none"> • A t-SNARE protein, role in COPII vesicle docking/fusion between ER and Golgi complex membranes • Facilitates fusion of COPII vesicles to each other 	<ul style="list-style-type: none"> • Literature (function) • Upregulated in plasma cells, +1.6-fold (RNAseq) • Associated with increased CHO cell production capacity 	CHO	100756021	Q13190 Q8K1E0	Harreither et al., 2015

Appendix Table C.3: Effector Gene Library Information: Other

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: <i>Human/Mouse</i>	References
Largen/ PRR16	<ul style="list-style-type: none"> • Regulatory element • Regulator of cell size • Enhances specific mRNA translation 	<ul style="list-style-type: none"> • Function: overexpression resulted in increased cell size 	Human	51334	Q569H4 A3KMN5	Yamamoto et al., 2014

Appendix Table C.4: Effector Gene Library Information: Transcription Factors

Gene Name(s)	Function	Rationale	Sequence	NCBI Gene ID	UniProt IDs: Human /Mouse	References
ATF6αc (cleaved version, cytoplasmic domain)	<ul style="list-style-type: none"> Major role within UPR, alters transcription of many ER components Downstream effects include chaperones, XBP1s, CHOP, lipid synthesis, proteolysis and ER expansion Widespread effector 	<ul style="list-style-type: none"> Previous overexpression studies, resulting in up to 1.5-fold improved productivity levels of DTE proteins in CHO cells Upregulated in plasma cells +2.3 (RNAseq) and +3.0 (Affymetrix) Associated with increased CHO cell production capacity 	Human	22926	<i>P18850</i> <u>F6VAN0</u>	Harreither et al., 2015; Johari et al., 2015; Pybus et al., 2013
CREB3L2/BBF2H7 (cleaved version, cytoplasmic domain)	<ul style="list-style-type: none"> ER localised transcriptional activator, functions under ER stress, part of UPR Major direct regulator of secretory capacity, upregulate many secretory protein machinery components Potential for widespread effects similar to XBP1s 	<ul style="list-style-type: none"> Direct literature (function) Upregulated in plasma cells, +6.3-fold (RNAseq) and +6.1-fold (affymetrix) 	CHO	100771872	<i>Q70SY</i> <u>Q8BH52</u>	Fox et al., 2010
TFE3 (active form: dephosphorylated version: TFE3 S108A)	<ul style="list-style-type: none"> Transcription factor, activated by Golgi complex stress response Dephosphorylated TFE3 binds the Golgi complex stress response element (GASE) which up-regulates Golgi complex structural proteins, glycosylation enzymes and vesicular transport elements 	<ul style="list-style-type: none"> Literature (function of interest, increases capacity of Golgi complex) 	CHO	100761176	<i>P19532</i> <u>Q64092</u>	Sasaki and Yoshida, 2015; Taniguchi et al., 2015
XBP1s	<ul style="list-style-type: none"> Part of UPR, working downstream of IRE1/ATF6, role in increasing protein folding and secretion within ER Increases the capacity of the ER, upregulates transcription of ER components Widespread effector 	<ul style="list-style-type: none"> Multiple previous overexpression studies resulting in improved productivity levels (Table 3.3) Upregulated in plasma cells +5.2 (RNAseq) and +2.9 (Affymetrix) Associated with increased CHO cell production capacity 	Human & CHO	7494	<i>P17861</i> <u>Q35426</u>	Cain et al., 2013; Harreither et al., 2015; Ku et al., 2007; Tigges and Fussenegger, 2006
YY1	<ul style="list-style-type: none"> Multifunctional transcription factor Regulates up and down transcription of many genes through epigenetic changes 	<ul style="list-style-type: none"> Previous overexpression improved productivity of IgG up to 6-fold in CHO cells 	CHO	100767279	<i>P25490</i> <u>Q00899</u>	Tastanova et al., 2016

Appendix Table C.5: Effector Gene Library Nucleotide Sequences

(C) denotes CHO protein and (H) denotes human protein

ATF6α c (H)	atgggagaacctgctggcgtggccggcaccatggaagccctttagccctggcctgttccaccggctggacgaggattgggactctgcccgtttgcccagctgggctactccaccgacaccgacgaactccagctggaagccgccaacgagacatacagagaacaactcgaacaactggacttcgacctggacctgatgccctggagctccgacatctggacatcaacaaccagatctgcaccgtgaaggacatcaagggccgagccccagcctctgcccctgcccagctctactccgtgtcctccccagatccgtggactcctactcctccaccagcagctgcccggaggaactggacctgtcctccagcagccagatgtccccctgtctctgtacggcgagaactccaactccctgtccagcgcggaaccctctgaaagaggacaagcctgtgaccggccctcggaacaagaccgagaatggcctgacccccaaagaaaagatccaagtgaactccaagccctccatccagcccaagccccctgctgctgctgctcctaagaccagaccaactcctccgtgcccggcaagaccatcatcatccagaccgtgccaccctgatgccctggtaagcagcagcccatcatcctccagcctgccctaccaagggccagacagtgctgtctcagcccaccgtggcagctccaggtcctggcgtgctgcatctgctcagcctgtgctgctgtgctggctggcagtgaccagctgctaatcacgtcgtgaacgtggcctgcccctccgccaactcctctgtgaaacggcaagctgtccgtgaccaagccagtgctccagtcaccatgcggaacgtggctccgatccgctgctgagaaggcagcagagaatgatcaagaacagagagtcggcctgcccggaaaaaagaaaagagtagatgctggcctggaagccggctgaaggccgcctgtctgagaacgagcagctgaagaaaagagaaacggcaccctgaagcggcagctgtaggtgtgtcggagaccagagactgaaggtgcccagccccagcggagatga
Bet1 (C)	atgcgtagagctggactggagaaggcggccctcctggcaattacggcaattatggctacgccggctctggctacaatgcctgcgaggaagagaacgacagactgaccgagagcctgcggtctaaagtaccgcatcaagcctgtccatcgagatcgccacgaagtgaaagaccagaacaagctgtgcccggagatggactcccagctcagacttaccaccggcttctggycaagaccatgggacagactgagatcctgtccagaggtcccagacaaagctgtgtgctacatgatgctgttctccctgtctgttctctgtgactactggatcatcaagctgcgctga
Bet3 (C)	atgtccagacaggctaacagagaccagctccaagaagatgtcctccagctgttaccctgacatggtcgtctggttaccagctgtgcaaggactacgagaacgacgagagctgaacaagcagctggaccgtagggctataatcgccgtgcccgtgatcgaaggactcctgtagatctaacgtgggagatgccacgactcagagaaaccgcccagatgtgatcggaaaggtggcctcaagatgacctgggcatcacccctccatccaattggagctcctgcccggcagcagcttccctgatcctggaaacaaacctctggtggacttcgtggaactgcccagctcactccagcctgatctacttaacctgctgtgtggcgtgctgagaggcgcactggaatggtgagatggcctggaagcgaatctgtgacggatacactgaaaggcagcggcgtgaccgagatccggatgcccgttctgcccagagaatcgagataacctgcccagcggcggaggatga
BiP (H)	atgaagctgctctggtggccgcatgctcctgctgctgctgcccgtagagccgaggaagaggacaagaagaggacgtgggcaacgctggtggcactcagctgggaccactcttgcctgctggcgtgtcaagaacggcgggtggaatcattgccaacgaccagggcaaccggatcacccctctactgctggccttacccccagggcgagagactgatcggcgacggcctagaagaaccagctgacctccaaccccagagaacacgctgtcagcgaagcggctgatcggcagaacctggaacgaccccctcctgagcagcagatacaagtctctgcccctcaaggtggtggaaaaaagaccaagccctacatccaggtggacatcggcggaggccagaccaagaacctcgcgccgaagagatctccgcatggtgctgaccaagatgaaggaaaccgcccagggcctaccctgggcaagaaagtgaccacgcccctgtgaccgtgcccgcctactcaacgacgctcagagacagccaccaaggacggcgacaatcgcccgtgaacgtgatgctgatcatcaacgagcccaccgctgctatcgctaccgctggataagcgggaaggcagagaagaacatcctggttctgatctggcggagggcaactcagctgctcctgctgaccatcgcaaacggcgtgttcaggtggtggtaccaacggcagataccacctgtggcggaggacttcgaccagagagtgatggaacacttcatcaagctgtacaagaagaaaaaccggcaaggacgtgcccgaaggacaaccggccgctgcagaaactcggagagaggtggaaaaggccaagcggccctgtcctctcagcaccagggcagaatcgagatcgagctctatgagggcagagattctccgagacactgaccgggccaagttcgaggaactgaacatggacctgtccggtccacaatgaagcctgtgcagaaggtgctggaagattccgacctgaagaagtcggacatcgacgagatcgtgctcgtggcggctccaagaatccccaaagatccagcagctcgtgaaagagttctcaacggcaagagaccctcccggggcatcaacctgatgagctgtggcttacggcgtgctgcaaggctggcgtgctgctggcagatcaccggcgatctggtgctgctggaagctgtcctctgacccttggaatcagacagtgggcggcgtgatgaccaagctgatcccagaacacggctggtgctcaagaagtgccagatctctcacccgctccgacaaccacgctaccgtgaccatcaaggtgtacgagggcgacggccctgaccaagaagatcaacctctgctgggcaaccttgacctgaccggcatccctcctgctcctagagggctgccacagatcgaagtgaccttgagatcagctgaacggcatcctgagagtgaccgctgaggacaagggcaccggcaacaagaacaagatcaccatcaccaacgatcagaacaggctgacacctgaaagagatcgagcggatggtcaacgatccgagaagttcggcaagaagataagaagctgaaagagcggatcgacaccggcaagagctggaatcctacgctactccctgaagaaccagatcggagacaaagagaagctggcggcaagctgtccagcagggacaagaacacatggaaaaagcctggaagagaagatcgagtggtggaatctaccaggacgctgacatcgaggacttaaggctaagaagaaagaactggaagaatcgtgcagcccacatctaaagctgtacggctcgtccggccctccacctaccggcagggagataccgcccgagaaggacgagctgga
Calr (H)	atgctgctgtctgttctcctgctgctgctggactgctggactggctgttctgctgagcccggctgtacttcaaagagcagttttggacggcgaacggctggacctccagatggatcgagctaaagcacaagctccgactcggcaagttcgtgctcctctggcaagttctacggcgacgagaaagggacaagggcctccagacctcagagcggcagattctacggcctgtctgctcttctgagccctcagcaacaagggccagacctggtgtgcagttaccctgaaagcagcagacatcagctgtggcggcgatcgtgaagctgttccctaactctctgaccagaccgacatgacggcactccagatcaacatcatgttccggcctgacatctgcccctggcaccagaagaagtgacgtgatcttcaactacaagggcaagaacgtgctgatacaaaaggacatccgggtgcaaggacgacaggtcaccacctgtataacctgatcgtgcccggccagacaacacctaagtgaaagatcgacaactccaggtggaaatccggcagcctggaagatgactgggactcctgctcctaagaagatcaaggaccccagcctcctaagcctgaggatggatgagagagccaagatcgacaccccagactcacaacagagactgggacaagcccagacatccctgatcctgatgccccaaagcccgaagattgggacgaagagatgatggcagtgggagcctcctgtgatccagaatcctgagtacaaaggcgaatggaagcccggcagatcgataaccccgactacaagcactggttacccccagattgacaaccccagactcctccgatccttcatctacgctcagacaactttggcgtgctggcctgatctgtgcaagtgaaagctccggaacatctctgataacttctgataccaacgacgagggcctacggccaagagttcggca

