The University Of Sheffield.

Chemical & Biological Engineering.

Cell Line and Process Development for Improved Transient Production of a "Difficult-to-Express" Fusion Protein by CHO Cells

YUSUF B JOHARI

Supervisor:

Prof. David C James

In partnership with:

Biogen Idec, USA

biogen idec®

A thesis presented to The University of Sheffield in fulfilment
of the thesis requirement for the degree of
Doctor of Philosophy in Chemical and Biological Engineering

May 2015

Declaration

I, Yusuf Johari, declare that I am the sole author of this thesis and that the research presented within is the result of my own efforts and achievements, unless acknowledged otherwise in the text. I confirm that this work has not been submitted for any other degrees.

Abstract

Despite the remarkable yield improvements of recombinant proteins produced in mammalian cells, some “difficult-to-express” (DTE) proteins achieve considerably lower production titres. The bottlenecks are exacerbated in the case of transient gene expression (TGE) systems as the host cells are easily overloaded with recombinant genes, hence necessitating intensive cell line and process development. The aims of this thesis are to study the limitations to high TGE yields of DTE proteins, and subsequently investigate strategies to efficiently alleviate the bottlenecks. For this purpose, we used a model DTE Fc-fusion protein (Sp35:Fc; proprietary of Biogen Idec) expressed in Chinese hamster ovary (CHO) cells, as well as secreted alkaline phosphatase and green fluorescent protein for comparisons. Through analyses of intracellular and extracellular Sp35:Fc polypeptides, we found that post-translational mechanisms were limiting in the cells. The saturation of Sp35:Fc expression coincided with the retention of folded proteins in the ER and the increase in disulphide bonded aggregates. Further *in silico* analysis via a mathematical model enabled identification of the relative importance of specific cellular process on Sp35:Fc productivity (*q*P). Based on these observations, three strategies aimed at improving the transient production of Sp35:Fc were investigated.

The first strategy involved the evaluation of functional performance of clonally derived cell populations to produce Sp35:Fc. We critically assessed the key intrinsic functional traits of the clones, and their impact on Sp35:Fc production. The data indicate that cell lines with the capability to accumulate high biomass while maintaining relatively high specific growth rate (μ) were likely to be high producers for DTE proteins. For the second strategy, we utilised a novel vector system specifically beneficial for DTE proteins by incorporating ER stress response elements (ERSE) into the SV40 vector expressing Sp35:Fc or the UPR transactivator ATF6c. The ERSE-SV40 vectors acted as a synthetic "amplifier/dual activator" circuit, where expressed ATF6c amplified both its own and Sp35:Fc expression via activation of ERSE-SV40, whilst generally transactivating cellular ER capacity via endogenous ERSE. In the third approach, we addressed the hypothesis that specific functional proteins and chemical chaperones could improve ER capacity for Sp35:Fc folding reactions, increase secretion rate and/or relieve host cells from ER stress. We employed cell/process engineering by co-expressing a variety of molecular chaperones or UPR transactivators with Sp35:Fc, as well as a range of chemicals and hypothermic condition. We observed that Sp35:Fc production could be improved via two distinct modes; (i) increase in *q*P correlated to repression of μ, and (ii) stimulation of μ with general reduction in *q*P. In this regard, genes and chemicals could work synergistically to provide an optimal solution.

Overall, this study illustrates that effective cell line and process development for DTE protein production requires a synergistic combination of vector, cell and process engineering strategies designed to alleviate cellular bottlenecks simultaneously, enabling the host cell to attain both high *q*P and cell density. Using two clonal variants and combinations of cell and process engineering, the transient Sp35:Fc production yield could be increased by more than six-fold. Rapid, high-throughput predictive mathematical tools would be particularly valuable in assessing the relative/synergistic impact of different engineering interventions.

Keywords: Chinese hamster ovary cells; transient gene expression; difficult-to-express protein; aggregation; unfolded protein response; mathematical modelling; genetic heterogeneity; screening; vector engineering; cell engineering; chemical chaperones.

Acknowledgements

Four years of my hard work, but this thesis, and my PhD, is also the outcome of direct and indirect touches of many people whom I owe my deepest gratitude to;

To David James, who made this possible at the first place via a PhD studentship, and whose supervision made it possible for the subsequent four years. I doubt I could find a better supervisor. Indeed, it was an honour to be your (overworked) student.

To Scott Estes and Christina Alves at Biogen Idec, as well as late Marty Sinacore who was involved during the early stage of the project. It was great to work with all of you, and you aspire me to be a *real* bioengineer.

To all post-docs and fellow PhD students in David's group through the years; Claire Bennet, Adam Brown, Claire Bryant, Joe Cartwright, Sarah Davies, Danielle Fairbrass, Alejandro Fernandez-Martell, Beata Florczak, Darren Geoghegan, Rhian Grainger, Joanne Longster, Joseph Longworth, Clare Lovelady, Jane McLeod, Olivia Mozley, Peter O'Callaghan, Leon Pybus, Camille Segarra, Katie Syddall, Ben Thompson, Christa Walther, Nathan West and Robert Whitfield. I really appreciate the helps, insights, tips and discussions, however small they were. Special thanks to Alejandro for assisting with several tedious rounds of plate-based Nucleofection even though the results are not included in this thesis.

To the thesis "reviewers"; Paul Dobson, Junaid Raja and again Claire Bryant, Katie Syddall and Nathan West. A big thank you for the kind and helpful feedbacks and comments.

To all friends in D72 office, especially Stephen Jaffe, Rahul Kapoor and Mahendra Raut who often provided company late at night when sensible people were fast asleep (or partying!).

.....

And above all, to my family. For everything.

Table of Contents

List of Figures

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Appendix

.

.

List of Tables

Abbreviations

2FI	two-factor interaction
ANOVA	analysis of variance
ATF4	activating transcription factor 4
ATF6	activating transcription factor 6
ATF6c	cleaved activating transcription factor 6
Bcl-2	B-cell lymphoma 2
Bcl-x_L	B-cell lymphoma x_L
BHK	baby hamster kidney
BiP	binding immunoglobulin protein
BSA	bovine serum albumin
bZIP	basic leucine zipper
cDNA	complementary deoxyribonucleic acid
CHO	Chinese hamster ovary
CHOP	C/EBP homologous protein
CMV	cytomegalovirus
CypB	cyclophilin B
DHFR	dehydrofolate reductase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DOE	design of experiment
DPBS	Dulbecco's phosphate buffered saline
DTE	difficult to express
DTT	dithiothreitol
EBNA1	Epstein-Barr nuclear antigen 1
EBV	Epstein-Barr virus
EDEM	endoplasmic reticulum degradation-enhancing α-mannosidase-like
EF1α	elongation factor 1 α
eIF2α	eukaryotic translation initiation factor 2 α
ELISA	enzyme-linked immunosorbent assay
EPO	erythropoietin
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum associated degradation
ERManI	ER mannosidase I
ERO1	endoplasmic reticulum oxidoreductin 1
ERO1Lα	endoplasmic reticulum oxidoreductin-1-like α
ERO1Lβ	endoplasmic reticulum oxidoreductin-1-like β
ERSE	endoplasmic reticulum stress response element
ETE	easy to express
FDA	U.S. Food and Drug Administration

GADD	growth arrest and DNA damage
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GDP	guanine diphosphate
GFP	green fluorescent protein
GS	glutamine synthetase
HDAC	histone deacetylation
HEK	human embryonic kidney
IE2	immediate early 2
IgG	immunoglobulin G
IRE1	inositol-requiring enzyme 1
IVCD	integral of viable cell density
LDC	limiting dilution cloning
LRR	leucine-rich repeat
MAb	monoclonal antibody
MCA	metabolic control analysis
M-CSF	macrophage colony-stimulating factor
mRNA	messenger ribonucleic acid
MSX	methionine sulphoximine
MTX	methotrexate
NaBu	sodium butyrate
NEM	*N*-ethylmaleamide
Neu5Gc	*N*-glycolylneuraminic acid
OFAT	one factor at a time
ORF	open reading frame
OriP	origin of replication
OST	oligosaccharyltransferase
PAT	process analytical technology
PBA	sodium 4-phenylbutyrate
PBS	phosphate buffered saline
PDI	protein disulphide isomerase
p-eIF2α	phosphorylated eukaryotic translation initiation factor 2 α
PERK	protein kinase-like endoplasmic reticulum kinase
PFA	paraformaldehyde
PPMCC	Pearson's product moment correlation coefficient
*q*P	cell specific productivity
qRT-PCR	quantitative real-time polymerase chain reaction
R&D	research and development
RC	response coefficient
rDNA	recombinant deoxyribonucleic acid
RNA	ribonucleic acid
SRP	signal-recognition particle
RSM	response surface methodology

SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEAP	secreted alkaline phosphatase
SV40	simian virus 40
TGE	transient gene expression
TMAO	trimethylamine N-oxide
TNFR	tumour necrosis factor receptor
tPA	tissue plasminogen activator
UPR	unfolded protein response
UPRE	unfolded protein response elements
UDP	uridine diphosphate
VCD	viable cell density
VPA	valproic acid
XBP1	X-box binding protein 1
XBP1s	spliced X-box binding protein 1
XBP1u	unspliced X-box binding protein 1
YFP	yellow fluorescent protein

Nomenclatures

Uppercase

$[\mathrm{P}]$	concentration of unfolded protein monomer in the ER (pp chain cell^{-1})
$[\mathrm{P}_2]_{\mathrm{ER}}$	concentration of protein dimer in the ER (molecule cell^{-1})
$[\mathrm{P}_2]_{\mathrm{G}}$	concentration of protein dimer in the Golgi complex (molecule cell^{-1})
$[\mathrm{P}]_{\mathrm{ubq}}$	concentration of ubiquitinated protein monomer in the ER (molecule cell^{-1})
$[\mathrm{P}_2]_{\mathrm{ubq}}$	concentration of ubiquitinated protein dimer in the ER (molecule cell^{-1})
$[\mathrm{P}_4]_{\mathrm{ER}}$	concentration of protein tetramer in the ER (molecule cell^{-1})
$[\mathrm{P}_4]_{\mathrm{G}}$	concentration of protein tetramer in the Golgi complex (molecule cell^{-1})
$[\mathrm{P}_6]_{\mathrm{ER}}$	concentration of protein hexamer in the ER (molecule cell^{-1})
$[\mathrm{P}_6]_{\mathrm{G}}$	concentration of protein hexamer in the Golgi complex (molecule cell^{-1})
N_{gene}	gene copy number (copy number cell^{-1})
R_{P2}	production rate of protein dimer (molecule cell^{-1} h^{-1})
R_{P4}	production rate of protein tetramer (molecule cell^{-1} h^{-1})
R_{P6}	production rate of protein hexamer (molecule cell^{-1} h^{-1})
S_m	specific transcription rate of mRNA (mRNA molecule h^{-1})
T_{P}	specific translation rate of monomer (pp chain (mRNA molecule)$^{-1}$ h^{-1})

Lowercase

$k_{\mathrm{agg,P4}}$	rate constant for dimer-dimer aggregate formation (cell molecule^{-1} h^{-1})
$k_{\mathrm{agg,P6}}$	rate constant for dimer-tetramer aggregate formation (cell h molecule^{-1})
k_{deg}	degradation rate of mRNA molecule (h^{-1})
$k_{\mathrm{ER,P2}}$	rate constant for dimer transport from the ER to the Golgi complex (h^{-1})
$k_{\mathrm{ER,P4}}$	rate constant for tetramer transport from the ER to the Golgi complex (h^{-1})
$k_{\mathrm{ER,P6}}$	rate constant for hexamer transport from the ER to the Golgi complex (h^{-1})
k_{fa}	rate constant for monomer-monomer assembly (cell molecule^{-1} h^{-1})
$k_{\mathrm{G,P2}}$	rate constant for dimer transport from the Golgi to the culture medium (h^{-1})
$k_{\mathrm{G,P4}}$	rate constant for tetramer transport from the Golgi to the culture medium (h^{-1})
$k_{\mathrm{G,P6}}$	rate constant for hexamer transport from Golgi to the culture medium (h^{-1})
$k_{\mathrm{ubq,P}}$	rate constant for protein monomer ubiquitination in the ER (h^{-1})
$k_{\mathrm{ubq,P2}}$	rate constant for protein dimer ubiquitination in the ER (h^{-1})
$k_{\mathrm{ubq,P4}}$	rate constant for protein tetramer ubiquitination in the ER (h^{-1})
$k_{\mathrm{ubq,P6}}$	rate constant for protein hexamer ubiquitination in the ER (h^{-1})

$[m]$	intracellular mRNA concentration (mRNA cell^{-1})
$t_{1/2,m}$	half-life of mRNA molecule (h)
$t_{1/2,[\mathrm{P2}]}$	time required for 50% of protein dimer to be ubiquitinated (h)
$t_{1/2,[\mathrm{P4}]}$	time required for 50% of protein tetramer to be ubiquitinated (h)
$t_{1/2,[\mathrm{P6}]}$	time required for 50% of protein hexamer to be ubiquitinated (h)
$t_{1/2,[\mathrm{P2}]\mathrm{ER}}$	half-time for ER-to-Golgi transport for protein dimer (h)
$t_{1/2,[\mathrm{P2}]\mathrm{G}}$	half-time for Golgi-to-culture medium transport for protein dimer (h)
$t_{1/2,[\mathrm{P4}]\mathrm{ER}}$	half-time for ER-to-Golgi transport for protein tetramer (h)
$t_{1/2,[\mathrm{P4}]\mathrm{G}}$	half-time for Golgi-to-culture medium transport for protein tetramer (h)
$t_{1/2,[\mathrm{P6}]\mathrm{ER}}$	half-time for ER-to-Golgi transport for protein hexamer (h)
$t_{1/2,[\mathrm{P6}]\mathrm{G}}$	half-time for Golgi-to-culture medium transport for protein hexamer (h)
$q\mathrm{P}_{\mathrm{dimer}}$	specific production rate of native dimer (pg cell^{-1} h^{-1})
$q\mathrm{P}_{\mathrm{tetramer}}$	specific production rate of tetramer aggregate (pg cell^{-1} h^{-1})
$q\mathrm{P}_{\mathrm{hexamer}}$	specific production rate of hexamer aggregate (pg cell^{-1} h^{-1})

Greek symbols

λ_{P2}	unit conversion factor of protein dimer (3.32×10^{-7} pg h molecule^{-1} day^{-1})
μ	cell specific growth/proliferation rate (h^{-1})

Chapter 1

Biopharmaceuticals and Their Production Systems

This chapter provides an overview of biopharmaceuticals and their expression systems. Specific emphasis is given to mammalian cell culture and challenges faced in the production process. The objective of this chapter is to contextualise the research presented in this thesis, with regard to the body of literature describing transient production and difficult-to-express proteins.

1.1 Protein-Based Therapeutics

1.1.1 Monoclonal antibodies and Fc-fusion proteins

The term biopharmaceutical may refer to therapeutic biological drugs derived from living cells such as antibodies, vaccines, hormones and nucleic acid-based products. Today, there are more than 200 approved biopharmaceuticals for use in diverse clinical settings. Among these, recombinant protein-based products particularly monoclonal antibodies (MAbs) and Fc-fusion proteins, constitute the most rapidly growing class of biopharmaceuticals and currently dominate the biologic drug market. In terms of market value, MAbs and Fc-fusion proteins together account for 54% of the US$ 140 billion global sales of therapeutic proteins in 2013 (Walsh, 2014). The majority of recombinant protein-based products are used against various types of cancer while other main treatment areas include autoimmune diseases such as arthritis, as well as transplantation and respiratory disorders (Huggett, 2013; Sommerfeld and Strube, 2005).