	atgagacatggggcgtgaccaaggccgctgagaagcagatgaaggacaagcaggatgaggaacagcggctgaagaggaa gaagaggacaagaagcgcaagaagaggaagagccgaggacaagaggacgacgaagataaggacgaggatgaaga agatgaagaggcaaaagaaggagcagggaaagggacgtccaggccaggctaaggatgagctgtaa
CANX (H)	atggaaggcaagtggctgctgtatgctgctggtgctggccacagctatcgtggaagctcacgacggccacgacgacgctgat cgacatcgaggacgacctggatgatgtatcgaggaagtgaagatagaagccgacaccacgctcctcatctagcccaa agtgacctacaaggctcctgtgctaccggcaggtgactctgcccattcttcgacagaggcaccctgtccggctgattctgtcta aggccaagaaggacgacacagacgacgagatcgtaagtagcagcgaagtgggaagtcgaggaaatgaaggaatccaag ctgcccggcgacaaaaggcctggtgctgatgtctagagccaagcaccacgcatctccgccaagctgaacaagcctcctgttoga caccaagcctgatcgtgacgtacgaagtgaactccagaacggcatcgagtggcggcgcttacgtgaagctgctctaaagac ccctgagctgaacctggaacctccactacaagacccttacaccatcatgttcggcggcaggactacaag ctgcactctatctccggcacaagaacccccagaccggcatctacgaggaaaaagcagccaagaggcctgacgcccagcctgaa aacctactcaccgacaagaaaaccacctgtacacctgatcctgaatcctgacaactcctcgagatcctgtggaccagctcgt ggtcaactctggcaacctgctgaacgatgaccctcctgtaacccctccagagagatcgaggaccctgaggacagaagcct gaggactgggacgagcggcctaagattcctgatccagagccctggaagcccagcactgggatgaagatgctcccgtaaagatc ccgacgaggaagctacaaagcctgaaggtggtggacgacgagcctgagatgtgccagatcctgatgcccgagaagccaga agattgggacgaagatagatggcagtgagggaagccctcagatcgcaatcctagatcgagctgctcctggatgtggcgtgtg ccagaggcctgtgatcgataacccccactacaaggcaagcctcctatgatcgagcctcctgacacccctcctgacacag aagccccggaagatccctaactcgtactctcgaggacgctgaacccctccggatgaccccttttccgcatcgccctggaactgtg gtctatgacctccgacatctctcagacaactcatctatcggccgaccggcggatcgtggatgattgggctaagatggctggggcct gaagaaagctgctgatgtgctgagctgaaccaggcgtcgtgggacagatgattgaggccggaagaaagacctgtgctgggt cgtgtacatcctgacagtgccctgctgttctcctggtcatcctgtctgctcctccgcaagaaacagacctccggcatggagta aaaagaccgacgctcccagcctgacgtgaaagaggaagaagaggaaaaagaagaagagaaggacaaggccgacgaag aggaagaggccgaagagaagctggaagaaaagcagaagtcgacgcccgaagaggatggcggcacagtgctccaagaaga agaagatcggaagcccaaggccgaggaagatgagatcctgaacagatccctcggaaacccctcgagagaatga
CERT (C)	atgtctgacaaccagctcctggaactcctccgctccgaggaagatcctgagacagagctcggcctcctgtggaagatcgcgct gctgtctaagtggaccaactacatccacggctggcaggacagatgggtcgtgctgaagaacaacacccctgctactacaagtccg aggacgagactgagtagcggctgcccgggatctatctgtctccaaggctgtgatcacccctcagacttcgacgagtgagattcg acatcctcgtgacgactcctgctgtgtatctgctgctcagaccctgatcacagacagcagtgatgacccatcgagcagcat aagaccgagctccggtatggctctgagctcctttgagaaggcagggcgtatgggtctctgtgtctgcccgtcttggtactctgcc acctccaccttagctcaagaaggccacagcctgagagaaaagctggccgagatggaaacctccgggacatcctgtgtagac aggtggacacccctgcaaaagtacttcgacgtgtgtgcccagcctgctcaaggacgaattgagagagacaaggtcgtcgagg acgacgagatgactcctaccacagatctgacggcagctcctgcaacaacacccaagcgaagaagctgttccctca cgtgaccctaaaggcatcaacggcatgattcaaggcgaagccatcacctttaaggccaccaccgctggcatcctggctaccc tgtctactgtatcgagctgatggtcaagcgggaagagctctggcagaagcggcagcacaagaaatggaaaagcggcggagg ctggaagaggcctacaagaatgctatggccgagctgaagaagaagcccagattggcggcctgactacgaggaaggcccaa ctctctgatcaacgaggaagatctttgacgcccgtggaagcccctctggacaggcagataagatcgaggaacagctcccagctc gagaaagtgcggctgattggcctacacctctgctcagcggcagccttttctcctggtggcaccacagattcgtgcagaaagtgg aaagatggtgcagaaccacatgacctacagcctgcaagatgctggaggcagcgttaattggcagctggtggtgaaagggcg agatgaaggtgacccgagagaggtggaagaaaacggatcgtgctggaccctctgagaccctctgagaccctctgagaccctcagag ccggccacgaagtgcaactactttggagcgtggactgcccgaacgactgggagacaacctcagagaactccacgtggtgga aacctggccgacaacgccatcatctatcagaccacaagagagtggtggcccagccagagagatgctgtactctgctc gcatccggaagatccctgctgaccgagaatgaccccagacatggatcgtgtgtaactctccctggaccacgacagcgcctc tctgaacaatagatgtgctggccaagatcaacgtggccatgatctgcagaccctggtctcctccagaggcaatcaagagat ctcccgggacaatatcctgtgcaagatcacctacgtggccaacgtgaacctggtggatgggctcctgctctgtgctgagagctgtg ccaagagagagatccccaagtctgaagcggctcacctcctacgtgcaagaaaagaccggcggcaagccatcctgttctag
CREB3L2 (C)	atggaagtggaaacctctcaacctctcctcctcctgacagcagcactctactccctgtgaggaacctcggctcagtc ccctttacacacgtggccacctctgacggctcaacgacgaagaggtggaatccgaaagtgtatctgctccctgactcctccctc gccaccgtgaaaacagagcctatcacagaggaacagccacctggcctggtgcatctgtgacctgacctacccgcatcagc accccttcgagaagaggaatccctctggacatgaatgcccggcgtgtcggactccagcagctatctacctcctggtatcctgcta ctcctccggcatggatggcatgacatgctcctctccgctcctggtgaccagctgattgcctccaacacctccatcctctcac tctccgattccgaggcagcctgtctcctaactcctggctgacacctttagcctgagccaggctcatagccctgcccagagctatgct agaggccctccgctctgtactcctcctctgctgaccgctcctcaagctgcaagatctggacctctggtgctgaccgaggaag agaagagaacctgatcggcagggctacccattctacaagctgcctctgaccaagtcggaggaaaagggcctggaagaag atccggcggaaagatcaagaacaagatcagcggccaagagtcggcgggaaaaagaagagatcatgactccctggaaag aaggtggaagctgctccaccgagaacctggaactcgggaagaaagtggaaagtctggaacacacaagaacctgctcc agcagctgcaaaagctccaggctctggtcatggcgaagtgtccaggctctgcaagctggctggcaccagaccggaacataa
CRELD2 (C)	atgcatttgctccagctgctgctgtgactgctgtgtgtgtgctcctccagccagatggcctccagaagcctaccatgtgcc agagatgcagagccctggtggacaagtcaatcaaggatggccaacaccgccaagaaactcggcggaggaataaccgccc tggaagagaagtccctgtctaagtagatctccgagatccggctgctggaaatcatggaaggcctgtgactccaacgacttc gagtgtaaccagctgttgagcagcagaggaacagttggaggctggtggcagaccctgaagaaagagtgcccaacctgttcg agtgttctcgtgcacaccctgaaggtgtgtgctcctggcaactcaggaacctgactgtcaagagtgcaagcggcctcagag gctgttctggaatggacattgcgacggcagcggctctagacaaggcagcggatctgtcagtgccacgtgggtataaggccc tctgtcatcagctgatggacggctactctccctgctcggaaacgagactattcctctgcaaccgctcggacgagctcctgaaga cctgttctggccctaccaacaaggctgctggaatgtgaagtcggcctggaccagagtggaagatgctgtgtgagatgtggacgagt gtgcccgtgagacacctcctgctcaacgtgcaatgtcagagaaactggaacggctcctacacctgtgaagagtgcgactctacct gtgtggctgtaccggcaaggccctgccaactgcaaaagagtgatctccggctactccaagcagaagggcagtgccggacat cgacgagtgctccctgagacaaaagtgtgcaagaaagagaacgagaactgtacaacacccctggcagctcctgtgctgctgt