The number of therapeutic MAb in development has continued to increase tremendously over the last three decades due to the flexible and robust nature of the immunoglobulin molecule, as well as advances in molecular biology and technology. Indeed, since the approval of the first MAb product Orthoclone OKT3® in 1986, MAb formats have evolved from entirely murine structure to "chimeric" form and later, fully humanised antibodies containing human sequences. Engineered human MAbs present a generally improved effector function, reduced immunogenicity and increased stability (Johnston, 2007; Sommerfeld and Strube, 2005). Gene cloning/synthesis and recombinant DNA technologies allow the production of new MAbs against a virtually unlimited number of target antigens (Igawa et al., 2011; Tachibana et al., 1999). Today, therapeutic MAbs boast the most active pipeline in the recombinant biopharmaceutical industry, with six of the top-ten selling biologics in 2013 being MAbs (Table 1.1).

Meanwhile, many other biologically active protein and peptide drugs have very short serum half-lives due to fast renal clearance that limit the therapeutic efficacy. Considerable efforts have been made to develop techniques to extend the half-life of these therapeutic proteins (Kontermann, 2011). Fc fusion technology, by which the Fc domain of IgG is joined to a therapeutic protein, is arguably the most effectual solution where it extends a protein's *in vivo* half-life via the neonatal Fc receptor recycling mechanism. In addition to the enhanced pharmacokinetics and pharmacodynamics, this technology also confers novel properties to the hybrid molecule such as Fc receptor and protein A bindings. With

Table 1.1: Ten global top selling therapeutic proteins in 2013 (Hugget, 2013; Walsh, 2014).

Product	Lead company	Molecule type	Production host	Main treatment	Sales ($ billions)
Humira® (adalimumab)	AbbVie	MAb	CHO	Arthritis	11.00
Enbrel® (etanercept)	Amgen	Fc-fusion protein	CHO	Arthritis	8.76
Remicade® (infliximab)	Johnson & Johnson	MAb	Murine myeloma	Arthritis	8.37
Lantus® (insulin glargine)	Sanofi	Protein	*E. coli*	Diabetes	7.95
Rituxan® (rituximab)	Roche	MAb	CHO	Arthritis, cancer	7.91
Avastin® (bevacizumab)	Roche	MAb	CHO	Cancer	6.97
Herceptin® (trastuzumab)	Roche	MAb	CHO	Cancer	6.91
Neulasta® (pegfilgrastim)	Amgen	Protein	*E. coli*	Neutropenia/ leukopenia	4.39
Lucentis® (ranibizumab)	Roche	MAb	*E. coli*	Macular degeneration	4.27
Epogen® (epoetin alfa)	Amgen	Protein	CHO	Anaemia	3.35

these advantages, Fc fusion technology is used in nine FDA-approved drugs for the treatment of chronic diseases including rheumatoid arthritis, platelet disorders and psoriasis whilst several more are awaiting approvals or in the development pipelines (Mei et al., 2013; Walsh, 2014).

1.1.2 "Difficult-to-express" recombinant proteins

A new range of engineered proteins, such as fusion proteins and bispecific antibodies, have begun to fill the development pipelines of many biopharmaceutical companies. These so called next-generation biologics can be significantly more ‘‘difficult to express’’ (DTE) compared to natural protein formats—generally host cells cannot correctly fold and process the recombinant polypeptide where factors such as translation rate and redox potential can be limiting, therefore resulting in low production titres. Additionally, many complex proteins involve post-translational modifications (glycosylation, phosphorylation, etc) that are absent or vary from expression host to expression host, and the problems can extend to protein insolubility, degradation and aggregation, thereby affecting product qualities (DePalma, 2012; Hall, 2007). Artificial fusion proteins, for instance, have not

coevolved in which the interdomain interactions might not contribute to overall conformational stability or could even destabilise the confirmation (Fast et al., 2009). Therefore, DTE proteins require effective cell line and process development to achieve a good balance between drug manufacturing costs (i.e. prices) and therapeutic benefits.

1.2 Mammalian Expression Systems

1.2.1 Mammalian cells as a preferred expression system

A wide range of hosts are available for the production of biopharmaceuticals including bacteria, yeast, insect, plant and mammalian cells. Prokaryotic systems, especially *E. coli*, have the advantage of low cost and high productivity compared to eukaryotic expression systems. However, bacterial cells are unable to carry out the post-translational modifications required for the biological function of many recombinant proteins (Pandhal and Wright, 2010). On the other hand, yeast, insect and plant cells have a limited capability in synthesising proteins that are similar to those naturally occurring in humans. For instance, yeast cells are known to provide *N*-linked and *O*-linked high-mannose-type glycans that could be immunogenic in humans. In this regard, mammalian cells become the preferential host as they have the correct molecular machinery to perform human-like post-translational modifications (Dasgupta et al., 2007; Durocher and Butler, 2009).

Accordingly, more than 55% of therapeutic proteins on the market are produced using mammalian cells, with Chinese hamster ovary (CHO) cells being the predominant host (Walsh, 2014). This is reflected in Table 1.1 where seven of the ten best-selling biologics are produced in mammalian cells, and of those seven, six are from CHO cells. Similarly, CHO cells are the major host for Fc-fusion protein production—six Fc-fusion molecules on the market are expressed in CHO cells, while two more are expressed in human embryonic kidney (HEK293) cells and one in *E. coli* (Mei et al., 2013; Walsh, 2014). Nevertheless, other established mammalian cell lines such as NS0 murine myeloma and baby hamster kidney (BHK) cells have been, and to a certain extent, continue to be used to develop biologics (Estes and Melville, 2014; Wurm, 2004).

1.2.2 Stable gene expression

Large scale manufacturing of therapeutic recombinant proteins involves development of recombinant cell lines. Principally, the generation of stably expressing cell clones consists

of the transfection of the cell with a gene of interest cloned into a plasmid vector, followed by the selection of host cells that have the recombinant gene integrated into the cell genome, and the characterisation of the clones for production process (Figure 1.1A; Lai et al., 2013). To isolate cells that have successfully integrated the transgene, a selection marker is used in which the gene of interest is flanked with a gene coding for a vital enzyme. The glutamine synthetase (GS) and dehydrofolate reductase (DHFR) systems are commonly used in the industry. In both systems, selection takes place in culture media lacking the appropriate metabolite(s), namely glutamine for GS and glycine, hypoxanthine, and thymidine for DHFR (Hacker et al., 2008).

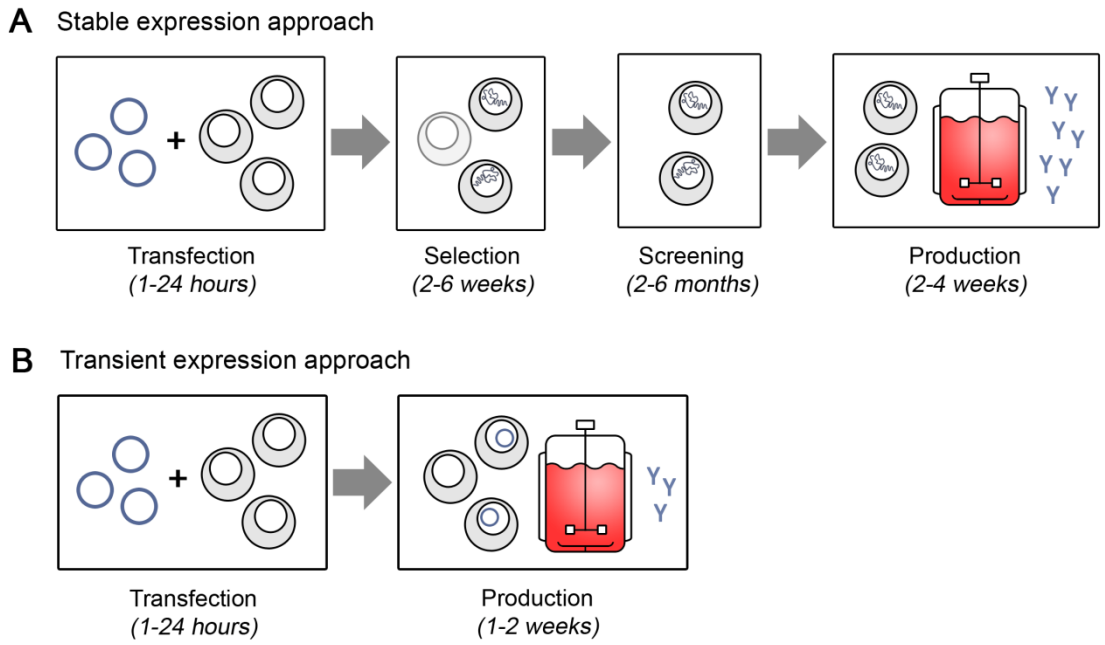

Figure 1.1: Stable expression and transient expression approaches to recombinant protein production. (A) The generation of stably expressing cell lines enables selection of clones with desirable production and growth characteristics for large-scale manufacturing of kilogram quantities. The plasmid DNA encoding for the protein of interest is randomly integrated in the genome of the host cell. **(B)** Transient gene expression represents an attractive alternative to stable expression in which small (gram) quantities of product can be generated in days rather than months. Unlike stably expressing cells, the recombinant DNA is maintained as an extrachromosomal unit within the cell nucleus that results in impermanent, lower productivity (Codamo et al., 2011; Lai et al., 2013).

Cells surviving selection are characterised by the integration of one or several copies of the expression vector into a transcriptionally active region of the host cell chromosomes. With both GS- and DHFR-based selection, the copy number of the integrated recombinant DNA can be amplified by exposure of the selected cells to increasing levels of methionine

sulphoximine (MSX) or methotrexate (MTX), respectively, that inhibits the enzymatic activity of the selection marker. This approach for cell line generation relies on the screening of hundreds to thousands of individual cell lines for the desired phenotype with regard to cell growth rate and recombinant protein productivity. The rare, superior cell lines are also eventually evaluated for the stability of protein production over time and in the absence of selection marker (Hacker et al., 2008). Consequently, the generation and selection of stable, high-performing mammalian cell lines is regarded as a major bottleneck in process development for the production of biologics (Browne and Al-Rubeai, 2007).

To shorten the overall developmental time frame for therapeutic proteins, an alternative approach using stable transfection pool technology is also being employed. The technology is similar to the stable gene expression approach except that the step of isolating and characterising clonal cell lines is omitted. Following selection and amplification, the stably transfected pools is used for bioreactor production of recombinant proteins, yielding gram quantities of recombinant proteins within two months post-transfection (Ye et al., 2010). In addition to the relatively high yields, stable pools offer an advantage over transient expression platforms in that their production does not require large amounts of high-purity, endotoxin-free plasmid DNA (Bertschinger et al., 2006; Schmid et al., 2001).

1.2.3 Transient gene expression

The vast expansion in the number of recombinant protein molecules needed both as therapeutic candidates and for research purposes demands a means of rapid high-throughput production of good-quality recombinant protein in mammalian cells (Figure 1.2). Transient gene expression (TGE) systems are an attractive alternative for rapid production of research-grade protein during the early stages of drug development due to the substantial time and resources associated with stable cell line generation. TGE methods can be employed to fast track the production of multiple biologics to perform biochemical study and early preclinical evaluation of drug candidates (Figure 1.1B).

TGE has been used for decades in cell biology laboratories for small-scale research and analytical purposes, and in the last 15 years or so the technique has been scaled-up for rapid supply of biopharmaceuticals. HEK293 cells were originally used for scaled up TGE owing to its well established system for episomal replication, whereas systems for episomal replication in CHO cells came later (Baldi et al., 2007; Geisse, 2009). Whilst

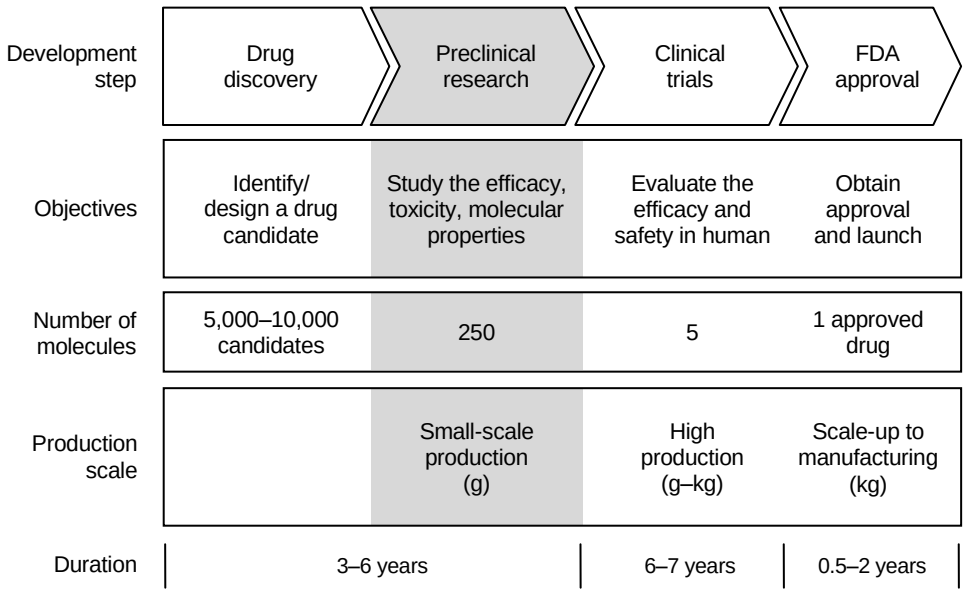

Figure 1.2: Research and development (R&D) process in the biopharmaceutical industry. The development of a new drug is a lengthy multi-stage process and costs $1.2 billion for one successful drug launch. A way to ease this predicament would be to reduce the time and the costs associated with the preclinical research by providing a rapid, high-throughput production platform including using transient gene expression technology. Adapted from PhRMA (2012).

HEK293 cells are regularly used in TGE for the early preclinical stage, it is desirable to align the early stage production with later manufacturing stage through use of the same cell line. For example HEK293 cells and CHO cells are known to have different post-translational modifications such as glycosylation, thereby reducing predictability of drug candidate development (Suen et al., 2010). Therefore, recent works on TGE processes have focused on CHO cells which is the industry's major cell line for biomanufacturing. To date, the highest TGE titre for HEK293 cells reported within the literature is 1.1 g/L (Backliwal et al., 2008), whereas more recent work using CHO cells achieved 2 g/L (Daramola et al., 2014).

A plethora of transfection methods have been described within the literature, with electroporation, cationic lipids and cationic polymers being the most commonly employed (Kim and Eberwine, 2010). With respect to the former, direct physical electroporation of expression plasmids into the cell nucleus ensures very high transfection efficiency and consistency. The recent advent of MaxCyte® technology (Fratantoni et al., 2003) has enabled large-scale electroporation for TGE, although the relatively expensive equipment and reagents practically limit its use. Similarly, despite their widespread use for microscale research applications, cationic lipids (commonly available in proprietary formulations such

as Lipofectamine®) are not used for large scale TGE due to its prohibitive cost. Consequently, the cationic polymer polyethylenimine (PEI) is arguably the most suitable method for large scale TGE. The technique requires low cost with relatively high transfection efficacy, and is now immerging as the leading TGE method for bioproduction (Backliwal et al., 2008; Daramola et al., 2014; Rajendra et al., 2011; Raymond et al., 2011).

1.2.4 Heterogeneity within cell populations

Heterogeneity in a clonal cell population is readily revealed by phenotypic marker analyses, which typically result in Gaussian distribution histograms for the abundance of a certain protein per cell in a population of cells (Figure 1.3). Within the histogram, the abundance of the protein in the cells with the highest and lowest expression level usually varies by 10 to 1000 orders of magnitude—disparities which could not be attributed to signal measurement noise (Brock et al., 2009; Chang et al., 2008). This variation in phenotype is often explained by genetic heterogeneity, which in turn, is thought to be promoted by genetic instability in cells (Huang, 2009; O'Callaghan and James, 2008). The term genetic, or genomic, instability refers to an increased mutation rate that alter the normal organisation and function of genes and chromosomes (Kuttler and Mai, 2006; Smith et al., 2010). In other words, the heterogeneity with respect to a given protein (or any other quantifiable functional parameters) between individual cells is due to alteration to the genes that control the protein expression in the cells.

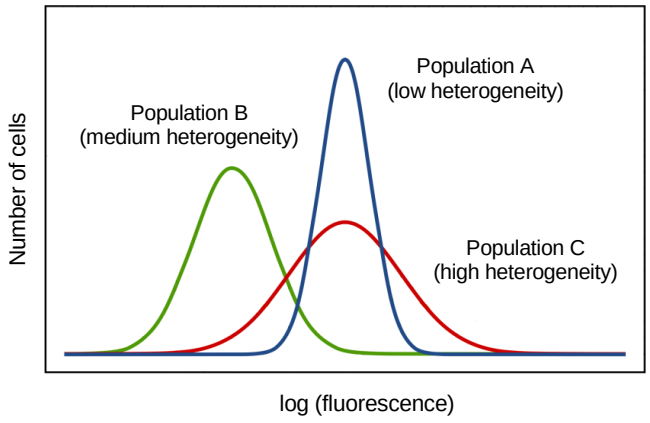

Figure 1.3: Heterogeneity in three clonal cell populations in fluorescent protein expression as manifested by the width of the Gaussian distributions. A cell line may also have two or more distinct subpopulations of cells (e.g. population A plus population B) having different levels of protein expression.

Although cell lines are generally required to maintain functional genetic stability throughout a period of 30–60 generations beyond the production of a manufacturer's working cell bank (Brown et al., 1992; Robinson and Chu, 2007), transformed mammalian cell lines are artificially immortalised cells, and thus have an inherent temperament of genetic instability (Barnes et al., 2006). With the exception of the study of Kaneko et al. (2010), it has been shown that mammalian production cell lines are genetically unstable, especially after long periods of cultivation and in the absence of selective agents (Derouazi et al., 2006; Heller-Harrison et al., 2009; Kim and Lee, 1999). CHO cells, for example, are known to have a very unstable karyotype due to chromosome rearrangements arising from homologous recombination and translocations, primarily in response to gene amplification (Kim et al., 2001; Yoshikawa et al., 2000).

In the absence of selective pressure, amplified genes localised to extrachromosomal double minutes are often lost by disproportionate segregation at mitosis (Kaufman et al., 1983; Wahl et al., 1982). Consequently, clonally-derived cell lines can display erratic and uncontrollable behaviour in culture such as variation in specific growth rate (Barnes et al., 2006), variations in protein modifications such as *N*-linked glycosylation (van Berkel et al., 2009) and loss of productivity (Heller-Harrison et al., 2009). With respect to the latter, cell lines may exhibit a relatively stable *q*P profile, gradual instability over numerous population doublings, or a dramatic loss of recombinant protein expression (Heller-Harrison et al., 2009). In some cases, the instability has also been shown to be transient, with productivity reverting to a specific level after a period of time in culture (Merritt and Palsson, 1993). Nevertheless, the rapid mutation capability of a host cell provides a mechanism to enrich clonal populations with cells that have survival advantages and the ability to grow autonomously, for instance with regard to "ease of adaptation" to a selective culture condition (O'Callaghan and James, 2008).

1.3 Mammalian Cell Factories for Recombinant Protein Production

1.3.1 Cellular constraints in mammalian cell hosts

Mammalian cell hosts, such as CHO, do not have a dedicated secretory phenotype and are poorly efficient to handle elevated trafficking load or complex secreted proteins (Dalton and Barton, 2014). At low levels of mRNA expression, cell-specific recombinant protein production and mRNA abundance are positively correlated. When mRNA expression is

high, this correlation often breaks down due to limiting translational and/or post-translational mechanisms, including for non-engineered proteins such as immunoglobulins (Barnes and Dickson, 2006; Carpentier et al., 2007; Le Fourn et al., 2014) and luciferase (Mead et al., 2009; Takahashi et al., 2011). The limitation can also be due to saturation of ER export machinery as reported by Hasegawa et al. (2011) for a CHO phenotype. In this particular case, the IgG synthesis and oxidative folding reactions exceeded the ER transport rate, resulting in accumulation of export-ready IgG in the ER lumen until a threshold concentration was reached to nucleate crystals.

The cellular bottlenecks are exacerbated in the case of DTE recombinant protein production using TGE systems as the host cells can be easily overloaded with transgenes and nascent recombinant polypeptides. As a result, for many recombinant proteins, productivity is unpredictably low, even MAb products in the same isotype/sub-class can display variable expression levels due to different translational and post-translational process rates (Pybus et al., 2014a,b). This can also be expected from artificial fusion proteins which have not coevolved; the two (or more) combined components are thought to have different folding and/or secretion requirements (Lee et al., 2007). On the other hand, stable transfectants capable of efficient DTE recombinant protein expression avoid/minimise unfolded protein response (UPR) induction by lowering the rate of recombinant gene transcription resulting in low stable expression system. Therefore, the amount of available recombinant mRNA often appeared to become the limiting event in stably producing cells, whereas translational and post-translational mechanisms are generally the bottlenecks in transient production of DTE proteins (Davies et al., 2011; Mason et al., 2012; O'Callaghan et al., 2010).

1.3.2 Protein translocation, folding and glycosylation

The transport of secretory proteins typically consists of several stages and involves various molecular chaperones along the processes. In eukaryotic cells, the transport into the ER lumen represents the first secretion step, and for the majority of these proteins (particularly those with more than 100 amino acids) this transport occurs during translation—a process referred to as co-translational translocation (Figure 1.4). The binding of the signal-recognition particle (SRP) to the sequence signal is thought to induce translational arrest to allow time for a ribosome-nascent chain complex to diffuse to the ER membrane. This would prevent premature folding of a secreted protein in the cytoplasm as well as enable the removal of the cleavage of the signal peptide, leading to the synthesis of properly

processed and translocated polypeptides. Moreover, the ER luminal Hsp70 and Hsp40 chaperones, namely immunoglobulin binding protein (BiP) and Sec63, respectively, have been shown to play important roles during co-translational translocation in mammalian cells. For example, the gating of the channel in the ER membrane is performed by BiP (Alder et al., 2005; Schäuble et al., 2012), whereas the initial insertion of several protein polypeptides into the Sec61 complex has been associated with Sec63 (Lang et al., 2012).

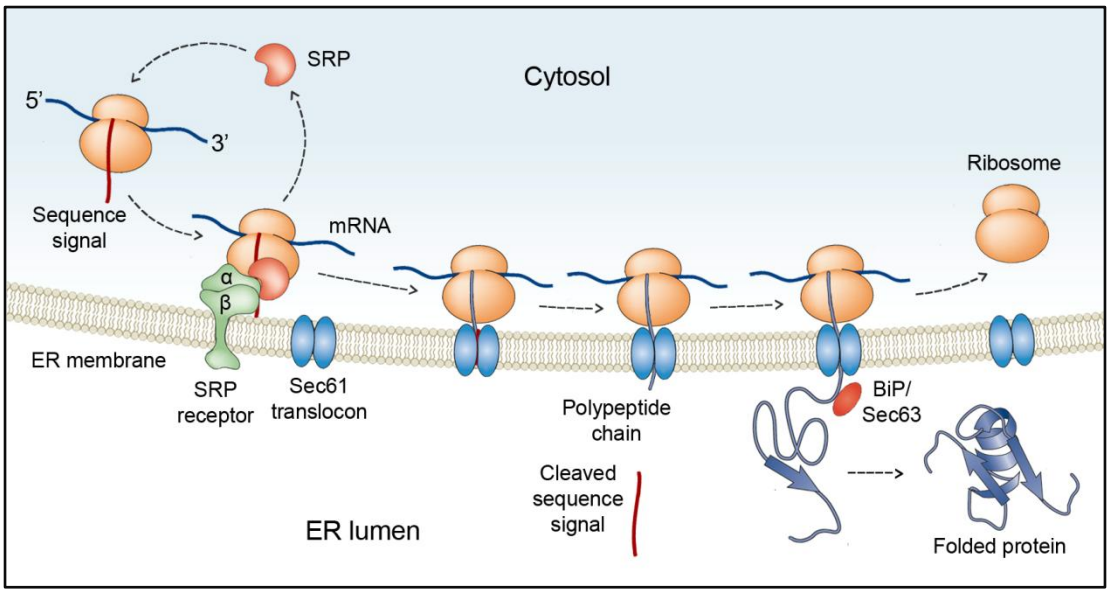

Figure 1.4: Co-translational translocation of secretory proteins into the ER. Translocation begins when a (signal-recognition particle) SRP binds to the signal sequence of a nascent polypeptide chain. The complex of SRP, nascent chain and ribosome then binds to the SRP receptor (consisting of α and β subunits) in the ER membrane. After SRP and its receptor dissociate from the nascent polypeptide chain, the signal sequence binds to the Sec61 translocon, thus opening the translocon gate that blocks the internal channel. The signal sequence is inserted as a loop into the central cavity of the translocon and the polypeptide chain elongates. Then the signal sequence is cleaved and rapidly degraded, while the chain continues to elongate until translation is completed and the chain is extruded into the ER lumen. ER lumen molecular chaperones BiP and Sec63 are known to assist the translocation process. Finally, the ribosome is released, the translocon gate shuts, and the secreted protein assumes its tertiary conformation. Adapted from Lodish et al. (2000) and Nyathi et al. (2013).

Furthermore, approximately two-third of therapeutic proteins and biologics candidates are proteins modified post-translationally by glycosylation (i.e., the attachment of oligosaccharide chains). This complex post-translational modification has multiple roles in the cell and is characterised by various glycosidic linkages, with *N*-linked and *O*-linked glycosylation being the most commonly observed types (Pandhal and Wright, 2010; Sethuraman and Stadheim, 2006). In the ER, glycosylation is performed to check the

status of protein folding, functioning as a quality control system to ensure that only correctly folded glycoproteins are exported to the Golgi complex. In the *trans* Golgi network, the sugar components can be bound by specific receptors to assist their delivery to the appropriate cellular destination. These sugar moieties can also potentially stabilise a protein (e.g. by enhancing the solubility), act as ligands for receptors on the cell surface to stimulate signal transduction pathways or mediate cell attachment, as well as regulate protein half-life (e.g. sialic acid may improve half-life) and biological activity (Walsh and Jefferis, 2006).

While glycosylation is generally characterised as a post-translational process, *N*-glycosylation often occurs co-translationally, in that the glycan is attached to the nascent polypeptide as it is being translated and translocated into the ER (Figure 1.5). *N*-linked glycans are essential in proper protein folding (90% of glycoproteins are *N*-glycosylated), in which the oligosaccharides trimming (i.e. sugar hydrolysis by glycosidases) is used to both monitor protein folding and specify when proteins must be degraded. The process involves the hydrolysis of two terminal Glc from the precursor glycan by glycosidases I and II, after which molecular chaperones calnexin and calreticulin bind to the nascent glycoprotein via the remaining Glc and assist the protein to fold. The final Glc is soon removed by glycosidase II, releasing the glycoprotein from the chaperones. If the protein fails to fold properly, glucosyltransferase transfers a Glc to the protein, and the chaperones again are bound to the protein to assist proper folding. This cycle of Glc removal and addition continues until the protein is properly folded, at which time it is not reglycosylated (Ellgaard et al., 2003; Roth et al., 2002). To this point, all *N*-linked glycoproteins have the same precursor glycan structure, and the glycoprotein is transported to the Golgi complex for further processing.

In addition to calnexin and calreticulin, various other molecular chaperones are involved in building efficient protein folding machinery within the ER. As nascent polypeptides enter the ER in the reduced state, disulphide bonds are rapidly formed to stabilise the folded structure of the protein. Protein disulphide isomerase (PDI) catalyses the disulphide-bond formation (i.e. oxidation) between cysteine residues intra- and intermolecularly, thereby allowing the nascent proteins to quickly find the appropriate configuration in their folded state. The reoxidation of PDI is accomplished by the ER oxidoreductins (ERO), although the lumen of the ER also provides an oxidising environment and substrate glutathione. Interestingly, the reduced form of PDI is able to catalyse a reduction of thiol residues, which is particularly important with terminally misfolded proteins that must be reduced before dislocation to the cytosol for proteasomal

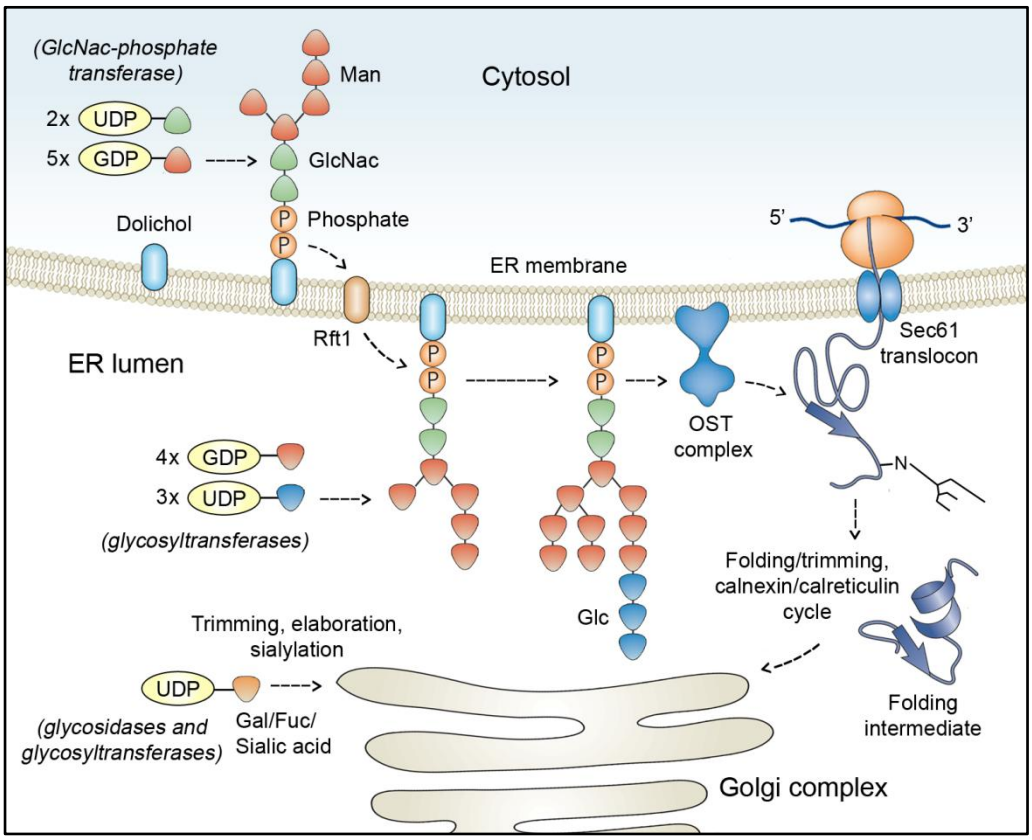

Figure 1.5: *N*-linked glycosylation in mammalian cells. Oligosaccharide chains attached through *N*-glycosidic linkages are derived from a 14-sugar precursor glycan composed of *N*-acetylglucosamine (GlcNAc), mannose (Man) and glucose (Glc) that are added in sequence onto dolichol. The first 7 sugars ($Man_5GlcNAc_2$) are provided by sugar nucleotides (UDP- and GDP-sugars) in the cytoplasm and bound to dolicholpyrophosphate (dolichol-PP). After the intermediate is completed, the entire unit is flipped into the ER lumen, after which Man- and Glc-P-dolichol molecules provide the remaining 7 sugars to make the $Glc_3Man_9GlcNAc_2$-PP-dolichol precursor glycan. The oligosaccharyltransferase (OST) then transfers the sugar moiety ($Glc_3Man_9GlcNAc_2$) to the nascent protein emerging from the Sec61 translocon, specifically the free amide group of the Asn (N) with sequence Asn-X-Ser/Thr. The glycan structure for all correctly folded glycoproteins that progress to the Golgi complex is $Man_9GlcNAc_2$. Adapted from Pandhal and Wright (2010).

degradation (Anelli and Sitia, 2008; Hatahet and Ruddock, 2009). BiP, the most abundant ER chaperone with multi-functions, has been demonstrated to cooperate with PDI and calnexin in oxidative folding and refolding of glycoprotein, respectively (Mayer et al., 2000; Stronge et al., 2001). On the other hand, ER mannosidase (ERManI) plays a key role in identifying proteins that are unable to fold correctly. The hydrolysis of mannose residues by ERManI would remove the glycoproteins from the reglucosylation and calnexin binding cycles, where they would be retro-translocated into the cytosol and eventually delivered to ERAD via ubiquitination, a process where ubiquitin binds to lysines on the protein and acts as a tag for proteasomal degradation (Frenkel et al., 2003; Tsai et al., 2002).