	cctgagggcttgaagaggacagacgggtgctgtgaccgactccagaagaagatctggcagaggcaagtcccacaccgccaactgccttaa
CypB (H)	atgtctcgggtgctgtgagcggaaacatgaagggtgctgtgcccgtgctgtgatcgccggctgtgttcttctgctgtgctgcccctctgcccgcgacgagaagaaaaagggccccaaagtgaccgtgaagggtactctgacctgcgatcggcgacgaggacgtgggcagagtgatcttccgctgttccgcaagaccgtgcctaagaccgtggacaactctggtccctggccaccggcgagaagggctcggctacaagaactccaagttccatcgagtgatcaaggactcatgatccaggcgggcactcaccagaggcgacggaacaggcggcaagtccatctacggcgagcggttccctgacgagaactcaagtgaaacactacggccctggctgggtgctcatggccaatgccggcaaggacaccaacggctcccagttctcatcaccaccgtgaaaaccgctggctggacggcaagcagtggtgttggcaagggtctggaaggcatggaagtcgtcggaaggtggaatccaccaagaccgactccgggacaagcccctgaaggacgtgatcattgccgactcggcaagatcgagggtggaagcccttcgctatcgccaaagagtga
Dad1 (C)	atgtctgcatccgtggtgctcgtgatctcccggttctggaagagtacctgagcagcaccctcagagactgaagctgctggatgacctgctgtacattctgctgaccggcgctctccagttcggctattgtctgctgtggcacccttcttcaattcttctgctcggctttatctcctgctgggtctttctatctggtgctgctgagaatccagatcaaccctcagaacaaggccgactccagggatctctctgtagagactctcgcgattctgcttaccatctgacactggtggtcatgaactctggtggctga
ERGIC53 (C)	atgctggtctagaagaagggaaacacaggtggcgctggcctttatctgtgccctgctgctgctcctcagcagattcgtgggctctgatggcatgggcgagatgctgctgctcctgggtgctgcttactctctgctgagctgccccacagaagattcagatcaagtaactcctcaagggccctacctgggtcagtgatggaaccgtgcttttggcccacgcccgaatgctatccctccagcagatcagatcagatcgccctagctgagtcagagggctgtgtggaccacaagccgctcagagaactgggaactggaagtgacactcagagtgaccggcaggaatcggagctgagctgtgtgacaccgagaatcggactcggcctgtgttccgactccttcgactccttcgacaacgacggcaagaagaacaaccccgcctcgtgatc atcggaacaacggccagatcaactacgaccaccagaacgatggcacaaccaggctctggcctctgcccagagggactccggaacaaaccttactctgctgcccgaagatcacctactaccagaaaaaccctgaccgtgatgatacaaatggctcaccctgacaa gaacgactcagagttctcgcgaaggtcgagaacatgatcatcctactcagggccactcggcatctctgctgactggtgggactggctgacgaccacgatgtgctgtcttctgacttccagctgaccgagcctggcaagagcctcctactggaagagacatctccgagaagagaagaaagaaagttcagaaggacaccccgcactccagggccagcctgctgagtgatggttccgactccagcaagagctggacaagaagaagaatcagaaggagaacagaatccactggaatcaagcagctgaaccggcagctggacatgatcctggacgagcagagaagatacgtgtcctcttgaccgaggaaatctctggagagggcgtggaactcctggacagcctggacaagtgtcccagcaagaactcgacaccgtgtca agaccagcagagatcctcggcaagtgaaacgagatgaaagaactcctatgctccgagacagtgccggtggtgctggtatccagc atcctggatctgctggtgacgagacaaccagcactcatggacatcaagaacaacctccacgtcgtgaagcgggacatcga caatctgcccagcggaaatcgccctcaacgagaagcctaagtgctccgagctgctcatttctagctgctgtctaccatccatctgctatctctggtggtgagaccgtgctgttaccggtacatcatgtaccggacacagaagaggtcggccaagaagttctctag
ERo1Lα (H)	atgggaagaggctgggcttctgttggcctgctgggagctgtgtggctgctgctccttggccatggcgaggaaacagcctccagagacagccgcccagagatgcttctgtcaggtgtccggctacctggacgactgcactgtgacgtggaaacctcagacaggttcaacaactacgggtgtcccccgctccagaagctgtggaatccgactactcgggtactacaagtgaaactgaagcggcctgcccctctggaaacgatcctccagtgccgagacagcgggactgcccgtgaagccttggcagctgatgaggtcccgcagcgaatcaagtcgcccctcaagtaactccaggaagcacaacctgatcgaggaatgcgagcagggccgagagactgggcccgtggatgagttccgttccgaagagacacagaagcggctgctccagtggaacagcagcagactcctccgacaactctcgaggccgacagatcagacccccagggcagtgatggtgacactgctgtaaccggagcggtaaccggctacaagggcctgatgctcgggaatctggaaactcagagaaaactgcttcaagccccagaccatcaagcggccccgaacctctggtctggtccaggacacagcgaagagaacaccttctacagctggctggaagcctggtgtggaagagcggcctttaccggctgatctccggctgcaagcctccatcaacgtcacctgtccgcccagatacctgctccaggaacatggtggaaaagaagtggggcccacaacatcaccgagttccagcagagattcagcggatcctgaccgagggcagggccctagaaggtgaaagaacctgtacttctgtatctgatcgagctgctggcctgtccaaggtgctgccattctcagcggcccggactccagctgttaccggcaacaagatccaggacgaagaaaacaagatgctgctgtgaaatctcgcacgagatcaagtccttcccctgactcagcagagaacagcttcttggcggcagaagaagagcccacaagctgaagaggacttccgctgacttccggaaactcctccggatcatggaactgctggtggtgttcaagtcggctgtgggcaagctgtgggcaagctccagcccagggactgggcaaccgctctgaaatctgttctccgagagaagctgatcggcgaacccggacccgcccctctacaggttccactgacccggcaggaatctgttccctgttcaacgccttccggcgatctccactcgtgaaagagctggaaaacttccgaacctgctccagaacatccactga
FKBP11 (H)	atgacactcagacctagcctgctgctcctgacactgtgtgctgtgtctctctgcccgtgtgtgtagagctgaggctggcctggaaa cagagttcctgtgcaacctccaggtgaaactctggtgaaactcctgagcctgtgcccgaacctgctgttggcgataacctg cacatccactataccggcagcctggtggacggccgattatcgataaccagcctgaccagagatcccctggtcatcgagctgggcca gaaacaagtgtccccggcctggaacagtcctgctggaatgtgtgtggcgagaagcggagagctatccttctcacttggc ctacggcaagagagaggttctccttctgtgctgctgacggcgtgtgagtagtggaaactgatcgcctgatcggcggccaacttggctgaagctgtcaagggaaactcgtcctctcgtcggcatggctatggtgctgctgctggtggcctgatcggctaccacctgtac agaaaggccaaccggcctaagggttccaagaagaagctgaaagaggaaaagcgaacaagctaaagaagaagtgga
Grp94 (H)	atgagagcttgggtgctgggctgtgtgctgctgaccttggctgtgctgcccggatgacgaagtgatgtggatggcac agtggaaagaggacctggcgaagtcagagaggctctagaacagacgacgagggtggtcagagagaggaaagggctattcag ctggacggcctgaacgctctcagatcagagagctgagagagaagtcggagaagttgcctccaagcgaagtgaaaccggatg atgaagctgatcatcaactcctgtacaagaacaagagatcttctgcccggagctgatctcaacgcctctgacgcccggacaa gatccggctgatctctcctgaccgacgagaacgcctgtctggcaatgaggaaactgaccgtgaagatcaagtcagacaagaga gaacctgctgcaactgaccgataaccggcgtgggcatgacaagagagggaaactgtaagaacctgggcaaacctcgaagtcgg gcacctcagatttctgaacaagatgaccaggctcaagaggacggcagctccacatctgagctgatcggacagtttggcgtgggcttctactcctcctggtggtgctgacaaagtatgctgacctcaagcacaacaacgacacccagcacatctgggagtcgactcc aacgagttcctgctgctcagagcctagaggcaataccctcggcagaggcaccacaatcaccctggtgctgaaagaagaggcct ccgactacctggaactggacacatcaagaatctctggaagaagtaactccagttcatcaactccccatctactggtgctcctcaa

	<p>gaccgagactgtggaagaacccatggaagaagaggaagccgccaagaagaggaagaaagaagagtcgacgacgaaagccgctcgcgaggaagaggaagaagagaagcctaagaccaagaaggtggaaaagaccgctctgggactgggagctgatgaacga catcaagcctatctggcagcggccctccaaagaggtcgaagagagcagatacaaggcctctacaagtcctcagcaaaagaag cgacgaccccatggcctacatccattcacagctgagggcgaagtgacctcaagtcctctgttctgctcctcctgctccaga ggctgtcgtgatgactgctcaagaagctgactacatcaagctgacgtgctggcggtgttcatcaccgacgactccacgac atgatgcccaagtacctgaactctgtaagggcgtggtggactccgacgacctgctgaaatgtgtctagagagacactccagcag cacaagctgtgaaagtgatccggaagaagctgctcagaagacccctcgacatgatcaagaagatcgccgacgacaagtaaca cgatacctctggaagaggtcggcaccataatcaagctggcgctgatcaggaccactccaacagaaccagactggctaagctg ctgctggctcagctctcaccatcctaccgacatcaccagcctgaccagctggaacggatgaaggaaagcaggataaga tctactctatggcggctccagccggaagagggcagaggttccattctggaagagctcctggaagagctcagaaagctac tacctgaccgagccagtgatgactgtatccaggctctgcccaggttcgacggcaagagattccagaacgtggccaaagaag gcgtgaagctgacgagctcggaaagaccaaagagagcagagagggcgttgagaagaggttgaccctctgtaattgatga aggacaagccctgaaggacaaaatcagaagggcgtggtgtccagagactgaccgaatctctgtgctggtggcctctcag tacggatgtcggcaacatggaagaatcatgaaggccagggcctaccagaccggcaaggacatctaccaactactacgcc agccagaagaaaacctcagatcaacctagacacctctgatccggacatgctgctggcgatcaagaggatgaggacga caagaccgtgctgacctggctgtggtgtgtttgagacagctaccctgagaagcggctcctgctgataccaagggcctacgg ccacagaatcgaagcggatgctgagactgtcctgaacatcgacctgacgccaaggtgaggaaccgagcaagcaacca gaggaaaccgctgaggataccaccgagataccgagcaggacgagacgagaagaatggacgtgggaccgagcaagaag aagaacagctaaagagctccaccgcccaggaaggtgagctgta</p>
<p>HSPA1a (C)</p>	<p>atggctaagtctaccgcatcgacctgggaccactattctgtgtggcggtgtccagcagggcaaggtggaatcattg ccaacgaccaggccaacagaaccacacctctacgtgacctaccgacaccgagagactgatcggcgagcggcgtaaaaac cagggtggccctgaatcctcagaacaccgtgttcgacgccaagcggctgatcggcagaagttggagatgccgtggtcaggccg acatgaagcactggccattcagggtgtaacgacgcgacaagcctaaggtccagggtgtcctaaggggcagacacgggcttt ctaccccgaagagatctcctcatgtgctgaccaagatgaaggaggtggcggaggtcctacctggccatctgtgaccaatgctg gatcaccgtgctgctcactcaacgactctcagagacggccaccaaggtgctggcgtgatcgtggactgaaactgctgagaa tcataacagagcccaccgctgctatcgctacggactggatgatccggcaagggcgaagaaatgtgctgactctgatctcg gcgaggcaccctcgatgtgctcctgaccatcgacgacggcatctcgaagtaaggtaccctggtgataccatctcggaa ggcgaggactcgacaacagactggtgtccactctggaagagttcaagcggaagcacaagaaggacatctccagaacaa gccggcggctgagagactgagaaccctgtgaaagagccaagcggaccctgtctagctctaccaggcctctctggaatcga atccctgttgaaaagccctgagagatgccaagctggacaagccagatccacgattggctgctgctggcggctcgaagaat ccctaagggtgcagaagctcaccagactctcaacggcggcctggaacaagagcatcaacctgatgagccctggtctatg tggctgctgctgagcggctgctatcctgatggcgacaagctgagaacgtgacggactgctgctgctggatggtctcctgtctctgg gactgagacagctggcggagtgatgaccgctctgatcaagaaacagcacaatcccccaagcagaccagacctcacc acctactcgacaatcagcccggctgctgattcagggtgacgtgctgagagagccatgaccagagacaacaacctgctgggaa gattcgagctgagcggatccctcctgctcctagaggcgtgccacagatcgaagtgacctcgacatcgatgccaaccgcatcctga acgtgaccgcccaccgataagctacaggcaaggccaacaagatcccatcaccacgataagggcggcctcgggaaagaaga gatcgagagaatggtcagaagggcagagatacaaggcggaggaagaggtgacaggggaagaggtgctgccaagaacg cctggaatcctacgctcacaatgaagctccgctggaagatgagggcctggaagggcctcagagctcagagggcctgagaaga aggtgctggacggctgcaagaagtgatctcctggctggatgctaacacctggcggacaaaagaagaggtctccacaagcggca agagttggagagagtgctggacatctgctgctgctgctcaaggtgacggcctcctggtgctggtgattggagcaaacgt cctaaagggcgtctggctgctggccctacaatcgaggaagtgactga</p>
<p>Hyou1 (H)</p>	<p>atggctgacaaagttcggagacagaggcctcggagaagagtggtgtggcctctgtgctgctgctgctgctgctgctgctgct gataccctggcgtgatgctgtggacctgggctctgaatccatgaaggtggcattgtgaagcccggcgtgcccattggaatcgtgc tgaacaagagctcggcgcaagaccctgtgatcgtgacctgaaagagaacgagcggctctcggcactgctccgctctatg gccataaagaatcctaaggctaccctgctcctcagctcctcagctgctggcagcagggcctgataatcctcagctgcccactgaccag gccagattctcgtgacgagctgaccttctgctccagcggcagaccgtgacttccagatctctagccagctcagctgaccctg aagaggtgctcggcatggtgctgaactactctagatccctggcggaggttccggcagcagcctatcaaggatgcccgtgatcaccg tgctgttctcaatcaagccgagagaagggcgtgctcaggtgctgagatggctgagctgaaggtgctcagctgatcaacg ataacaccgcccaccgctctgcttaccgctgcttcaagcgaaggaatcaacaccaccgctcagaacatgatgtttacgacatgg gctccggctcaccctgtgacacatcgtgacataccagatggtcaagaccaagagggcggcctgagcagcccaggtgacgatcag agcgtgggcttggatagaaccctcggcggcctggaatggaactgagactgagagaagactggcggcctgttcaacgagca gagaaaagccagagagccaagggacgtccgagaaatctcagagctatggccaagctgtagagagagccaacagactgaa aaccgtgctgctgccaacggcaccatggctcagatcagggactgatggacgactggactcaaggccaaggtgaccag agtggaaattgaggaactgtgcccagctgtcagagaggttctggacctgtcagcagggcctccagctgctgagatgctcctgg atgagatcgagcaagtgatcctgctgcccggagctaccagatgctcagagtgcaagaagtgctgtagggcctggtggcaag aggaactgggcaagaacatcaacggcagatgaggtgctgctatggcctgcttaccagccgctgctgctcgaagcctcaaa gtagccttctgctgctgcccagcggctggtgatacctctggtcaggttaccagagaggtggaagaggaacctggcatccaca gcctgaagcacaacaagcgggtgctgttctcccggatgggcccctatcctcagcggaaagtgatcacctcaaccgctactcccacg actcaactccacatcaactacggcagctggttctcggcctgagatctgagaggttctggcctcagagcctcagaccctgaccgg taaagctgaaagcgtgggagactcctcaagaagctaccggactcagagctcaagggcctcagagccactcaaccctgagag tctggcgtgctgagcctggatagagtgaaagcgtgttcgagacactggtggaagattccggcggaggaagagagcaccctgacca agctgggcaacacatcagttctgttccggcggagccacacctgacgctaaagagaatggcaccgacaccgtgcaagagg aagaagagctcctgcccagggctcaagatgagcctggcgaacaggtggaactgaaagaagaggtgagggcctgtggaa gatggtctcaacctcctccactgaacctaaagggcagcctacacctgagggcggagaagggccacagagaagaaacggcgc acaagagcagggcccagaagccatctgaaaagggcgaagctggaccagaaggtgctgctcctgctccagaagggcagaga agcagaagcctgcccagaagcggcggatggtcgaagagattggagtggaactggtggtgctgacactgctgctgcccagaag</p>