Glycan processing in the Golgi complex involves both trimming and addition of sugars to generate diverse oligosaccharides on individual glycoproteins for different functions (see below). As with precursor glycan biosynthesis in the ER, this maturation pathway to diversify the glycans is highly ordered. The final glycan structures can be broadly classified into three types; (i) complex oligosaccharides (multiple sugar types), (ii) high-mannose oligosaccharides (multiple mannose residues), and (iii) hybrid-branches which consist of both complex and high mannose oligosaccharides. *N*-glycosylation does not prohibit other types of glycosylation from occurring, as *O*-glycosylation normally occurs on glycoproteins that were *N*-glycosylated in the ER. Unlike *N*-glycosylation, *O*-glycosylation occurs post-translationally on Ser and Thr side chains in the Golgi complex and the glycans typically have much simpler oligosaccharide structures. Additionally, in *O*-glycosylation the sugars are added one-at-a-time to Ser or Thr residues, which is in contrast to *N*-glycosylation in which the precursor glycan is transferred *en bloc* to Asn (Hossler et al., 2009).

Proper glycosylation is a critical parameter for the manufacturing of glycoprotein therapeutics as it can affect protein stability, bioactivity, pharmacokinetics, immunogenicity and protein clearance in the circulation system (Walsh and Jefferis, 2006). For example, the clearance mechanisms for recombinant Factor VII have been associated with *N*-glycosylation (Appa et al., 2010) while the removal of *N*-glycosylation sites on recombinant erythropoietin (EPO) were shown to significantly reduce its *in vivo* activity (Delorme et al., 1992). Additionally, for both *N*-linked and *O*-linked oligosaccharide chains, sialic acid is typically the desired terminal sugar, with *N*-acetylneuraminic acid (Neu5Ac) and its *N*-glycolylneuraminic acid (Neu5Gc) derivative being the two most common forms (Borys et al., 2010; Chen et al., 2011). However, Neu5Gc is not typically expressed in humans (but abundantly expressed in most mammals) and has been suggested to elicit human immune responses (Irie et al., 1998; Virki et al., 2001). In this regard, anti-NeuGc antibodies is detectable in 85% of healthy humans (Zhu and Hurst, 2002), whereas high Neu5Gc levels on a chimeric CT4-IgG fusion protein were linked with faster protein clearance profiles *in vivo* (Flesher et al., 1995).

1.3.3 Intracellular protein aggregation

Protein aggregation continues to be a major complication in the development and manufacturing of biopharmaceuticals. The term "protein aggregation" may be referred to as a broad class of protein species with higher molecular weight (oligomers) than the

desired native or native-like protein monomer (Mahler et al., 2009; Philo, 2006). These protein aggregates may exhibit adverse effects such as immunogenicity, reduced or no biological activity, as well as production, formulation and storage problems (Wang et al., 2010). The potential for protein aggregate formation presents at all stages of mammalian-based biomanufacturing including during the protein folding reaction in the ER.

Studies suggest that the poorly populated protein folding intermediates are precursors in protein aggregation (Figure 1.6) where these intermediates, for example, can have higher flexibility and expose more hydrophobic patches than the folded state. In contrast, unfolded (nascent polypeptide chains) and completely folded proteins do not aggregate easily (Wang, 2005). Under ER stress (e.g. accumulation of unfolded proteins in the ER), proteins can be misfolded, which can then undergo physical interaction (e.g. hydrophobic interaction) or cross-linking reactions that lead to formation of protein aggregates (Ioannou et al., 1992; Kopito, 2000; Tyedmers et al., 2010). With respect to the latter, the most common cross-linking reaction is the intermolecular disulphide bond exchange/formation, although other non-disulphide cross-linking pathways such as oxidation, dityrosine formation and formaldehyde-mediated cross-linking have also been described (Wang et al., 2010).

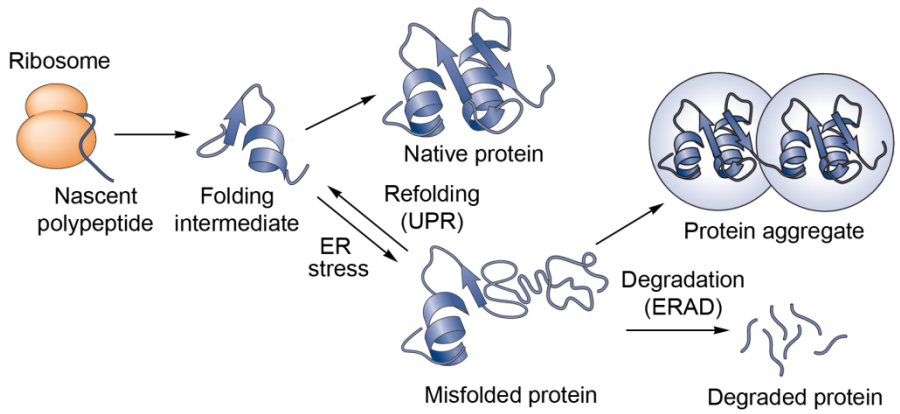

Figure 1.6: Overview of cellular protein folding, misfolding and aggregation. After a polypeptide is synthesised at the ribosome, it folds through an intermediate to its native, three-dimensional conformation. Inefficient ER-folding machinery to fold the synthesised proteins as they accumulate in the ER can cause ER stress and protein misfolding. The misfolded intermediates can be refolded to their native state via activation of UPR mechanism, or be degraded by ER-associated degradation (ERAD) and other cellular proteolysis systems that avert the accumulation of misfolded proteins. If the quality-control network is inundated (e.g. through increased amounts of aberrant proteins) aggregates can form. Adapted from Kopito (2000) and Tyedmers et al. (2010).

As mentioned earlier, the formation of disulphide bonds is a critical step in the maturation of therapeutic proteins in which they fold into their native conformation vital for their structure and activity (Cromwell et al., 2006; Zhang et al., 2011). However, in the case of MAbs, which require coordinated synthesis and folding of multiple polypeptides, *in vitro* studies have shown that the folding rate is relatively slow (Goto and Hamaguchi, 1981; Lilie et al., 1994), whilst comparative analyses revealed that different CHO cell lines (O'Callaghan et al., 2010) and different MAbs (Pybus et al., 2014a) have significantly different folding rates. To overcome the folding limitation, mammalian cells have a number of mechanisms that give them protection against protein misfolding and aggregation, primarily the UPR and ERAD (Dinnis and James, 2005). During the UPR, molecular chaperones such as BiP assist protein folding at high concentrations through binding of unfolded nascent polypeptides, which also helps in preventing aggregation by sheltering hydrophobic surfaces from forming unwanted intra- or intermolecular contacts (Kopito, 2000). However, depending on cell lines and protein products, the capacity of these chaperones can become overwhelmed, resulting in decreased protein folding efficiency. In many cases, inefficient ER-folding machinery to fold and assemble the synthesised proteins as they accumulate in the ER, often result in the formation of aggregates with anomalous intra- and intermolecular disulphide linkages (Gomez et al., 2012; Marquardt and Helenius, 1992; Schröder et al., 2002).

Accordingly, while disulphide bonds formed from the coupling of unpaired free thiols on cysteines are vital for the correct configuration of glycoproteins, they may also be the basis of protein misfolding and covalent aggregate formation. The presence of free thiols can also affect long-term stability of protein products where they are able to propagate aggregation by forming disulphide bonds during storage. This has been observed in the case of a lyophilised MAb stored at 30°C (Liu et al., 2005) and an Fc-fusion protein in liquid solution stored at –30°C (Wang et al., 2013). The formation of disulphide mislinkages may further promote physical aggregation of proteins, for example via hydrophobic interactions (Cabra et al., 2008; Wang et al., 2013), whereas correctly-folded disulphide-bonded proteins with no free cysteine residues can still undergo aggregation via disulphide exchanges through β-elimination (Wang, 2005).

Various approaches have been taken by manufacturers to control aggregates in protein-based products. Although it is technically possible to remove the aggregates at a later stage during downstream processing, preventing/minimising the accretion of aggregation-prone misfolded proteins is the first and most effectual intervention point to control protein aggregates (Cordoba-Rodriguez, 2008; Cromwell et al., 2006; Tyedmers et

al., 2010). MAbs susceptible to aggregation due to elevated free thiol can be prevented, for example, by adding low amounts of oxidising agent copper sulphate to the culture medium, which have been shown to result in a 10-fold decrease of free thiols and improved disulphide bond formation (Chaderjian et al., 2005). It is also evident that mutational changes can improve the stability of recombinant proteins which inclined to aggregate under stress-induced condition. In the study of Lu et al. (2008) for instance, when the serine residue 241 of an IgG4 produced in CHO cells was converted to proline, the hinge-mutant antibody appeared to be more resistant to freeze-thaw-induced aggregation compared to its wild type.

1.3.4 ER stress and the UPR

When the ER protein folding demand is excessive relative to the protein folding capacity, unfolded/misfolded proteins would accumulate in the ER, resulting in ER stress. To overcome ER stress, a cell activates several signalling pathways, collectively known as the UPR, to coordinate between the load of client proteins and the capacity of the protein folding machinery. The UPR promotes an adaptive cellular response to ER stress and re-establishes homeostasis in the ER, as well as induces an apoptosis programme. Double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK), activating transcription factor 6 (ATF6) and inositol-requiring enzyme 1 (IRE1) are the three main sensors of unfolded proteins in the ER and are therefore key transducers of UPR. PERK, ATF6 and IRE1 are activated when unfolded protein levels increase and BiP is released from the three proteins (Figure 1.7; Chakrabarti et al., 2011; Dinnis and James, 2005; Schröder and Kaufman, 2005).

The main function of the PERK pathway is to modulate translation following the accumulation of unfolded/misfolded proteins in the ER, although it is also responsible for the transduction of both pro-survival and pro-apoptotic signals. Dissociation of BiP from the N-terminus of PERK triggers dimerisation and autophosphorylation of the protein kinase domain. Activated PERK phosphorylates the α-subunit of eukaryotic initiation factor 2 (eIF2α) at Ser51, inhibiting the guanine nucleotide exchange factor eIF2B that recycles eIF2 to its active GTP-bound form. Lower levels of active eIF2 globally attenuates translation initiation as well as downregulates gene transcription (DuRose et al., 2009), hence reducing the influx of protein into the ER and reducing the folding load. Additionally, phosphorylated eIF2α paradoxically induces translation of the transcription factor ATF4 mRNA. ATF4 upregulation stimulates the expression of pro-survival functions including

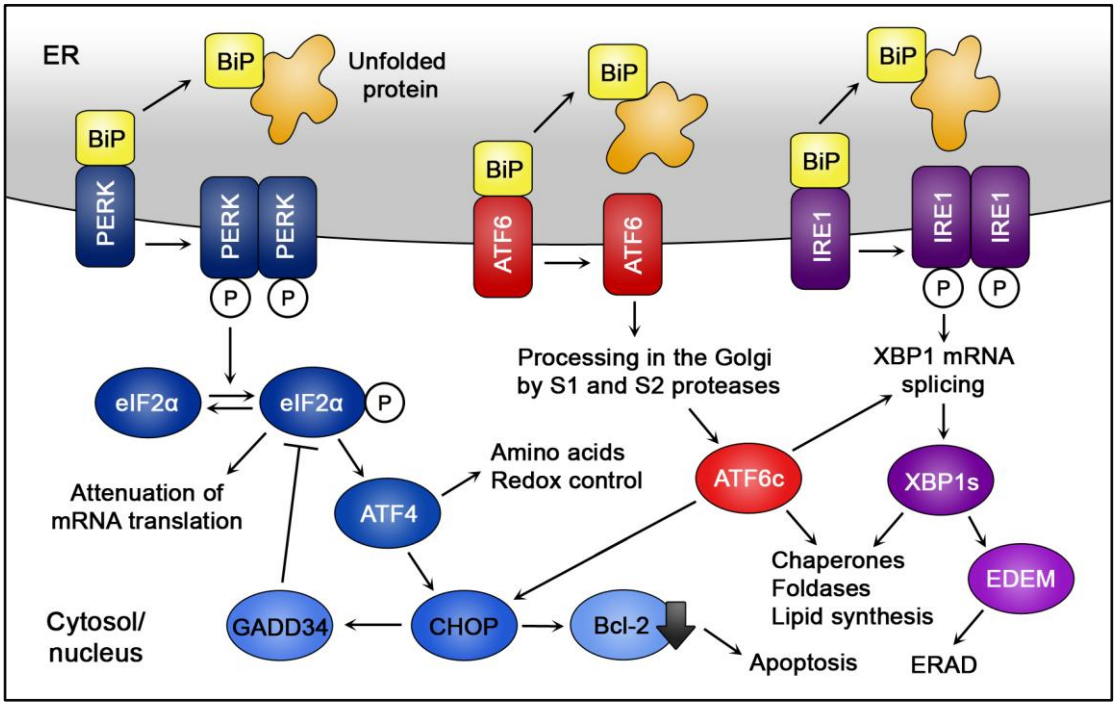

Figure 1.7: The major intracellular signalling during the UPR in mammalian cells. In unstressed cells, BiP binds to the lumenal domains of PERK, ATF6 and IRE1. Upon accumulation of unfolded/misfolded proteins in the ER lumen, PERK released from BiP dimerises, autophosphorylates, and phosphorylates eIF2α. eIF2α phosphorylation attenuates global translational initiation and paradoxically induces translation of ATF4 mRNA. ATF4 regulates pro-survival functions such as amino acid import and ER redox control and activates transcription of CHOP which leads to apoptosis via downregulation of anti-apoptotic Bcl-2 protein. CHOP also serves in a negative feedback loop ultimately reducing the phosphorylation of eIF2α. Secondly, ATF6 released from BiP translocates to the Golgi complex where for cleavage by S1 and S2 proteases to create a cytosolic domain ATF6c that travels to the nucleus to stimulate transcription of genes encoding ER chaperones and foldases. Similarly, IRE1 released from BIP dimerises to activate its kinase to initiate XBP1 mRNA splicing, yielding a strong transcriptional activator to further enhance transcription of chaperones and foldases. Additionally, key targets that need IRE1 or XBP1 pathway for induction include genes encoding functions in ERAD, particularly EDEM.

amino acid import and ER redox control (Harding et al., 2003) as well as the expression of pro-apoptotic factors through the transcription factor C/EBP homologous protein (CHOP). Whilst CHOP induces apoptotic cell death by repression of B-cell lymphoma 2 (Bcl-2) expression (McCullough et al., 2001), it also induces a negative feedback loop involving the direct dephosphorylation of eIF2α by promoting the expression of growth arrest and DNA damage-inducible protein 34 (GADD34; Novoa et al., 2001).

In addition to PERK, the general control non-derepressible-2 (GCN2) kinase, heme-regulated inhibitor (HRI) kinase, and protein kinase R (PKR) can also activate this pathway (i.e. phosphorylation of eIF2α) independently of ER stress, and therefore this branch of the

UPR is termed the integrated stress response (ISR; Ron and Walter, 2007). A well-characterised precedent for such mechanism is the eIF2α kinase GCN2. Besides inhibiting global rates of mRNA translation, GCN2 activation by amino acid starvation or exposure to oxidants and the subsequent eIF2α phosphorylation preferentially promotes the synthesis of ATF4 (see above). Similarly, low levels of iron lead to depletion of heme in erythroid cells, activating HRI autophosphorylation, which in turn phosporylates eIF2α (Han et al., 2001). PKR participates in an antiviral defense mechanism triggered by interferon, in which double-stranded RNA accumulation in virus-infected cells leads to PKR autophosphorylation and enhanced p-eIF2α (Dey et al., 2005).

The second branch of the UPR consists of a complex series of ATF6 translocation to the Golgi complex and irreversible proteolytic processing steps, eventually driving the upregulation of a pro-survival gene expression programme. Unlike the PERK and IRE kinase pathways, ATF6 activation does not involve phosphorylation but cleavage by serine protease site-1 (S1) and metalloprotease site-2 (S2) pr6789oteases. Cleaved ATF6 (ATF6c), in conjunction with other basic-leucine zipper (bZIP) transcription factors and required co-regulators, induces a transcriptional programme that increases ER chaperone activity and CHOP (Kokame et al., 2001; Yoshida et al., 2000). Furthermore, ATF6c can enter the nucleus and bind ER stress response elements (ERSE) of X-box binding protein 1 (XBP1) to increase its transcription which, after processing by activated IRE1α, also induces the expression of chaperones, foldases, lipid biosynthesis enzymes as well as other control elements (Yoshida et al., 2001).

On removal of BiP, IRE1 (like PERK) homodimerises and undergoes autophosphorylation of the kinase domain. The endoribonuclease activity of IRE1 removes a 26-base pair intron from XBP1 transcripts, which generates a potent 41 kDa transcription factor XBP1s (Yoshida et al., 2001). XBP1s then travels to the nucleus and binds to ERSE and unfolded protein response elements (UPRE) of a variety of ER chaperones and ERAD related genes. With regard to ERAD, ER degradation-enhancing α-mannosidase-like (EDEM) protein is known to directly involve in recognition and targeting of unfolded and misfolded proteins for degradation, in which the induction is dependent on XBP1 and stimulated by ATF6. Since activation of ATF6 precedes XBP1 splicing, it has been proposed that the UPR takes place in two phases, (i) ATF6-dependent induction of ER chaperones to facilitate protein folding, and (ii) both protein refolding and XBP1-dependent induction of ERAD to degrade unfolded/misfolded proteins (Hosokawa et al., 2003; Yoshida et al., 2003).

Chapter 2

Cell Line and Process Development for the Production of Biopharmaceuticals

This chapter provides an overview of cell line development and engineering strategies for the production of biopharmaceuticals. Specific emphasis is given to approaches utilised to improve or optimise the production platform, as well as computational tools to assist the development and engineering processes. The objective of this chapter is to contextualise the data presented in this thesis, with regard to their industrial applications.

2.1 Cell Line Development

2.1.1 Development of novel cell lines

CHO cells have the most established history of regulatory approval and thus are an attractive, safe option for biomanufacturing and remain the industry "workhorse" (Estes and Melville, 2014). Yet, the increasing demand and complexity of therapeutic proteins call for novel host cell types to meet the challenges of high productivity. In recent years, human-derived cell lines, particularly retina derived PER-C6® cells and amniocyte derived CAP® and CAP-T® cells, have emerged as an attractive alternative. These cell lines do not require amplification of the recombinant genes, thus providing enhanced stability in the absence of selection pressure. Additionally, their ability to confer human post-translational modifications reduces the potential immunogenic reactions against non-human epitopes (Schiedner et al., 2008; Swiech et al., 2012). The CAP-T cell line, developed for high transient gene expression (TGE), had been demonstrated to possess intrinsic capabilities to synthesise and secrete complex proteins that could not be expressed in sufficient yield or adequate quality in CHO and HEK293 cells (Fischer et al., 2012). PER-C6 cells are currently being used for production of many therapeutics in clinical trials, and with CAP cells they are expected to be a major production platform in the near future (Durocher and Butler, 2009; Swiech et al., 2012).

2.1.2 Harnessing genetic heterogeneity for super-producers

Genetic heterogeneity and clonal variation are the basis of the development of mammalian cell production systems. Whilst randomly arising clonal genetic heterogeneity contributes significant negative side effects, the increased genetic instability/mutation may also grant a benefit by producing genetically variant cells with improved phenotypes. This form of mutation may be supposed to encompass the regular adaptation of cell lines to suspension culture (Astley and Al-Rubeai, 2008), and to serum (Sinacore et al., 2000), cholesterol and protein-free media (Birch et al., 1994; Hartman et al., 2007) through growth in selective culture conditions. In this way, it can be hypothesised that increased genetic instability is an inherent trait to clonally-derived cell populations and is directly linked to cell metabolism as well as recombinant gene expression. Therefore, from a cell line development perspective, it is desirable to take advantage of this natural variation to generate advantageous cell hosts capable of supporting biomanufacturing processes. The host cell

factory is a major manufacturing parameter itself, and enhancement of parental cell functional capacity, as well as elimination of ineffectual variants from the cell population, enable the selection of production process attuned cell lines.

The maintenance of cell culture involves drastic, periodic contractions in population size when fresh media are inoculated during serial passage. A dilution ratio of below 0.1, which is a standard practice, diminishes the possibility that rare beneficial mutations are established (Wahl et al., 2002). In order to generate superior cell lines, desirable novel phenotypic traits are often selected through cloning and extensive screening, or by applying a suitable selective pressure. In the context of the latter approach, the genetic changes may characterise a developed adaptation of the overall population to a new surrounding condition, though in many studies it entails death of most of the cells and outgrowth of the surviving cells—allowing the formation of a genetically and phenotypically different sub-population that can tolerate that particular culture condition (Browne and Al-Rubeai, 2011; Hartman et al., 2007). In fact, even a slight advantage in cell survival (i.e., higher cell proliferation rate) will allow a clonal sub-population to overgrow a culture within a limited number of generations (Kromenaker and Srienc, 1994; Lee et al., 1991).

Similarly, an iterative process of extended batch culture eliminates inferior cells with low resistance to environmental stresses, and thus allows the selection of cells which exhibit improved and/or new characteristics (Browne and Al-Rubeai, 2011; Prentice et al., 2007). Successful examples of selective strategies include the isolation of variant CHO cell lines with the ability to be transfected with adenovirus (Condon et al., 2003), the ability to survive harsh culture conditions (Keightley et al., 2004; Prentice et al., 2007), and improved growth rate (Prentice et al., 2007) and specific productivity (Pichler et al., 2011). Moreover, genetically diverse cell populations can be induced artificially, for instance by temporarily suspending the DNA mismatch repair mechanism, which has been shown to generate high-osmolarity resistant cells (Liu et al., 2010). Although such host-cell enhancements may also be achieved via a more direct approach of genetic engineering (e.g., introduction of anti-apoptotic genes, knock-out of lactate dehydrogenase gene), this strategy requires extensive knowledge of the genes and cellular mechanisms involved and proves to be largely intractable (Dietmair et al., 2011; Dinnis and James, 2005; Lim et al., 2010).

Nevertheless, it is also important to note that the selection of functionally superior cell lines and managing the consequences of cell-specific functional/genetic instability is time-consuming and resource intensive, and with its own limitations. As mentioned earlier, all

inherent traits in mammalian cells may vary over time by means of point mutation, gene loss, etc. Consequently, the heritability of the desirable phenotypic/functional trait may be considered an issue; i.e. does the screened phenotypic/functional trait remain stably heritable over numerous sub-cultures? The rate at which genetic and its associated functional characteristic drift away is itself clone-specific, and therefore is likely to be the key parameter that needs to be evaluated.

2.2 Process Engineering Strategies

2.2.1 Optimisation of culture modality

The final recombinant protein titre is a function of two culture parameters, (i) the integral of viable cell density, IVCD and (ii) the cell specific productivity, *q*P. For fed-batch manufacturing processes, the optimal combination is a rapid accretion of productive host cells to high culture density for prolonged duration (Bibila and Robinson, 1995). Indeed, improvements in culture conditions for recombinant protein biomanufacturing have been remarkable in the last three decades, leading to over 100-fold increase in product yield. In the late 1980s, standard batch culture production runs had a production phase of 1 week and a peak cell density of around 2 million cells/mL, with a product titre of 50–100 mg/L. By comparison, current fed-batch production runs start with a lower cell density and last 3 weeks with a peak cell density of over 25 million cells/mL, while the product titre can reach 10–13 g/L (Huang et al., 2010). Much of these improvements have come from innovative and progressive developments of media and feeds of fed-batch operations, leading to more and healthier cells over a longer cultivation duration (De Jesus and Wurm, 2011).

In addition to the media compositions, researchers also found that specific MAb production rate of hybridoma cells increased when the cell division rate was slowed or halted by either small molecule inhibitors (Suzuki and Ollis, 1990) or nutrient limitation during chemostat culture (Miller et al., 1988). This apparently inverse correlation between *q*P and cell growth formed the basis of a controlled proliferation strategy to enhance recombinant protein expression (Fussenegger et al., 1998). The logic behind this approach is that a cell inevitably draws energy and resources away from recombinant protein production in order to generate new cellular biomass during mitosis. To circumvent the drawback of lower cell density, a "biphasic" cell culture is often employed. According to this scheme, cell growth is unrestrained during the first phase to accumulate viable cell density,

whereas during the second phase cell cycle is arrested to allow an extended period of high productivity (Dinnis and James, 2005).

Specific productivity may also be induced by increasing medium osmolarity, for instance through the addition of sodium chloride. Even though cell growth is depressed and apoptosis is induced at elevated osmolality, this can be resolved by the two-stage culture or by use of osmoprotectants as well as adaptation of cells to hyperosmotic pressure (Ryu et al., 2000). The main objective is to utilise a way of arresting cell division without inducing cell death and (in)directly interfere with the increase in recombinant protein productivity (Dinnis and James, 2005). Similarly, controlled proliferation by the use of lower culture temperature causes, in many cases, an increase in *q*P. This is particularly advantageous as hypothermic condition effectively extends culture longevity. However, several studies have shown that the enhancement effect of low temperature on recombinant protein production is not due to the G0/G1-phase growth arrest. Instead, the improvement is largely as a result of improved mRNA stability and elevated mRNA levels, and that a cell may still exhibit growth-associated productivity in hypothermic condition (Becerra et al., 2012; Fox et al., 2005; Kou et al., 2011).

2.2.2 Chemical chaperones

Chemical chaperones are a group of low molecular weight compounds with a common feature mimicking the chaperone function of molecular chaperones. These compounds are usually osmotically active, such as glycerol and dimethyl sulfoxide (DMSO), but other classes such as histone deacetylase (HDAC) inhibitors are also members of the chemical chaperone group. The exact mechanisms by which chemical chaperones function are not fully understood, but they are known to suppress ER stress, improve the folding capacity of ER, inhibit protein aggregation, enhance protein secretion, among others (Perlmutter, 2002). Owing to their positive effects, various osmolytes and HDAC inhibitors are in clinical use/trials for folding problem-related disease (e.g. Alzheimer's disease) and cancer treatments (Kim and Bae, 2011; Papp and Csermely, 2006).

In cell culture processes, chemical chaperones have been reported to grant a wide range of benefits as illustrated in Table 2.1. For example, HDAC inhibitors sodium 4-phenylbutyrate (PBA), sodium butyrate (NaBu) and valproic acid (VPA) are useful in enhancing the production of recombinant proteins in CHO cells (Hwang et al., 2011a; Sung and Lee, 2005; Wulhfard et al., 2010), and at the same time are able to reduce the

Table 2.1: Examples of effects of various chemical chaperones on cell culture processes.

Cell	Chemical chaperone	Effect	Reference
BHK, Neuro2a	DMSO, glycerol, TMAO	Reduce aggregate formation and cell death	Yoshida et al. (2002)
CHO	Betaine	Enhances cell viability and recombinant protein production	Follstad and Potter (2006)
CHO	Betaine	Suppresses protein aggregation and improves protein trafficking	Roth et al. (2012)
CHO, HEK293	PBA	Rescues cells from ER stress and apoptosis from misfolded myocilin	Yam et al. (2007)
CHO	DMSO	Enhances hepatitis B surface antigen production	Wang et al. (2007)
CHO	DMSO	Enhances expression of various exogenous genes	Liu et al. (2001)
CHO	DMSO	Enhances recombinant HBsAg intracellular accumulation	Ma et al. (2008)
CHO	Glycerol	Enhances recombinant M-CSF production	Liu and Chen (2007)
CHO	Glycerol	Reduces protein aggregation	Kim and Lee (1993)
CHO	Glycine betaine	Protects cells in hyperosmotic culture	Kim et al. (2000)
CHO	NaBu	Enhances recombinant human thrombopoietin production	Sung and Lee (2005)
CHO	NaBu	Enhances recombinant tPA production	Palermo et al. (1991)
CHO	NaBu	Minimises Neu5Gc content of fusion protein	Borys et al. (2010)
CHO	NaBu	Enhances production and minimises Neu5Gc content of antibody	Chen et al. (2011)
CHO	PBA, proline, glycerol, DMSO	Enhance production and minimise aggregation of FCA1 protein	Hwang et al. (2011a)
CHO	PBA	Restore the functionality of misfolded mutant low-density lipoprotein receptor	Tveten et al. (2007)
CHO	Trehalose	Suppresses recombinant antibody aggregation	Onitsuka et al. (2014)
CHO	VPA	Enhances recombinant mRNA and MAb levels	Wulhfard et al. (2010)
COS-1, human neuroblastoma	PBA	Suppresses Pael receptors aggregation and ER stress	Kubota et al. (2006)
E. coli	Arginine, sorbitol	Improve recombinant protein solubility in cells	Prasad et al. (2011)
E. coli	Proline	Inhibits P39A protein aggregation	Ignatova and Gierasch (2006)
HEK293	PBA	Reduces ER stress and prevents mutant HFE aggregation	de Almeida et al. (2007)
HEK293	PBA, VPA	Promote assembly of functional α7 AChRs	Kuryatov et al. (2013)
HT-1080, SW-1353	Glycerol, TMAO	Reduces thermolabile collagen II mutants intracellular accumulation	Gawron et al. (2010)
Hybridoma	Betaine	Protects cells from hyperosmotic stress	Berg et al. (1991)
Mice fibroblasts	PBA	Rescues trafficking incompetent mutant α-galactosidase A	Yam et al. (2007)

potentially immunogenic *N*-glycolylneuraminic acid (Neu5Gc) content of glycoproteins (Borys et al., 2010; Chen et al., 2011). The latter results are of particular importance as the chemical(s) can conveniently be employed to minimise or control Neu5Gc in the manufacturing processes, thereby improving product quality.