<p>PDI (H)</p>	<p>atgctgagaagggccctgctgtgctggccgtggctgctctctgctgagccgatgctcctgaggaagaggaccagctgctggtgctg cggaaagccaacttgcgagggctgctggccgcccacaagctgctggtgaggtctacgcccctggtgctgagccactgtaagacc ctggcccctgagtagcctaaggcctgctgaaagctgaaggcagggctctgagatccggctggtaaggtggacgccaaccgag gaatctgacctggcccagcagatggtgctgctgggctaccccacatcaagttctccggaacggcgacaccgctcccccaag agtataccgcccagagaagccgacgacatctgtaactggtgctgagaagagaaccggcccctgcccaccacactgctgctg ggtgctgctgctgagctccctggggaatctctgaggtggccctgctgctctcaaggaagctggaatccgactccgccaagcagt tctccaggccgcccagggctatcgacgatacccttggcctacacctcaactccgacgtgttctcaagatgactggacaagga cggcgtggtgcttcaagaagttcgacgagggccggaacaactcgagggcgaagtgaccaaagagaacctgctggactcatc aagcacaaccagctgcccctgctgactgagttcaccgagcagacccgcccgaagatctcgccgagagatcaagaccacatc ctgctgttctgcccagctgctgactacgacggcaagctgtccaactcaagaccgctgagctcctcaaggcaagatcc tgtctatctatcgactccgaccacaccgacaaccagcgcctctggaatcttctggcctgaagaaagaggaatgcccctgctg gctgacaccctggaagaggaaatgaccaagtacaagcccagctgctgaggaactgaccgagcggatcaccgagttctgcca cagattctggaaggcaagatcaagcctcacctgctgctccaggaactgcccagggactgggacaagcagcccgtgaaagtgtc gtgggcaagaacttgaggacgtgctcgcagagagaagaaacgtgtctggtgagtttctgctcctggtggtggcattgcaagc agctggccccatctgggataagctgggagacatacaaggaccagagaacatctgctgccaagatggactccaccgccc aacgaggtggaagccgtgaaggtgacagcttcccaccctgaagttttcccgcctctgcccagcggcgtgactgattacaac ggcgagagaacctggacggcttaagaagttcctggaatctgctggccaggaagcggcgtgctgacgatctggaggac ctggaagaagctgaggaacctgacatggaagaagatgacgaccagaaagctggaaggacgagctgta</p>
<p>PDIA4 (H)</p>	<p>atgaggcctagaagggcttttctgctgctcctgctgggactgtgctgctgttggctgttctgctggcgtgagggccctgacgaggatt ccttaacagagagaacgccatcgaggacgaggaagaagaagaggaaagggacgacgacgaagaagaggacgacactgga agtgaagaagaagaaacggcgtctgctgctaacgacgccaactctgataactctgctggccgacaaggacaccgtgctgctga gtttatgctccttggctggccactgcaagcagtttcccctgagtagagaagatcgccaacatcctgaaggacaaggaccctcct attcctgtggcaagatcgacgctacctctgctgctgctggcctccagattcgatgtgctggctaccccaccatcaagactcgaag aaaggccagccgtggactcagaggcagcagaacaagaagagatctgctgccaagagtgctgaggtgctcccaactgatt ggacccctccactgaaagtgaccctggctgaccaaagagaactctgacgaggtctgtaacgatccgacatcctggtgctga gttctacgcccctggtgctgacattgcaagaagctggctcccagtagcgaagggccgccaagagctgccaagcgtgcccctc caattccactgctaaagtgacgcccaccgctgagacagatctggccaagagattgacgtgctccgctatctcaactcaagatct cagaaagggcagaccctacgactacaacggccctcgcgagaagtagcggatctggtattatgatcgagcagagcggccctc ctagcaagagatctgacactgaaacagggtgcaagagtttctgaaggacggcgacgacgtgatcatcctggagttcaaggg cgagagcagccctgctaccagcagtagcaggtgcccgaacaatctgagagaggactacaagttccaccacacctctctacc gagatctgaagttcctgaaagtgctccaggccagctggtgctatcgacgctgagaagttcagctcaagatcagcagcctggag ccactgatggacgtgacggctctaccaggaactccgccaagactctgctgctgaaagtatgcccctgctgctggccag aaaggtgccaacgatgccaagaggtacaccgacacctggtggtggtgactactcctggactcagctcgcactacagagc cgccacacagttctggcgttaaggtgctggaagtgccaaggacttctgtagtacacctcgctatcgccgacgaagaggatta cgccggcgaaggaagatctggactgctgctgctggcagagatgtaatgcccattctggacgagatccggaagaattc ccatggaacctgaagagttcgacagcagaccctgagagaatttggaccctcaagaagggaagctgaagcccgtgatca agagccagcctgctcctaagaacaacaaggccctggaaggtgctggtggcaagacctcgactccatctgtaggaccccaa gaaagacgtgctgattgagtttctgacacctggtggtggcactgaaacagctggaacccctgacactcctcgcaagaagta caagggccagaaaggcctgctgacgccaagatggtgacccagcgaacgatgctgcccctcgacagatacaaggtggaaggt ccctaccatctactgcccctccggcgacaagaaaaacccctggaagttcgaagggcgacagagatctggaacacctgtcc aagttcatcgaggaaacacgccaccaagctgcccggacaagaaggagactgtag</p>
<p>PRDX4 (H)</p>	<p>atggaagctctgctctgttggccgctaccacacctgatcagcgagacatagaaggctgctgctgctccctctgctgctgtttgtg ctgctggcctgctgcaaggtgggagacagaggaaagaccccgacagagagaggaatgccactttatgctgctggccag gtttaccctggcagggctctagagtgctgctggccgatcactccctgacactgccaaggccaagatctccaagcctgctcctattgg gagggcaccgctgctgacgagcggcgagtttaagagctgaagctgaccgactaccggggcagatcctggtgttcttctacccctc tggactcacctctgctgctcccaccagatcattgctcggcagacagctggaagagttccgctcctcaatcaaggtggtggc gctgctcctggtgactcctcagttaccacacctgctgctgacacaccctagaaggaagcggcagctggccctattagaaatccact gctgctcctgctgaccaccagatctctaaggactacggcgtgactctggaagatagcggccatacactcggggcctgttcatcat cgacgacaagggcatcctgctgagatcaccctgaatgatctgctgctggcagatcctggacgagacactgagactggtgag gcctccagtagaccgataagcagcgaagctgctgctgctgctggaagcctggctccgagacaatcattcctgctcctgccc aagctcaagtagctgacaagctgactga</p>
<p>PREB (C)</p>	<p>atgggaagaagaagggcgtgagctgtacagagccccttctctgtacgcccctgcaagtgaccctaagaatggcctgttgatt gctgctggcggaggcggagctgtaagaccggaatcaagaacggcgtgacactctgagcctggaacagatcaacggcctgctg ctgctctgctgactctcagataccgagacaagagaccatgaatctggccctgctgctgctgctgctgctgctgctgctgctg tgctagatgccagctgctgagattccactgctcaccagcagaagcggaaacagaccgagaagctccggctccaaagagcagggcc ctagacagagaaaaggcggcctctgctgagaagaaatctggcctggaagtgacccccgaaggcgtggaactgaaagtgaa aacctggaagcctgagacagacttctcaccgagcctctgcaaaaggtcgtgcttcaaccagcacaacaccctgctgctac cggcggaaacagatggacacgttctgagtggaaggtcccagcctggaagaggtgctggttcaaggctcacgagggcgagat cggagatctgctgctggacctgaggaagctggtcacagctggatggatctgaagggcctcctgctggcagaagagcagctg gttacacagctccagtggaagagaacggccctgctctctgacaccccttacagataccagcctgtagattcggcaaggtgcca ctcagcctggcggcctgagactgtttaccgtgacatccctcacaagcggcgtgagacacccctcctgcttaccctgaccgctgg actcttactctctgctctgagaaccagacctgcccggcacaagtgatctctgctgctgctgctgctgctgctgctgctgctg ctgggacagtgacaggtcctgctgctctatatacgcctcagcctgcaacggctgactacgtgaaagaggtcacggcatcgtc gtgaccgacgtgaccttctgctgagaaggtgctggtgacctgagctgctggacacctcagagacagctctgtttctgctgctgga ctcccgtgctagctgctgctgctcctctagaagatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg atcctgctgtgagctcggccttctgctgctgctg</p>

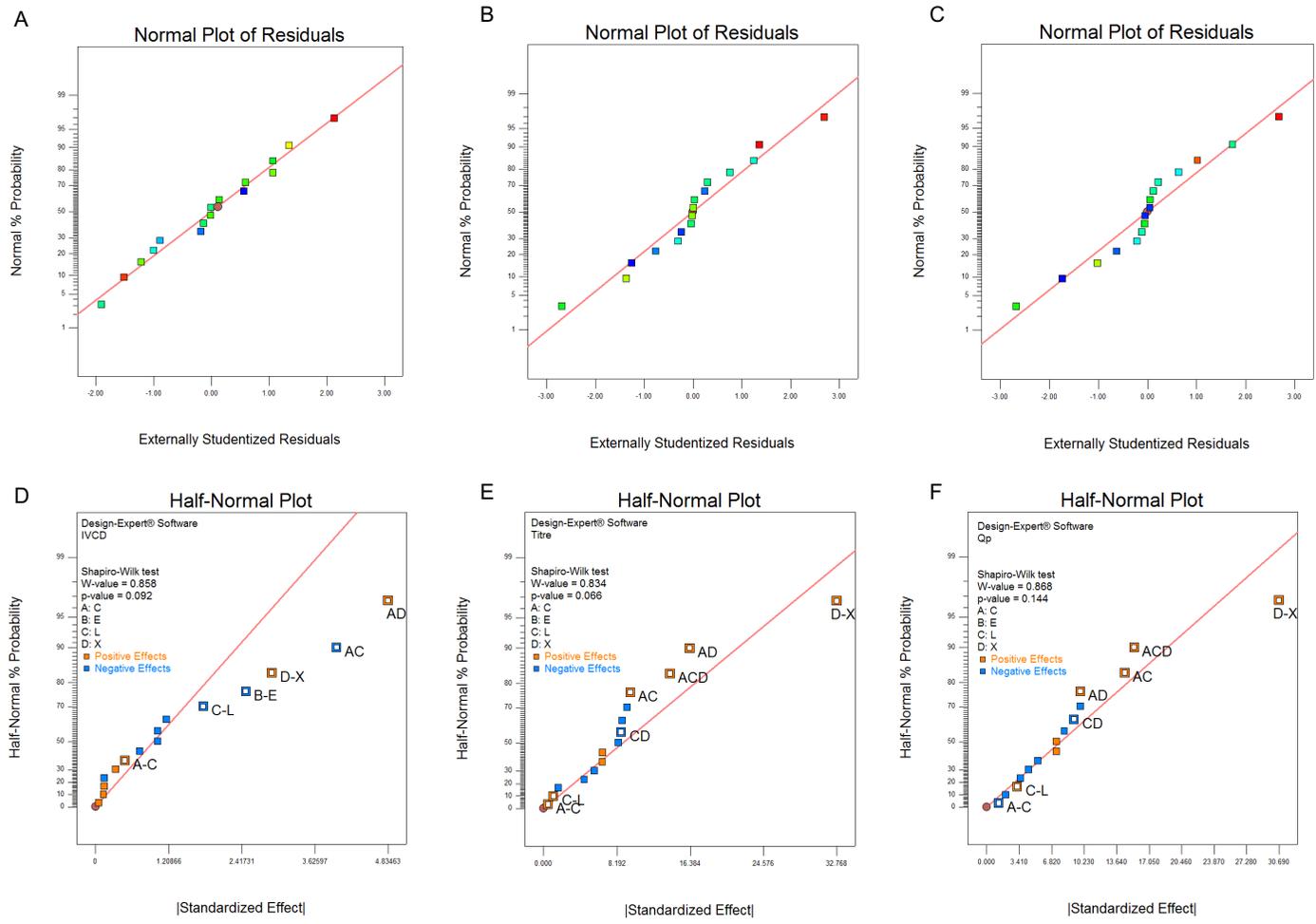
Stx5a (C)	atgattcctcgggaagagatacggctccaagaacaccgatcagggcgtgtacctgggctgtctaagactcaggtgctgtcctctgta ccgcccgtgtcctcctctctgatacaccctctgctacacctgtggctctggctcctcctcctgacaccatgtctgcccgggacaga acccaagagttcctgtccgctgcaagtcctccagtcacagacagaaacggcatccagaccaacaagcctgtctgagagccgca gacagagatccgagttaccctgatggccaagcgatcggaagggacctgttaaacacctcgcaagctggaaaagctgacat cctggccaagagaaagtcctgtcgacgacaagccgtggaatcgaggaaactgacctacatcaaacagggtctctgctga
TFE3 ^Δ (C)	atgtctcatgctgcgaacctgccagagatggcgtggaagcctctgtggaaggcctagagctgtgtctgctgctggaagagaga aggcctgccgattctgctcagctgtcctcaatcctctgctgctgagctggcatcgtggccgatalcagattggagaacattctg gaccccgacagcttctacagagctgaagtctcagccctgctctgagatcctctgccaatctctgcaagctaccctaccctct gccacctgtgtctctagctctgtgcccgatctggaccctgctatgtccgctagtctctctcgggtgctgctgaggcagcagtt gatgagagcccaggctcaagagcaagagcgcagagagagaagagagcaggtctgctccgctcctttcctaactcctgctcctg tctcctgccatcagcgtgatcggagttcagctggggacatacactgggacagcctccaccagctcaggtgccaagagaagtgt gaaggtgcagaccacctgaaaaccccaccagataacctcctcagcagggcagcagcaagtgaaagcagtaacctgtcta ccacctgggcccctaagctggcctcaggtttgactcctcctcctggacctgctctgctcaacctgctcctgctccagagacagct atgtaccggacctaccggctctgcccctaattctcctatggctctgctgacctcctcctccagcagaaaagagtaacctgctgtcc cacagatcgacgacgtgatcgacgagatcatcctcctggaatccagctacaacgacgagatgctgtcttacctgctggcggcaca gctggattgcagctgctctacactgctgtcggcaacctgctggcagctgactctaactcaggggtgctgctaccctgctatcacc gtgtctaactctgtcccggcagctgcccacaatcaagagagagattccgagacagaggttaagccctgtgaaagagcggc agaagaaggacaaccacaacctgatcgagcggcggaggcgggtcaacatcaacgacagaatcaaagagctgggcaactgat cccaagagcaacgaccggagatgagatggaacaaggcaccatcctgaaggcctcctggaactacatccggaagctccag aaagagcagcagcggcgaagcctggaatctagacagagatccctggaaacagggccaacagatctcctcagctgcgcatca agagctggaactccaggctcagattcagggactgccaagtctccaacacctggactgctgctcactggctacctctcctgctccgac tctctgaagcctgagcagctggacatcgaggaagagggcagaccctctaccaccagctttcatgtgtctgcccgcctatccagaat gcccctcaacaacacctcctgctccacctctgacgcctgctggatctgcactttcatctgatcactggcgatctggcgagccc ttttatctggcctcgaggacatcctgatggaagaagaggcgtctgctggaggactttccgggtgaaacattgtctcctctgagagcc gcctctgatcccctgctgagttctgtctcctgcccgtgccaaggcctcctctagacggctccagcttccatggaagaggaatcctga
Tor1a (H)	atgaagctgggtagagctgttctgggactgctgctgctgctcctctgtggtcaggccgtggaacctatctctcctggactggctctg ctggcgtgctgaccggctatctaccctagactgactgctgctgctcctgagctgctgagcagagctgctgagagagc cctccagaagagacctggacgataaacctgtttggccaagcctgccaagaagatcctcctgaaacgcccgtgtcggcttcaaca acccaagcctaagaagcccctgacactgtctctgcatggatggaccggcaccggcaagaactcgtgtccaagatcattgcccag aacatctacgaaggcggcctgaatagcactgctgacactgtttgtgctaccctgactcctcctcagcctcaacatcaccctgt acaaggatcagctccagctgtggatcagaggcaacgtgtccgctgcccagatccatcttcatctcagcagatggacaagatg cacgcccggcctgatcgacccatcaagccccttctgactactacgacctgggtggacggcgtgctcctaccgaaagccatgttcatct ttctgccaacgctggcggcagagaatcaccgatgtggcctggattttggcggagcggcaagcagagagaggacatcaagct gaaggacatcgacagcccctgctcctgctcctgctcaacaacaagaactcggctcctgactcctcctctgactcagccggacct cctcagattactcgtgcccctcctgctcctgagtaacaagcctggaagatgctcatcagagatgctcagagctcagctcgggctacgag atcgacgaggacattgtgtccagagtgccggaggaaatgacattctccaaaagaggaacgggtttctccgacaagggctgcaa gaccgtgtcaccagctggactattactacgacgactga
XBP1s (C)	atgggtggtggctgcttctcctctgctgccactgctgctcctaaaggtgctgctgtgtgtgccaacctgctgctgatggtagaccctg cctctgatggtgctgctgatctagagctgcccggatctgaggcctaagtcgcccctcaggctagaagagacagcggctgaccacct gtctccagaggaaaaggcccctgcccgggaagctgaagaatagagtggtgcccagaccgagacagaaagaaaagcccg gatgtccgagctggaacagcaggtcgtggacctggaagagaaaaccagaagctgctgctggaatcagctgctgagagaaa agaccacggcctgctcagcagaatcaagagctgagaacccgctcggcagctgctgacaaacagagagagccctgag acagagctccaaaggcaatgctgtagacctgtgcccggctctgctgaaagtgtgctggtgctggacctggtgacctctcctgagc atctgctatggactccgacaccgtgactcctgactccgagctgatacctgctgggcatcctggacaagctggaccccgtgatg ttctcaagtgcccctctctgagctcggcaatctggaagaactgcccgaagtgatccccggaccttctctgctcctcactgtctct gtccgtggcaccatcttctgccaagctggaagccatcaacgagctgatcagattcgaccacgtgtacaccaagcctctggtgctcga gatcccctccgagactgagctcagaccaactggtggtcaagatcgaagaggcccctctgctcctccagcaggaagatcacccctg agttcatcgttccgtgaagaaagagcccctgaaagaggactcctccccgagcctgcatcctaacctgctgctcctagccactg ctgaaagcctagctctgctgctggaagcctactctgattggtgtaagggcagccctcctcttctccgatatgctcctcctctg gaaatcgaccactcctgggaagatacctcgcgaatgagctgtcctcctcagctgatctcagtgta
XBP1s (H)	atgggtggtggcgcctcctaatcctgcccagatggcaccctaaaggtgctgctgctgctgcccagcctgctctgctgctggtgctc ctgctggacaggcccctgctctgatggtgctgctcagagaggccttctcctgaggctgcttctggcggactgcccagggcagaa agagacagagactgaccacactgtccccgaggaagggcctgcccgggaagctgaaagaacagagtggtgcccagaccg ccagggaccggaagaaagccagaatgtccgagctggaacagcaggtggtggaacctggaagaggaagaaaccagaagctgctgct ggaaaaccagctgctgagagaaaagaccacggcctggtggtggaaaatcaggaactgcccagcggcggctgggcatggatgctc tggtgctgaggaagaggccgagggcaagggcaatgaagtgcacctgtggcggctctgcccgaagtgctgctggcgtgagc ctgctgaccctcctgagcactgctcctatgactcggcagctcggcagctcctcctcctcctgactcctgctgctgctgggctcctg acaacctggaccccgtgattctcaagtccccagcccctgagcccctcctggaagaactgctgaggtgtacctgagggc ccctctctgctgctcctcctgagctgctcctgggcaacctctccgcaagctggaagccatcaacgagctgatcagattcgacca catctacaccaagcccctggtgctggaatcccctcggagacagagctccaggccaacgtggtctgaaagatcagaggaagccc cctgtcccctagcagaaacgaccacctgagttcatctgctcctgaaagaggaaccccgtggaagatgacctggtgcccgagctg ggaatctcaacctgctgctcctcagccactgcccgaagcctgctcctgctgctggaagcctcctcctgctgctggaagcctcctg ctctgctcccctcctcagacatgctcctcctgctggagtgaaacctcctggaggacacctcggcaatgagctgttccccagctg atctcagtgta