Similarly, enhancement in productivity can also be obtained using glycerol, DMSO and proline, even though they are probably more recognised for their ability in suppressing aggregate formation (Hwang et al., 2011a; Liu and Chen, 2007; Wang et al., 2007). On the other hand, methylamine osmolytes such as betaine and trimethylamine N-oxide (TMAO) are associated with a protective effect on cells, for instance against hyperosmotic condition (Kim et al., 2000), although betaine has also been reported to suppress aggregation and improve trafficking of factor VIII in CHO cells (Roth et al., 2012). Other chemical chaperones that have been shown to yield positive effects in cell culture include arginine, sorbitol (Arakawa et al., 2006) and trehalose (Onitsuka et al., 2014). Despite the versatility of general and specific effects, chemical chaperone treatment strategy is still underutilised. As low yield and aggregation have become recognised as critical issues in a growing number of complex and difficult-to-express proteins, the use of chemical chaperones is theoretically attractive because of their applicability to a broad range of cell culture processes. Moreover, the strategy has its practical advantage in which the chemical dosage can be easily and accurately titrated or applied at any time point of culture.

2.3 Vector and Cell Engineering Strategies

2.3.1 Vector engineering

Examples of constitutive promoters that are often used in mammalian expression systems include the human/mouse cytomegalovirus (CMV) promoter, the simian virus 40 (SV40) promoter and the non-viral elongation factor 1α (EF1α) promoter. As the nature/design of a promoter has profound effects on transgene expression level and overall productivity (Backliwal et al., 2008; Davies et al., 2011), developments in mammalian gene expression technology to increase transcriptional activity have occurred subsequently. For example, Running Deer and Allison (2004) cloned a Chinese hamster EF1α gene and its flanking sequences to generate CHEF1 vectors that had 6–35-fold higher expression levels than vectors utilising the CMV or EF1α promoter. On a similar note, Chatellard et al. (2007) cloned an enhancer from the mouse CMV immediate early 2 region (IE2), identified as a

potent expression-promoting element. The IE2's unique bi-directional promoter architecture can be used to efficiently express multi-chain proteins, which was shown to result in high recombinant protein production levels. Furthermore, recent work in this laboratory successfully created a library of numerous synthetic promoters designed to precisely regulate recombinant gene expression in CHO cells over a large transcription range (Brown et al., 2014).

Significant efforts have also been carried out to increase the recombinant gene copy number, especially in the transient gene expression (TGE) system. In the stable gene expression system, the transgene is inherited by the daughter cells upon mitosis, and thus allowing a rather constant transgene copy number in cells of succeeding generations. In contrast, in the TGE system the transgene copy number per cell is reduced during cell division, and it is therefore inevitable that the cell specific productivity decreases with subsequent generations. To overcome this issue, several episomal TGE systems have been developed that allow autonomous replication of the recombinant plasmid within mammalian cells. A widely used system is the Epstein-Barr virus (EBV) nuclear antigen 1 (EBNA1), which consists of cells constitutively expressing the EBNA1 protein combined with an expression vector containing the EBV origin of replication (OriP) for plasmid retention. Other episomal systems include the SV40 large T-antigen and the Py large-T antigen (*Epi*-CHO), where the former supports the episomal amplification of plasmids containing the SV40 OriP, while the latter supports the amplification of plasmids containing the Py virus and EBV OriPs (Kunaparaju et al., 2005; Piechaczek et al., 1999).

The ability of plasmids containing OriP to be maintained stably in antigen-positive cells makes such systems very efficient. For example, EBNA1 has a specific binding ability to the OriP, and with its ability to anchor EBV plasmid to chromosomal site, allows the plasmid to mediate replication and (non-random) segregation of the episome during division of the host cell. However, as EBNA1 is the only EBV protein necessary for plasmid maintenance but lacks any enzymatic activities (e.g. DNA helicase), the replication is mainly performed by the host cell replication apparatus (Van Craenenbroeck et al., 2000; Yates et al., 1985). This is consistent with studies which demonstrated that the total EBV chromosome did not require the replication-initiation function of OriP in latently infected proliferating cells. Therefore, the vital function of EBNA1 would be the maintenance (segregation) of the circular EBV chromosome, most likely in concurrence with the extrachromosomal maintenance function of OriP (Lee et al., 1999; Little et al., 1995).

2.3.2 Engineering of chaperone machineries

Several investigators observed that the expression of a number of ER chaperones were upregulated under stressful culture conditions where recombinant protein productions were enhanced (Dorner et al., 1989; Mazzarella et al., 1994). More recently, functional proteomic profiling of MAb-producing NS0 cell lines with varying production rate revealed a positive relationship between *q*P and the relative abundance of various molecular chaperones and foldases including immunoglobulin binding protein (BiP), protein disulphide isomerase (PDI) and endoplasmin (Dinnis et al., 2006; Smales et al., 2004; Stansfield et al., 2007). Similarly, high-producing CHO cells (Carlage et al., 2009) and HEK293 cells (Jones et al., 2005) were also found to have elevated expression of BiP. These observations show that high-level recombinant protein production links to upregulation of the ER-resident proteins involved in protein folding and secretion.

As such, the most ordinary strategy in secretion engineering to date has been one involving overexpression of ER chaperones like BiP and PDI (Table 2.2) which catalyses intra- or intermolecular disulphide bond formation. Unfortunately, while such an approach effectively increases heterologous protein production in insect cells (Ailor and Betenbaugh, 1998; Hsu and Betenbaugh, 1997) and yeast (Shusta et al., 1998; Smith et al., 2004), genetic upregulation of discrete chaperones in mammalian cells to increase *q*P has not always been successful. For example, BiP overexpression in mammalian cells has been reported to generically decrease the secretion of recombinant proteins it associates with (Borth et al., 2005; Dorner and Kaufman, 1994; Dorner et al., 1988; Dorner et al., 1992). Accordingly, the reduction of the amount of endogenous BiP, rather than its overexpression, was found to improve secretion of proteins in CHO and BHK cells (Brown et al., 2011; Dorner et al., 1988). Similarly, inducible overexpression of PDI did not improve antibody production in hybridomas (Kitchin and Flickinger, 1995) and caused reduced secretion of a disulphide rich Fc-fusion protein in CHO cells (Davis et al., 2000).

Despite previously published data, oxidation reactions catalysed by PDI that are vital for recombinant protein folding/assembly remain a rationale target. Certainly, specific reaction steps could limit the rate of folding of particular proteins (especially multimeric, difficult-to-express molecules) more than others. More recent reports by Mohan et al. (2007), Pybus et al. (2014a) and Borth et al. (2005) had specifically shown that PDI overexpression in CHO cells increased *q*P by 15–40%, where the latter specifically showed that this was achieved by decreasing intracellular retention of MAb heavy chain polypeptides. Additionally, the active site of PDI contains disulphides that become reduced

Table 2.2: Examples for co-overexpression and down-regulation of molecular chaperones and foldases in mammalian cells.

Host	System	Molecule	Recombinant protein	Results (titre)	Reference
BHK	Stable	BiP	Human factor VIII	50% decrease	Brown et al. (2011)
BHK	Stable	BiP down-regulation	Human factor VIII	2-fold increase	Brown et al. (2011)
BHK	Stable	BiP	Procine factor VIII	No effect	Brown et al. (2011)
BHK	Stable	BiP down-regulation	Procine factor VIII	1.3-fold increase	Brown et al. (2011)
CHO	Stable	BiP	MAb	35% decrease	Borth et al. (2005)
CHO	Stable	BiP	M-CSF	10–50-fold increase	Dorner et al. (1992)
CHO	Stable	BiP	Von Willebrand factor	3–4-fold increase	Dorner et al. (1992)
CHO	Stable	BiP	Factor VIII	5–30-fold increase	Dorner et al. (1992)
CHO	Stable	BiP down-regulation	Mutant tPA	2–3-fold increase	Dorner et al. (1988)
CHO	Stable	BiP	Firefly luciferase	Decreased secretion	Dorner and Kaufman (1994)
CHO	Transient	BiP	MAbs	0.9–1.5-fold change	Pybus et al. (2014a)
CHO	Stable	BiP	Antibody	2-fold increase	Nishimiya et al. (2013)
CHO	Stable	Calnexin and calreticulin	Thrombopoietin	1.9-fold increase	Chung et al. (2004)
CHO	Transient	CypB	MAbs	0.8–1.3-fold change	Pybus et al. (2014a)
CHO	Stable	SRP14	MAbs	Increased secretion	Le Fourn et al. (2014)
CHO	Stable	ERO1Lα and XBP1s	MAbs	5.3–6.2-fold increase	Cain et al. (2013)
CHO	Transient	ERO1L	Antibody	37% increase	Mohan and Lee (2010)
CHO	Transient	ERO1L and PDI	Antibody	55% increase	Mohan and Lee (2010)
CHO	Stable	PDI	Thrombopoietin	No effect	Mohan et al. (2007)
CHO	Stable	PDI	Antibody	15–27% increased secretion	Mohan et al. (2007)
CHO	Transient	PDI	MAbs	0.7–1.1-fold change	Pybus et al. (2014a)
CHO	Stable	PDI	MAb	37% increase	Borth et al. (2005)
CHO	Stable	PDI	Interleukin 15	No effect	Davis et al. (2000)
CHO	Stable	PDI	TNFR:Fc	Decreased secretion	Davis et al. (2000)
CHO	Stable	PDI	MAb	No effect	Hayes et al. (2010)
COS-1	Transient	ERO1Lβ	Antibody	1.5-fold increase	Nishimiya et al. (2013)
COS-1	Transient	PDI	Antibody	2-fold increase	Nishimiya et al. (2013)
Rat hepatoma	Transient	PDI	Apolipoprotein B	37% increased secretion	Grubb et al. (2012)
Rat hepatoma	Transient	ERp57	Apolipoprotein B	33% decreased secretion	Grubb et al. (2012)
Rat hepatoma	Transient	ERp72	Apolipoprotein B	33% decreased secretion	Grubb et al. (2012)

Table 2.3: Examples for co-overexpression of UPR and anti-apoptotic proteins in mammalian cells.

Host	System	Molecule	Recombinant protein	Results (titre)	Reference
BHK	Stable	HBx	Antibody	5.3-fold increase	Jin et al. (2010)
BHK	Stable	XBP1s	Antibody	2.3-fold increase	Jin et al. (2010)
BHK	Stable	XBP1s and HBx	Antibody	14-fold increase	Jin et al. (2010)
CHO	Stable	ATF4	Human antithrombin III	2-fold increase	Ohya et al. (2008)
CHO	Stable	ATF4	MAb	2.4-fold increase	Haredy et al. (2013)
CHO	Transient	ATF6c	MAbs	0.8–1.5-fold change	Pybus et al. (2014a)
CHO	Stable	Bcl-2	Antibody	2–3-fold increase	Kim and Lee (2000)
CHO	Transient	Bcl-x_L	Fusion protein	70–270% increase	Majors et al. (2008)
CHO	Stable	Bcl-x_L	ErbB2	Increased expression	O'Connor et al. (2009)
CHO	Stable	Bcl-x_L	EPO	No effect	Kim et al. (2011b)
CHO	Stable	Bcl-x_L	EPO	10–30% increase	Han et al. (2011)
CHO	Stable	Bcl-x_L	MAb	90% increase	Chiang and Sisk (2005)
CHO	Stable	30Kc6	MAb	3.8-fold increase	Wang et al. (2012)
CHO	Stable	GADD34	Human antithrombin III	1.4-fold increase	Omasa et al. (2008)
CHO	Transient	XBP1s	MAb	28% increase	Codamo et al. (2011)
CHO	Stable	XBP1s and ERO1Lα	MAbs	5.3–6.2-fold increase	Cain et al. (2013)
CHO	Transient	XBP1s	MAbs	0.8–1.7-fold change	Pybus et al. (2014a)
CHO	Stable	XBP1s	SEAP	6-fold increase	Tigges and Fussenegger (2006)
CHO	Stable	XBP1s	Secreted α-amylase	4-fold increase	Tigges and Fussenegger (2006)
CHO	Stable	XBP1s	MAb	1.4-fold increase	Becker et al. (2008)
CHO	Stable	XBP1s	MAb	1.4-fold increase	Becker et al. (2010)
CHO	Stable	XBP1s and XIAP	MAb	2.1-fold increase	Becker et al. (2010)
CHO	Stable	XBP1s	MAb	No effect	Ku et al. (2008)
CHO	Transient	XBP1s	EPO	2.5-fold increase	Ku et al. (2008)
CHO	Transient	XBP1s	Factor VIII	No effect	Campos-da-Paz et al. (2008)
COS-1	Transient	CHOP (GADD153)	Antibody	2-fold increase	Nishimiya et al. (2013)
HepG2	Transient	XBP1s	Factor VIII	No effect	Campos-da-Paz et al. (2008)
NS0	Stable	XBP1s	MAb	No effect	Ku et al. (2008)
NS0	Stable	XBP1s	Interferon γ	No effect	Ku et al. (2008)
NS0	Transient	XBP1s	EPO	2-fold increase	Ku et al. (2008)

during catalysis and their reoxidation is thought to be catalysed by the oxidoreductase enzyme ER oxidoreductin 1 (ERO1; Sevier and Kaiser, 2006). ERO1 overexpression therefore may create an altered oxidising environment suitable for MAb (or other disulphide-containing proteins) production within the CHO cells (Cain et al., 2013).

2.3.3 Engineering of the UPR and apoptosis

Given the complexity of mammalian cellular regulation, engineering a single component of protein folding may not always lead to the desired results especially if the secretory pathway suffers from several limitations (Delic et al., 2014). A strategy that can modulate the secretory machinery in a more global manner holds promise to increase concentrations of several chaperones in a functionally meaningful ratio. Accordingly, basic leucine zipper (bZIP) transcription factors ATF4 (activating transcription factor 4), ATF6 (activating transcription factor 6) and XBP1 (X-box binding protein 1) can be more effective gene targets compared to particular ER chaperones (Table 2.3). XBP1 for example is a key regulator of the UPR where expression of the protein induced a wide spectrum of secretory pathway genes including BiP and PDI (Lee et al., 2003). XBP1 is also known to physically expanded the endoplasmic reticulum and the Golgi compartments, resulting in an increase in the overall production capacity (Shaffer et al., 2004; Tigges and Fussenegger, 2006).

Several recent approaches have been described to introduce heterologous XBP1 into mammalian host cells resulting in higher production rates of various recombinant protein products (Becker et al., 2008; Ku et al., 2008; Pybus et al., 2014a). It is interesting to note that the study of Ku et al. (2008) highlighted that overexpression of XBP1s had no detectable effects on EPO productivity in stable CHO cell lines but significantly enhanced transient production in EPO-saturated CHO cells. Pybus et al. (2014a) on the other hand had shown, using a panel of difficult-to-express MAbs, that the effects of XBP1s were more pronounced in cells with limiting folding and assembly reactions, and that no effect was observed with an easy-to-express (ETE) MAb. Together, these data suggest that overexpression of XBP1 can be more beneficial when the accumulated level of the nascent recombinant polypeptide exceeds the secretory capacity of host cells. Overexpression of UPR transcription factors AFT4 and its target CHOP and GADD34 have also been reported to positively affect recombinant protein production—likely by promoting translation through dephosphorylation of eIF2α, thereby enhancing the protein secretion (Nishimiya et al., 2013; Ohya et al., 2008; Omasa et al., 2008).

Despite the positive effects, induction of apoptosis upon transfection of the UPR-related genes could be a drawback for this cell engineering strategy (Becker et al., 2010; Pybus et al., 2014a). Becker et al. (2010) overcame this issue by co-expressing the caspase-inhibitor x-linked inhibitor of apoptosis (XIAP) with XBP1s, resulting in both improved *q*P and CHO cell survival. Similarly, apoptosis induced by various stresses such as hyperosmolarity, nutrient depletion and chemicals (e.g. DMSO and butyrate) can be suppressed by overexpression of B-cell lymphoma 2 (Bcl-2) or its Bcl-x_L counterpart, leading to extended culture longevity and higher recombinant protein production (Chiang and Sisk, 2005; Kim and Lee, 2000; Majors et al., 2008). Bcl-2 and Bcl-x_L proteins, both commonly known to be apoptotic inhibitors, are also thought to play a vital function in autophagy through inhibitory interaction with Beclin-1 protein (Levine et al., 2008; Thorburn, 2008).

2.4 Computational Approaches for Understanding and Improving Biopharmaceutical Productions

2.4.1 Modelling of mammalian cells and cell culture processes

Mathematical models are beneficial in elucidating biological phenomena (O'Callaghan et al., 2010; Stockholm et al., 2007), organising disparate information such as in the case of high-throughput data (Hatzimanikatis et al., 1999; Schilling et al., 1999) or rationally identifying optimal strategies to improve a production process (Kontoravdi et al., 2007; Pybus et al., 2014a). Reported models in the literature show various degrees of biological process and mathematical complexity, though in mammalian cell culture substantial works are directed towards kinetic modelling. A kinetic model is usually represented by a set of differential mass balance equations that are integrated over time, either to a pre-defined time duration or a steady state. Kinetic models therefore can account for dynamic behaviour and are able to represent the complex biochemistry of cells in a more complete way (Almquist et al., 2014).

Conventionally cell culture models can be divided into two categories, namely, "unstructured" and "structured" models (Figure 2.1). The former describes culture dynamics encompassing cell growth/death, nutrient consumption, metabolite production and recombinant protein synthesis, and their dependence on various culture environmental factors (Jang and Barford, 2000; Pörtner et al., 1996). In contrast, the latter is based on single cell behaviour and attempt to explicitly describe the detailed phenomena of

intracellular metabolism (Bibila and Flickinger, 1991; McKinney et al., 1995). The unstructured and structured models can be further categorised into "unsegregated" or "segregated" biophases (Figure 2.1) where the latter can be applied for population balance modelling (Kromenaker and Srienc, 1994; Lee et al., 1991). In the recent years, several hybrid models for MAb production have been also developed in which the unstructured model of average cell population behaviour is linked to the structured synthetic model of MAb (Ho et al., 2006; Kontoravdi et al., 2005).

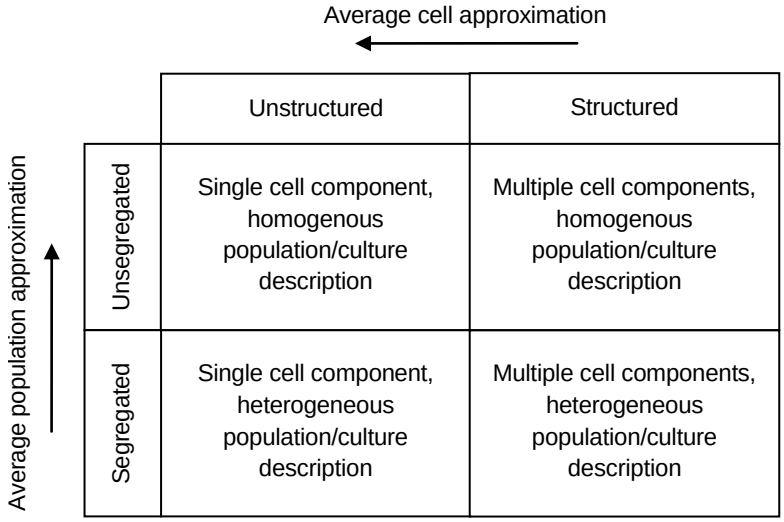

Figure 2.1: Model classification for cell culture systems. The term “structured” designates a formulation in which the cellular model is composed of multiple components (e.g. endoplasmic reticulum vs. Golgi complex), while the term “segregated” indicates explicit accounting for the existence of heterogeneous cell variants within a population (e.g. producing vs. non-producing cells). Adapted from Bailey (1998) and Sidoli et al. (2004).

Once a model is developed, they may be used for understanding and predicting the effects of adding, removing, or modifying molecular components of a mammalian cell factory, as well as for supporting/optimising the design of a culture process (e.g., feed strategies, medium formulation). With respect to the former, the analysis of fundamental cellular or metabolic processes and their ensuing elucidation (Karra et al., 2010; O'Callaghan et al., 2010) may lead to the design of more efficient cell culture systems, both from a bioprocess control and cell engineering perspectives. Importantly, several recent studies have successfully tested *in silico* predictions by implementing the suggested process engineering (Kiparissides et al., 2011; Kontoravdi et al., 2010) or cell engineering (Ho et al., 2012; Pybus et al., 2014a) strategies *in vivo*, resulting in improved recombinant

protein production. However, compared to chemical processes, biotechnology processes involve numerous reactions of the host cells with many of them being unknown or too complex for model application. Therefore, challenges still remain before kinetic modelling reaches the level of maturity to be a routine practice in the bioindustry.

2.4.2 Process development using design-of-experiment approach

Traditionally, experiments are conducted to establish the effect of a single input factor upon one output response—known as one-factor-at-a-time (OFAT) method. The design-of-experiment (DOE) technique develops models in a systematic way in order to obtain the maximum information from an experimental apparatus being modelled whilst minimising the number of experiments (i.e., time and resources) required (Franceschini and Macchietto, 2008). Several types of DOE methodology exist, each one has specific functions and benefits. Among the commonly used methodology in process development are factorial designs and response surface designs. The former are used for *screening* the factors affecting the responses, as well as for identifying the effects of interactions between the factors. On the other hand, the response surface methodology (RSM) consists of more advanced DOE designs aimed at *mapping* responses across a defined range of input factors (Figure 2.2). Thus, the designs capable of predicting the responses of a given coordinate of the design space that is not experimentally tested (Mandenius and Brundin, 2008).

Despite the obvious benefits, the DOE methodology is not fully adopted in the bioindustry. A recent report, for example, showed that only about 5% of all biological assays were developed using DOE approach, even though the general anticipation was that considerable cost reduction (3X) could be achieved by employing DOE. The lack of use of the DOE method is mainly due to two reasons. Firstly, many biologists are not well familiarised with the technique which involves complex statistics, even with the help of statistical packages. Secondly, the significantly more complex biological processes compared to chemical processes lead to many researchers believing that DOE methods can lead to "illogical" biological recommendations (Comley, 2009). Nevertheless, reports have shown that the FDA actively encourages biopharmaceutical companies to adopt a DOE driven methodology towards process development and validation including in its Process Analytical Technology (PAT) initiative (Baseman, 2012; FDA, 2004).

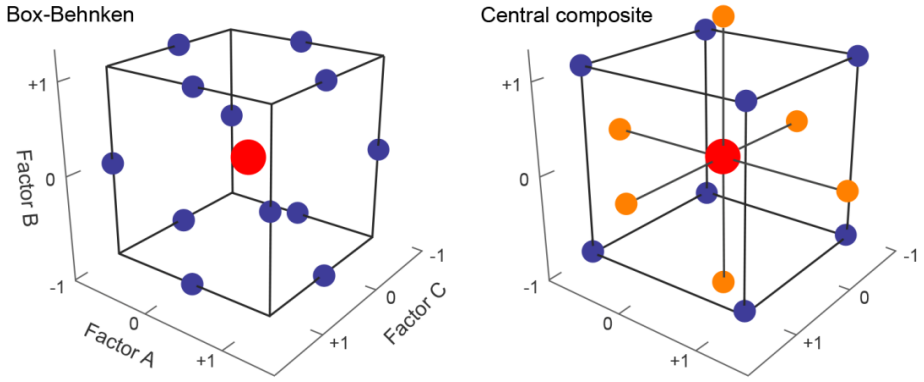

Figure 2.2: Response surface methodology (RSM) for three factors with three levels using the Box-Behnken and central composite designs. By conducting experiments at the designated points, RSM allows navigation inside and around the design space to study effects of interactions between the factors, and identify combinations of factors for optimal response(s). The centre point provides the estimation of curvature and pure error and axial points that allow accurate modelling of the process responses. For central composite design, the distance of the axial points from the centre point can vary and specify the nature of the experiment.

One common use of DOE in the production of recombinant proteins from mammalian cells is the optimisation of basal variables underpinning PEI-mediated transfection (Bollin et al., 2011; Daramola et al., 2014; Mozley et al., 2014; Thompson et al., 2012). The method enables the identification of optimal transfection conditions for high transfection efficiency and minimal polyplex cytotoxicity, and reveals that choice of host cell line is a critical parameter that should rationally precede the transient production platform. Other applications of DOE in process development include the optimisation of medium formulations to increase cell growth and production (Chun et al., 2003; Jiang et al., 2012; Kim and Lee, 2009; Sandadi et al., 2005) as well as the investigation of the combined effects of medium components on MAb *N*-glycan processing (Grainger and James, 2013). Such applications could further boost the multi-gram titres currently obtainable in the bioindustry and at the same time achieve high product qualities. Additionally, the DOE methodology has also been shown to be valuable in assessing scaled down culture systems such as cultiflask (Strnad et al., 2010) and microbioreactor (Legmann et al., 2009) as high-throughput tools for process development.

2.5 Thesis Overview

Very low productivity is often associated with DTE recombinant proteins produced in mammalian cells. In Chapter 4, the Sp35:Fc cellular production is first studied to identify

the limiting processes, and the impact on the UPR and protein aggregation. The study provides a detailed understanding of the cellular constraints normally associated with DTE protein, whilst the data obtained enables the identification of rationale engineering targets to improve the production process. In Chapter 5, we subsequently utilise mathematical modelling approach to systematically and quantitatively describe the cellular kinetics and evaluate the relative impact of each cellular process. A model framework for Sp35:Fc synthesis and secretion processes is developed based on the empirical data obtained in Chapter 4. The *in silico* analysis reveals cell functions in maintaining cellular homeostasis and importantly predicts the relative significance of different engineering targets.

In Chapter 6, we explore the possibility of exploiting the genetic and functional heterogeneity within the CHO cell population for improved Sp35:Fc production by isolating numerous clonal variants. The phenotypic/functional variations are characterised at different generations over prolonged culture, leading to the identification of several clones with inherent, superior characteristics for Sp35:Fc production. The data provides important knowledge for CHO host cell choice to express difficult to produce recombinant proteins. Chapter 7 describes a novel vector engineering approach specifically beneficial for DTE protein production. The system incorporates ER stress elements into the SV40 vector and manipulates endogenous and exogenous UPR transactivors. We show that the synthetic amplifier circuit produced a dynamic recombinant protein expression that permits rapid cell growth and high Sp35:Fc production in fed-batch culture.

In order to resolve the Sp35:Fc expression bottlenecks, Chapter 8 investigates the effects of various cell and process engineering strategies including the co-expression of functional proteins and treatment of cells with chemical chaperones. The interactions between different strategies are also studied including using DOE methodology. Based on the preliminary data, we formulate integrated cell and process engineering strategies for fed-batch culture in which up to a six-fold increase in volumetric titre was achieved. Additionally, the integrated strategies are tested on two clones obtained in Chapter 6 as well as on the production of SEAP as a model ETE protein. In the final chapter of this thesis (Chapter 9), work within previous chapters is summarised as a whole, followed by a brief description on the prospect of future work.

Chapter 3

Materials and Methods

The materials and methods used for experiments described throughout this thesis are described in this chapter. All materials used were of the highest purity available, unless otherwise stated.

Acknowledgements

The plasmid vector pCMV SEAP2-Control (Appendix A) encoding secreted alkaline phosphatase (SEAP) under the control of a CMV promoter was provided by Dr Nathan West. The plasmid was modified from pSEAP2-Control vector (Clontech, Mountain View, CA) which originally used SV40 promoter as described in West (2014).

3.1 Mammalian Cell Culture

3.1.1 Cell line and cell culture

A CHO-S cell line was provided by Biogen Idec (Cambridge, MA). The cells were cultured in CD CHO medium (Life Technologies, Paisley, UK) supplemented with 8 mM L-glutamine (Life Technologies) within Erlenmeyer shake flasks (Corning Incorporated, Acton, MA) maintained at 37°C under 5% CO_2 and shaken at 140 rpm. Cells were seeded at 2×10^5 viable cells mL^{-1} and were sub-cultured every 4 days, unless otherwise stated. Spent medium was removed by pelleting the cells using centrifugation at 200×g for 5 min. Cell culture viability and viable cell concentration were measured using a Vi-CELL™ Cell Viability Analyser (Vi-CELL XR; Beckman Coulter, Brea, CA) using Dulbecco's phosphate buffered saline (DPBS; Life Technologies) as diluent where necessary. The growth/proliferation rate (μ; day^{-1}) and the integral of viable cell density (IVCD; cell day mL^{-1}) were determined using Equations 3.1 and 3.2, respectively:

$$\mu = \frac{\ln(X_2/X_1)}{\Delta t} \tag{3.1}$$

$$\mathrm{IVCD} = \left(\frac{X_2 + X_1}{2} \times \Delta t\right) + \mathrm{IVCD}_{2-1} \tag{3.2}$$

where X_2 and X_1 are the viable cell concentration at second time point and first time point, respectively, and t is the time.

3.1.2 Cryopreservation and cell revival

Cells were harvested at mid-exponential culture (day 4), centrifuged at 200×g for 5 min and supernatant was removed. Cell pellets were re-suspended in cold culture media supplemented with 10% v/v dimethyl sulfoxide (DMSO; Sigma-Aldrich, Poole, UK) to 6.67×10^6 cells mL^{-1} at 1.5 mL per vial. Cryovials were frozen to –80°C for 24 h before

transferring to cryostats containing liquid nitrogen (–196°C) for long term storage. Cells were brought up from liquid nitrogen storage by thawing the cells in a waterbath at 37°C and resuspending in warmed culture media.

3.2 Plasmid DNAs

3.2.1 Gel electrophoresis

1% w/v agarose gel was used to visualise or separate DNA fragments of different sizes. Each gel block was prepared by dissolving 0.4 g of agarose into 40 mL of 1X TBE buffer by heating the mixture up, followed by the addition of 10 µL of ethidium bromide. After the gel was set, 100–500 ng of DNA was loaded into each well using 5X loading buffer (Bioline Reagents, London, UK) with a final volume of 10 µL. Nuclease-free water (Qiagen, Crawley, UK) was used as diluent where necessary and HyperLadder™ I (Bioline Reagents) was used as a molecular weight marker. The gels were run at a constant voltage of 110 V for 1.5 h and then visualised using a UV transilluminator (ImageQuant™-RT ECL™ Imager; GE Healthcare, Amersham, UK).

3.2.2 Plasmid DNA construction

The Sp35:Fc gene was provided by Biogen Idec and inserted into pcDNA™3.1(+) vector (Life Technologies) at the *Nhe*I and *Eco*RI restriction sites. Briefly, digestion was performed on the GeneArt® pMA cloning plasmid containing the Sp35:Fc gene and the pcDNA3.1 vector at *Nhe*I and *Eco*RI restriction sites using the respective enzymes and Buffer C from Promega (Southampton, UK) according to the manufacturer's protocol. The Sp35:Fc gene was separated from the pMA plasmid backbone and the linearised pcDNA3.1 vector was separated from the small DNA fragment using gel electrophoresis. The Sp35:Fc gene and linearised pcDNA3.1 vector were extracted from the gel slices using QIAquick® Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. The Sp35:Fc gene was then inserted into the linearised pcDNA3.1 vector by *Nhe*I and *Eco*RI sticky-sticky ligations using T4 DNA ligase (Promega) according to the manufacturer's protocol. To verify the plasmid construct (Figure 3.1), the purified product (see below) was subjected to single and double digests followed by gel electrophoresis prior to visualisation. The recombinant gene insert was confirmed by DNA sequencing using nine sequencing primers designed using Vector NTI® software (Life Technologies).