YY1 (C)	atggcatctggcgataccctgtatatcgccaccgacggctctgagatgcctgccgagattgtggaactgcacgagatcgaggtgga aacaatcccgtggaaccatcgagacaaccgctgtggcggaagaggaagatgaggacgaagaggatgaagatggtggtggc ggtggcggcggaggtggacatggacatgctggacaccatcaccatcatcatcaccaccatcaccatccacctatgatcgcccca gcctctggtcaccgatgatcctacacaggtgcaccaccaccaagaagtatcctggtccagaccagagaggaagtcgtcggcgg cgacgattctgatggactgagagctgaggacggctcgaggaccagatcctgattcctgttctgctccagccggcggagatgacga ttatatcgagcagaccctctgaccgtggccgctgctggaaaatctggtggcggaggatctagttcctctggcggaggcagagtga gaaaggcggcggaaagaagtccggcaagaagtcttattgtctgctggcgtgctggtgcagcaggcggaggcggagctgatcctgg aaacaagaagtgggagcagaagcagggtccagatcaagaccctggaaggcgagttctccgtgaccatgtgtcctccgacgaga agaaggacatcgaccagagacagtgggtggaagaacagatcatcggcgagaactcccacctgactactccgagatcatgaca ggcaagaagctgcctcctgcgccatccctggaatcgacctgtctgacctaagcagctggccgagttcgcccggatgaagcctag aaagatcaagaggacgacgcccctggacaatcgctgtcctcataagggctgcaccaagatgttccgggacaactccgcatg cggaaacatctgcacaccacggacctagagtgcacgtgtgcgctgagtggtgcaaggccttctggaatccttaagctgaagcg gcatcagctggtgcataccggcgaagaagccttccagtgcaacttgaaggctcggcaagcgggtctcctcgacttcaatctgaga accacgtgcgcatccacaccggcgtatagacctacgtgtgcctttcgacggctgcaacaagaattcgcccagtcaccaacct gaagtcccacatcctgacacacgccaaggccaagaacaaccagtga
---------	--

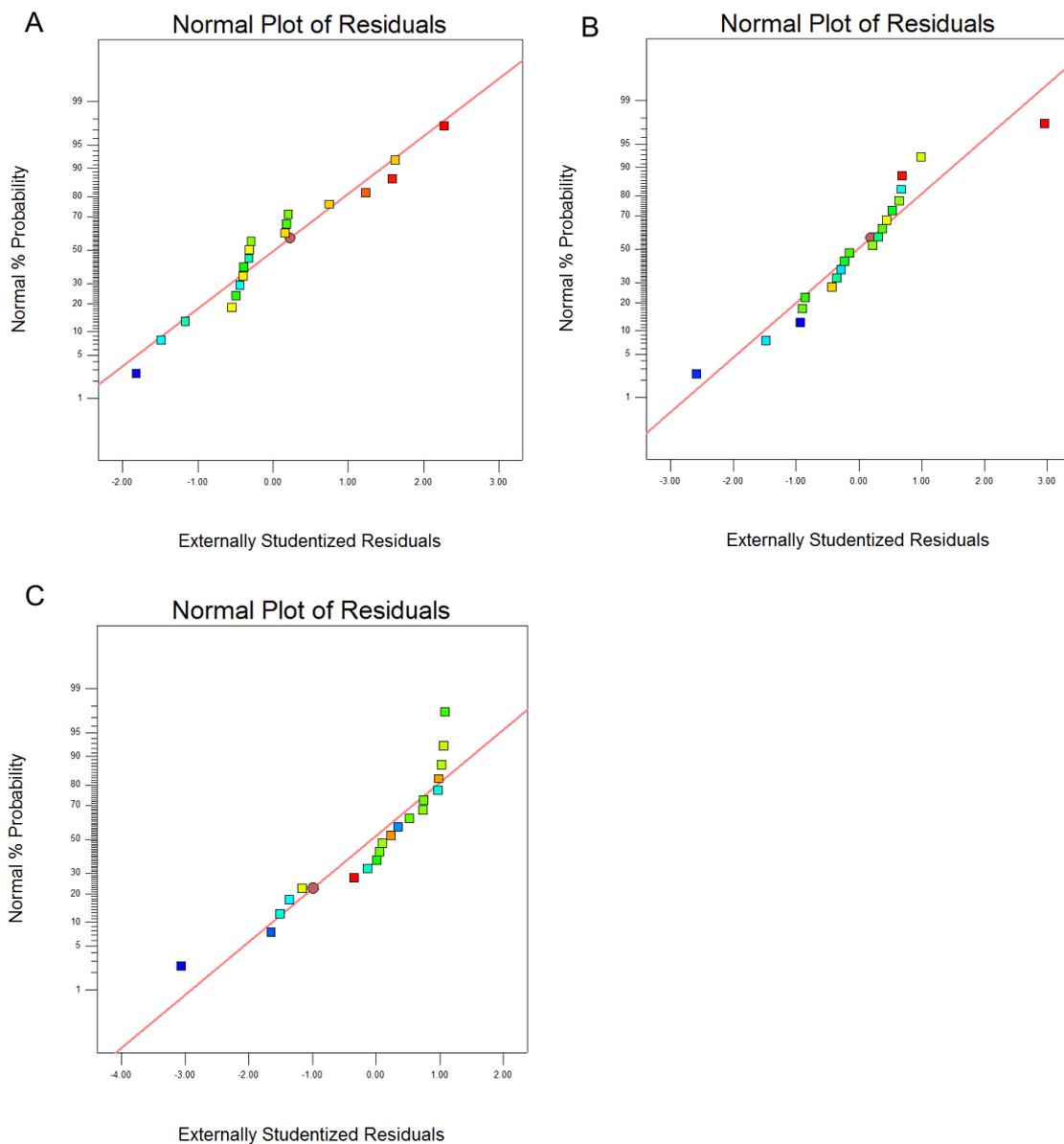
Appendix Table C.6: Summary Matrix for DoE Two-Level Factorial Design: Co-Expression of CypB (C), ER α (E), Largen (L) and XBPs (X)

The factor combinations investigated, output responses and resultant calculated effects are shown.

Run	Individual factors				Factor combinations (interactions)											Response (Relative to Transfection Control %)		
	A: C	B: E	C: L	D: X	CE	CL	CX	EL	EX	LX	CEL	CEX	CLX	ELX	CELX	IVCD	Titre	Qp
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+	100.000	100.000	100.000
2	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-	99.183	83.476	84.319
3	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-	96.422	84.759	87.959
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+	95.041	78.492	83.396
5	-	-	+	-	+	-	-	-	+	-	+	-	+	+	-	99.908	106.251	106.219
6	+	-	+	-	-	+	+	-	+	-	-	+	-	+	+	92.586	74.573	81.027
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+	98.781	106.977	108.494
8	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-	90.900	98.129	109.469
9	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-	96.849	151.212	157.941
10	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+	104.856	124.768	118.607
11	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+	93.595	112.019	121.068
12	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-	105.666	124.502	117.510
13	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+	99.189	103.312	105.658
14	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-	101.856	153.530	151.424
15	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-	97.803	97.133	101.424
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	96.334	128.327	132.768
IVCD effect	0.48	-2.49	-1.78	2.92	-0.15	-3.99	4.83	0.055	0.15	0.34	-1.02	0.13	-0.73	-1.17	-1.03			
Titre effect	0.52	-8.35	1.13	32.77	6.62	9.70	16.35	6.57	-9.36	-8.68	-5.67	-1.65	14.14	-4.55	-8.82			
Qp effect	-1.28	-5.39	3.21	30.69	7.33	14.50	9.83	7.35	-9.83	-9.17	-4.39	-1.99	15.50	-3.58	-8.16			



Appendix Figure C.1: Normal and Half Normal Probability Plots for DoE Two-Level Factorial Design: CypB, ERO1L α , Largen and XBPs
Normality and half Normality (with model terms highlighted) plots for: IVCD: (A) and (D) Titre: (B) and (E) and qP: (C) and (F), respectively.



Appendix Figure C.2: Normal Plots of Residuals for DoE-RSM: CypB, Largen and XBP's Normality plots for: IVCD (A) Titre (B) and qP (C).

Appendix Table C.7: Results of Ordinary One-Way ANOVA With Dunnett's Multiple Comparisons Test

ANOVA for experiments described in Section 7.3.6.2. All 28 effector genes, tested at four different gene doses were statistically compared to control transfections. Gene loads represented as 1, 2, 3 and 4 represent plasmid doses 15%, 7.5%, 3.75%, 1.88% (w/w) of recombinant product plasmid loading, respectively.

Dunnett's multiple comparisons test	IVCD			Titre			Specific Productivity		
	Significant?	Summary	Adjusted P Value	Significant?	Summary	Adjusted P Value	Significant?	Summary	Adjusted P Value
Control vs. Dad1 1	No	ns	0.9994	No	ns	0.9822	No	ns	0.9916
Control vs. Dad1 2	No	ns	0.9997	No	ns	0.91	No	ns	0.9801
Control vs. Dad1 3	No	ns	0.9981	No	ns	0.999	No	ns	0.9981
Control vs. Dad1 4	No	ns	0.9996	No	ns	0.9979	No	ns	0.9982
Control vs. ERGIC53 1	No	ns	0.9995	Yes	**	0.0077	No	ns	0.1536
Control vs. ERGIC53 2	No	ns	0.999	No	ns	0.1181	No	ns	0.7951
Control vs. ERGIC53 3	No	ns	0.9999	No	ns	0.9631	No	ns	0.9908
Control vs. ERGIC53 4	No	ns	0.9999	No	ns	0.8393	No	ns	0.979
Control vs. HSPA1a 1	No	ns	0.9982	No	ns	0.9999	No	ns	0.999
Control vs. HSPA1a 2	No	ns	0.9769	No	ns	0.9999	No	ns	0.9914
Control vs. HSPA1a 3	No	ns	0.9985	No	ns	0.9993	No	ns	0.9997
Control vs. HSPA1a 4	No	ns	0.9982	No	ns	0.9988	No	ns	0.9999
Control vs. PRDX4 1	No	ns	0.9988	No	ns	0.8608	No	ns	0.7011
Control vs. PRDX4 2	No	ns	0.9914	No	ns	0.9996	No	ns	0.9992
Control vs. PRDX4 3	No	ns	0.9979	No	ns	0.9999	No	ns	0.9988
Control vs. PRDX4 4	No	ns	0.9803	No	ns	0.999	No	ns	0.9996
Control vs. PREB 1	No	ns	0.9986	No	ns	0.9997	No	ns	0.9988
Control vs. PREB 2	No	ns	0.8124	No	ns	0.9988	No	ns	0.9991
Control vs. PREB 3	No	ns	0.4033	No	ns	0.9988	No	ns	0.2129
Control vs. PREB 4	No	ns	0.9913	No	ns	0.9994	No	ns	0.984
Control vs. PRKD1 1	No	ns	0.9833	No	ns	0.9994	No	ns	0.9994
Control vs. PRKD1 2	No	ns	0.6562	No	ns	0.7475	No	ns	0.9998
Control vs. PRKD1 3	No	ns	0.2134	No	ns	0.8076	No	ns	0.9998
Control vs. PRKD1 4	No	ns	0.9819	No	ns	0.9986	No	ns	0.9997
Control vs. SGMS1 1	No	ns	0.149	Yes	**	0.0016	No	ns	0.778
Control vs. SGMS1 2	No	ns	0.0893	No	ns	0.4297	No	ns	0.9998
Control vs. SGMS1 3	Yes	*	0.0317	No	ns	0.9803	No	ns	0.998
Control vs. SGMS1 4	No	ns	0.8827	No	ns	0.9411	No	ns	0.9997
Control vs. SRP14 1	No	ns	0.9989	No	ns	0.9996	No	ns	0.9986
Control vs. SRP14 2	No	ns	0.9984	No	ns	0.9996	No	ns	0.9986
Control vs. SRP14 3	No	ns	0.9983	No	ns	0.982	No	ns	0.788
Control vs. SRP14 4	No	ns	0.9916	No	ns	0.9999	No	ns	0.9983

Control vs. Stx5a 1	No	ns	0.3247	No	ns	0.9995	No	ns	0.9815
Control vs. Stx5a 2	No	ns	0.8813	No	ns	0.9996	No	ns	0.9778
Control vs. Stx5a 3	No	ns	0.7343	No	ns	0.999	No	ns	0.7431
Control vs. Stx5a 4	No	ns	0.7177	No	ns	0.9994	No	ns	0.8992
Control vs. TFE3 1	No	ns	0.1069	Yes	****	0.0001	No	ns	0.293
Control vs. TFE3 2	Yes	*	0.0454	No	ns	0.2359	No	ns	0.9997
Control vs. TFE3 3	No	ns	0.1732	No	ns	0.9981	No	ns	0.9986
Control vs. TFE3 4	No	ns	0.0549	No	ns	0.5581	No	ns	0.9996
Control vs. FKBP11 1	Yes	*	0.011	Yes	**	0.0017	No	ns	0.9979
Control vs. FKBP11 2	No	ns	0.0684	Yes	*	0.019	No	ns	0.9981
Control vs. FKBP11 3	Yes	**	0.0086	No	ns	0.8298	No	ns	0.9984
Control vs. FKBP11 4	Yes	*	0.0205	No	ns	0.8486	No	ns	0.9986
Control vs. P115 1	No	ns	0.9777	No	ns	0.9997	No	ns	0.9989
Control vs. P115 2	No	ns	0.3755	No	ns	0.9995	No	ns	0.6157
Control vs. P115 3	No	ns	0.1854	No	ns	0.9997	No	ns	0.5034
Control vs. P115 4	No	ns	0.7648	No	ns	0.9998	No	ns	0.998
Control vs. Bet3 1	No	ns	0.5708	No	ns	0.9997	No	ns	0.991
Control vs. Bet3 2	No	ns	0.3452	No	ns	0.9984	No	ns	0.1654
Control vs. Bet3 3	No	ns	0.0859	No	ns	0.9921	Yes	*	0.0196
Control vs. Bet3 4	Yes	***	0.0002	No	ns	0.9991	No	ns	0.056
Control vs. Bet1 1	Yes	*	0.0345	No	ns	0.1468	No	ns	0.9997
Control vs. Bet1 2	No	ns	0.099	No	ns	0.9992	No	ns	0.962
Control vs. Bet1 3	Yes	*	0.0251	No	ns	0.9985	No	ns	0.7814
Control vs. Bet1 4	Yes	*	0.0359	No	ns	0.9986	No	ns	0.8471
Control vs. CREB3L2 1	No	ns	0.5225	No	ns	0.8508	No	ns	0.9996
Control vs. CREB3L2 2	No	ns	0.0781	No	ns	0.367	No	ns	0.9997
Control vs. CREB3L2 3	No	ns	0.7412	No	ns	0.6992	No	ns	0.9991
Control vs. CREB3L2 4	No	ns	0.1571	No	ns	0.778	No	ns	0.9996
Control vs. CRELD2 1	No	ns	0.9993	No	ns	0.9983	No	ns	0.9991
Control vs. CRELD2 2	No	ns	0.7575	No	ns	0.1519	No	ns	0.9985
Control vs. CRELD2 3	No	ns	0.9996	No	ns	0.3206	No	ns	0.6684
Control vs. CRELD2 4	No	ns	0.9984	No	ns	0.9988	No	ns	0.9997
Control vs. Rab1a 1	No	ns	0.9996	Yes	*	0.011	No	ns	0.0796
Control vs. Rab1a 2	No	ns	0.9999	No	ns	0.5242	No	ns	0.903
Control vs. Rab1a 3	No	ns	0.9999	No	ns	0.7914	No	ns	0.9783
Control vs. Rab1a 4	No	ns	0.9991	No	ns	0.9916	No	ns	0.9994
Control vs. Cert 1	No	ns	0.9996	Yes	***	0.0003	Yes	*	0.0117
Control vs. Cert 2	No	ns	0.9996	Yes	*	0.0381	No	ns	0.3601
Control vs. Cert 3	No	ns	0.9997	No	ns	0.4302	No	ns	0.937

Control vs. Cert 4	No	ns	0.8808	No	ns	0.9808	No	ns	0.8425
Control vs. YY1 1	No	ns	0.9999	No	ns	0.998	No	ns	0.9985
Control vs. YY1 2	No	ns	0.978	No	ns	0.9987	No	ns	0.9791
Control vs. YY1 3	No	ns	0.9986	No	ns	0.9999	No	ns	0.9994
Control vs. YY1 4	No	ns	0.9986	No	ns	0.9993	No	ns	0.9999
Control vs. Tor1a 1	No	ns	0.9982	Yes	****	0.0001	Yes	****	0.0001
Control vs. Tor1a 2	No	ns	0.9994	Yes	*	0.0195	No	ns	0.1137
Control vs. Tor1a 3	No	ns	0.9996	No	ns	0.9993	No	ns	0.9992
Control vs. Tor1a 4	No	ns	0.998	No	ns	0.9989	No	ns	0.9993
Control vs. PDIA4 1	No	ns	0.9059	No	ns	0.9996	No	ns	0.9982
Control vs. PDIA4 2	No	ns	0.9996	No	ns	0.9999	No	ns	0.9999
Control vs. PDIA4 3	No	ns	0.9998	No	ns	0.998	No	ns	0.9988
Control vs. PDIA4 4	No	ns	0.397	No	ns	0.7939	No	ns	0.9998
Control vs. MCFD2 1	No	ns	0.9219	No	ns	0.9599	No	ns	0.9996
Control vs. MCFD2 2	No	ns	0.7294	No	ns	0.4678	No	ns	0.9992
Control vs. MCFD2 3	No	ns	0.4668	No	ns	0.9908	No	ns	0.9997
Control vs. MCFD2 4	No	ns	0.8193	No	ns	0.9981	No	ns	0.9999
Control vs. CANX 1	No	ns	0.9837	No	ns	0.4234	No	ns	0.4151
Control vs. CANX 2	No	ns	0.9986	No	ns	0.9992	No	ns	0.9999
Control vs. CANX 3	No	ns	0.9835	No	ns	0.9986	No	ns	0.9999
Control vs. CANX 4	No	ns	0.6919	No	ns	0.9908	No	ns	0.9999
Control vs. Calr 1	No	ns	0.9998	Yes	*	0.0219	No	ns	0.2452
Control vs. Calr 2	No	ns	0.9993	No	ns	0.9983	No	ns	0.9993
Control vs. Calr 3	No	ns	0.9994	No	ns	0.998	No	ns	0.9991
Control vs. Calr 4	No	ns	0.9996	No	ns	0.998	No	ns	0.9989
Control vs. Hyou1 1	No	ns	0.9999	Yes	****	0.0001	Yes	**	0.0016
Control vs. Hyou1 2	No	ns	0.9979	No	ns	0.8832	No	ns	0.7551
Control vs. Hyou1 3	No	ns	0.9655	No	ns	0.9999	No	ns	0.9983
Control vs. Hyou1 4	No	ns	0.9998	No	ns	0.9987	No	ns	0.9994
Control vs. Grp94 1	No	ns	0.9999	No	ns	0.9996	No	ns	0.9996
Control vs. Grp94 2	No	ns	0.9998	No	ns	0.6465	No	ns	0.9269
Control vs. Grp94 3	No	ns	0.9997	No	ns	0.684	No	ns	0.9787
Control vs. Grp94 4	No	ns	0.999	No	ns	0.3018	No	ns	0.9772
Control vs. XBP1s (C)1	No	ns	0.9813	Yes	****	0.0001	Yes	****	0.0001
Control vs. XBP1s (C) 2	No	ns	0.9992	Yes	****	0.0001	Yes	****	0.0001
Control vs. XBP1s (C) 3	No	ns	0.9914	Yes	***	0.0005	Yes	****	0.0001
Control vs. XBP1s (C) 4	No	ns	0.9999	No	ns	0.5635	No	ns	0.9989
Control vs. Rab11a 1	No	ns	0.9996	No	ns	0.9982	No	ns	0.999
Control vs. Rab11a 2	No	ns	0.9997	No	ns	0.9987	No	ns	0.9993
Control vs. Rab11a 3	No	ns	0.9999	No	ns	0.992	No	ns	0.9988
Control vs. Rab11a 4	No	ns	0.9999	No	ns	0.9982	No	ns	0.9988