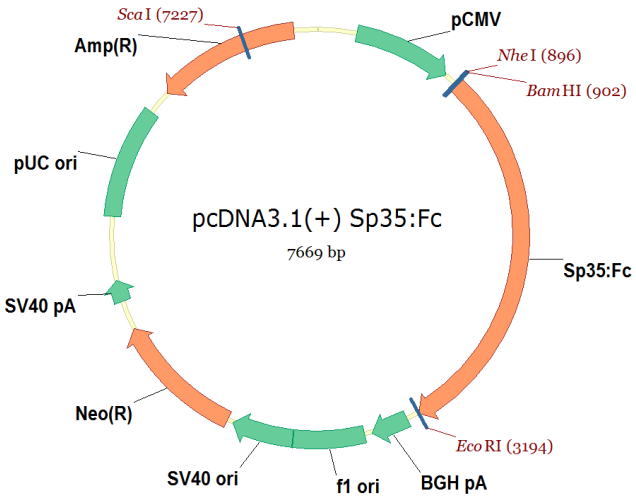

Figure 3.1: pcDNA3.1 plasmid vector containing the Sp35:Fc gene. The Sp35:Fc protein is encoded under the control of a CMV promoter.

3.2.3 Plasmid DNA transformation and purification

Plasmid DNA was transfected into *E. coli* MAX Efficiency® DH5α™ competent cells (Life Technologies) according to the manufacturer's instructions and successful transformants were selected by spreading the cells on LB agar (Fisher Scientific, Loughborough, UK) containing 100 µg mL^{-1} ampicillin. A single colony was picked and inoculated in 5 mL LB broth containing ampicillin and incubated for 8 h at 37°C and 250 rpm. Following amplification, the plasmid was purified using the QIAGEN® Plasmid Maxi Kit (Qiagen) according to the manufacturer's instructions and stored in TE buffer pH 8.0 (Sigma-Aldrich, Poole, UK) at –20°C. Plasmid DNA concentration (ng μL^{-1}; Equation 3.3) was determined by measuring the A_{260} and the purity was determined via the A_{260}/A_{280} ratio using a Biomate 3 Spectrometer (Thermo Scientific, Cramlington, UK). For accurate spectrophotometric DNA quantification, the A_{260} readings were taken in the range of 0.4 and 0.6 absorbance unit using Nuclease-free water (Qiagen) as diluent. To control the purity of the DNA, only preparations with a ratio of greater than 1.8 were used.

$$\text{DNA conc.} = 50 \times A_{260} \times \text{dilution factor} \qquad (3.3)$$

3.2.4 Plasmid DNA copy number

The plasmid DNA copy number per nanogram of a given vector was determined from its length using the following equation;

$$\text{Number of copies} = \frac{\text{amount (ng)} \times N_A}{\text{length (bp)} \times 650 \times 10^9} \quad (3.4)$$

where N_A is the Avogrado's number (6.022×10^{23} molecules mol^{-1}), 650 is the average weight of a base pair in g mol^{-1} (Dalton), and 10^9 is unit the conversion from gram to nanogram.

3.3 Transient Transfections

3.3.1 Electroporation

Electroporation-mediated transfection was conducted using the Amaxa® Cell Line Nucleofector® Kit V system (Lonza, Basel, Switzerland). Cells were seeded at 1.5×10^5 viable cells mL^{-1} and were sub-cultured every 2 days for 2 passages prior to transfection. Transient transfection was conducted at Day 2 of culture when the cell density was $<1\times10^6$ cells mL^{-1}. 4.5×10^6 cells per cuvette were centrifuged at 100×g for 8 min. Cell pellets were resuspended in the Nucleofector solution and transfected with up to 4.6 µg plasmid DNA using programme U-024. Immediately after electroporation, cells were diluted in 700 µL of culture medium and transferred to a TubeSpin (TPP, Trasadingen, Switzerland) containing 10 mL pre-warmed culture medium at a seeding density of 2.5×10^5 cell mL^{-1}. To scale up, samples from 3 cuvettes were pooled and cultured in a 125 mL Erlenmeyer flask. After 1–2 h of incubation at 37°C and 5% CO_2, sample was taken to determine cell viability and anti-clumping agent (Life Technologies) was added at 1:200 dilution. To determine the transfection efficiency, 4.5×10^6 cells were transfected with 1–4 µg of pmaxGFP® vector (Lonza) using the same procedure and analysed using flow cytometry.

3.3.2 Lipofection

Lipofectamine-mediated transfection was conducted using Lipofectamine® LTX with PLUS™ reagent (Life Technologies). One day before transfection, cells were sub-cultured at a seeding density of 3.7×10^5 cells mL^{-1} in CD-CHO supplemented with 8 mM glutamine in 1 L Erlenmeyer flask. The cells were grown for 24 h to a density of 1.0×10^6 cells mL^{-1} and aliquots of 25 mL were added to each 125 mL Erlenmeyer flask. For each transfection, 20 µg of plasmid DNA was diluted in 1.2 mL Opti-MEM® I Reduced Serum medium (Life Technologies) with 20 µL PLUS reagent, and then combined with 70 µL Lipofectamine pre-diluted in 1.2 mL Opti-MEM® medium. The Lipofectamine/DNA mixture was allowed to

stand at room temperature for 5 min before being added to the culture. Anti-clumping agent (Life Technologies) was added 24 h post-transfection at 1:200 dilution. To determine the transfection efficiency, cells were transfected with pmaxGFP® vector (Lonza) using the same procedure and analysed using flow cytometry.

3.4 Flow Cytometry

Cells were fixed with paraformaldehyde (PFA; Sigma-Aldrich) prior to flow cytometry analysis. 4% w/v PFA was prepared by adding 2 g PFA to 50 mL 1X PBS (Fisher Scientific). An aliquot of cells to be fixed were harvested by centrifugation at 200×g for 5 minutes and washed with 1X DPBS (Sigma-Aldrich). The cells were incubated in cold 4% w/v PFA solution at a concentration of 1×10^7 cells mL^{-1} for 15 minutes. After incubation, the cells were then centrifuged and resuspended in cold DPBS and stored at 4°C until use (<2 weeks). Flow cytometry was performed using Attune® Acoustic Focusing Cytometer (Life Technologies). A sample of 10,000 cells was measured and negative control cells without GFP plasmid were used to determine cellular auto-fluorescence. Data was analysed using Attune® Cytometric software (Life Technologies).

3.5 Recombinant Protein Quantitation

3.5.1 Determination of volumetric titre

To determine the recombinant protein titre, 0.5 mL of culture medium was filtered through a 0.22 µm Costar® Spin-X® Centrifuge Tube Filter (Sigma-Aldrich) by centrifugation at 10,000×g for 5 minutes. Sp35:Fc protein concentration was measured using FastELISA® Human IgG Quantification Kit (RD-Biotech, Besançon, France) whereas SEAP protein concentration was measured using SensoLyte® pNPP Secreted Alkaline Phosphatase Reporter Gene Assay Kit (AnaSpec, Fremont, CA) according to the manufacturer's instruction in both cases. A PowerWave™ XS microplate reader (Bio-Tek, Bedfordshire, UK) was used for absorbance measurement while KC4 3.1 software (Bio-Tek) was used to generate the Sp35:Fc and SEAP standard curves and interpolate the recombinant protein concentrations (Figure 3.2).

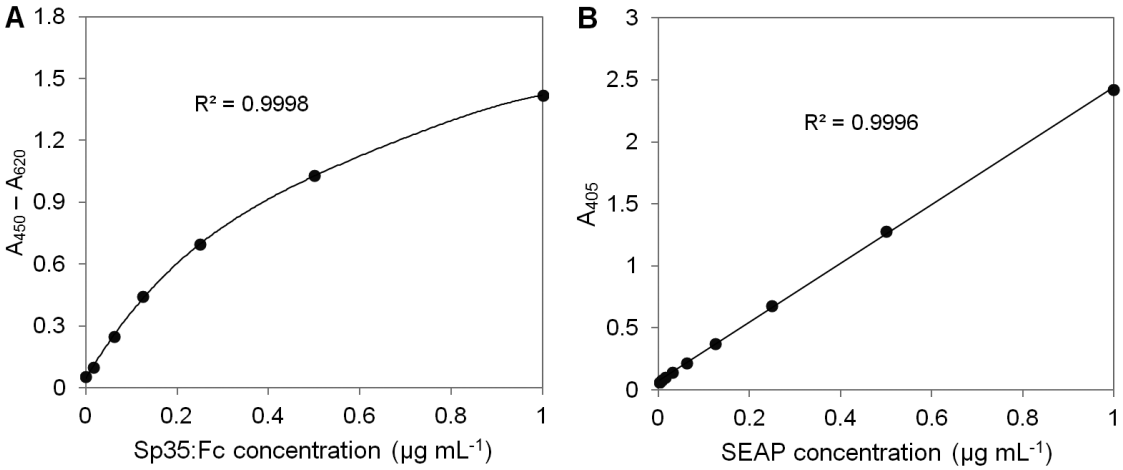

Figure 3.2: Example of standard curves generated using ELISA-based methods. Absorbance measurement was used to determine the concentration of recombinant Sp35:Fc fusion protein **(A)** and SEAP **(B)** in culture media. For sample concentrations above 1 µg mL^{-1}, serial dilution was performed on the samples.

3.5.2 Determination of specific productivity

The daily cell-specific productivity (*q*P; pg cell^{-1} day^{-1}) based on viable cell number, *X* (cells mL^{-1}) and volumetric titre, *T* (pg mL^{-1}) was calculated from the titre at second time point and first time point using the following equation:

$$qP = \frac{T_2 - T_1}{\left(\frac{X_2 + X_1}{2}\right) \times \Delta t} \tag{3.5}$$

The average *q*P through culture was determined from the volumetric titre and IVCD as follows:

$$qP = \frac{T_2 - T_1}{\mathrm{IVCD}_{2-1}} \tag{3.6}$$

3.6 Measurement of Sp35:Fc Protein Polypeptides

3.6.1 SDS-PAGE

For intracellular polypeptides, 5×10^6 cells were harvested and lysed in 500 µL of RIPA buffer (50 mM Tris–HCl, pH 8.0, with 150 mM sodium chloride, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate (Sigma-Aldrich)) and incubated at 4°C for 10 min. Lysates were centrifuged at 8,000×g for 10 min at 4°C,

supernatant fractions removed and stored short-term at 4°C. For analysis of extracellular polypeptides, 1 mL of 0.22 µm filtered culture supernatant was stored short-term at 4°C. SDS-PAGE was performed using 3–8% NuPAGE® Tris-Acetate gels (Life Technologies) with LDS sample buffer (Life Technologies) under reducing and non-reducing conditions, in which the latter contained 25 mM *N*-ethylmaleimidine (NEM) as an alkylating agent. For absolute quantification method, protein samples were resolved alongside a serial dilution of purified Sp35:Fc protein of known concentration at constant 150 V for 1 h. Biotinylated protein ladder (Cell Signaling, Danvers, MA), Novex® Sharp (Life Technologies) and HiMark™ marker (Life Technologies) were used for estimation of molecular weights where the latter was stained with SimplyBlue™ SafeStain (Life Technologies) according to the manufacturer's instructions.

3.6.2 Western blot

Resolved Sp35:Fc protein samples were transferred to nitrocellulose membranes by iBlot® semi-dry blotting (Life Technologies) according to manufacturer's instructions. Each membrane was washed with 20 mL 1X TBS-T for 5 min and blocked in 10 mL 5% w/v milk/TBS-T for 1 h at room temperature. After three washings with 15 mL 1X TBS-T (5 min each), the membrane was incubated overnight at 4°C with HRP-conjugated goat anti-human IgG Fc γ fragment in 10 mL 4% w/v BSA/TBS-T solution (1:20,000; Jackson ImmunoResearch, West Grove, PA), followed by three washings and incubation with HRP-conjugated anti-biotin antibody in 5% w/v milk/TBS-T solution (1:1000; Cell Signaling) for 1 h at room temperature. Finally, the membrane was washed three times to remove unbound antibodies prior to quantification.

3.6.3 Quantification of Sp35:Fc protein polypeptides

Detected polypeptides were visualised with an Immobilon™ Western chemiluminescent HRP substrate according to the manufacturer's instructions (ECL; Millipore, Watford, UK) and chemiluminescence signals collected using a CCD camera (ImageQuant™-RT ECL™; GE Healthcare, Amersham, UK). Western blot images were analysed using ImageQuant TL image analysis software (GE Healthcare) and Sp35:Fc polypeptides were quantified by relative quantification method or by external calibration using the band densities (known-concentrations) of the purified Sp35:Fc standards ran on the same gel (Figure 3.3).

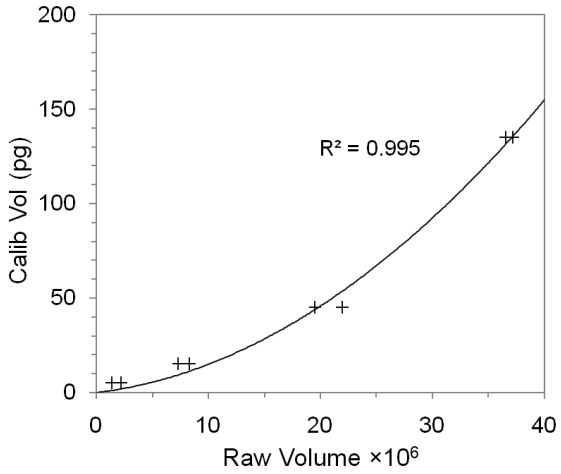

Figure 3.3: Example of standard curve generated by ImageQuant TL software for external calibration. Band densities of purified Sp35:Fc standard at different concentrations were used to quantify Sp35:Fc polypeptide samples.

3.7 Statistics

SigmaPlot 12.0 software (Systat Software, San Jose, CA) was used for statistical analysis. To test for linear correlation, Pearson's product moment correlation coefficient (PPMCC) was used. To test for statistical significance, an unpaired, two-tailed student's *t*-test was used. A *P* value of <0.05 was considered to be significant.

Chapter 4

Cellular Mechanism of a Difficult-to-Express Fusion Protein Production in CHO Cells

This chapter provides a detailed understanding of the cellular constraints normally associated with difficult-to-express protein production, as well as broader insights to the potential repercussions such as formation of protein aggregates. The aims are to identify the CHO cell constraints in manufacturing Sp35:Fc fusion protein at elevated transgene expressions, as well as to study the impact of the unfolded protein response. The data obtained, in turn, enable the identification of engineering targets to improve the production process.

4.1 Introduction

Major successes have been achieved in the production of recombinant therapeutic proteins from cultured mammalian cells, particularly with monoclonal antibodies (MAbs) and Fc-fusion proteins. Amplification methods to increase the recombinant DNA (rDNA) copy number yield efficient production of many recombinant products, with typical specific productivities (*q*P) of 10 pg/cell/day a decade ago have now increased to around 50 pg/cell/day (Butler and Meneses-Acosta, 2012). Considerable efforts have also been directed to targeting rDNA to transcriptionally active sites (Nehlsen et al., 2009), and to include chromatin opening elements (Antoniou et al., 2003). In the case of transient expression systems, powerful transfection method like Nucleofection (Zeitelhofer et al., 2007) and optimisation of transfection conditions (Mozley et al., 2014; Thompson et al., 2012) have led to very high transfection efficiency and low toxicity, thus enabling very high rDNA copy number delivered per cell. Transient expressions can also be enhanced by using post-transcriptional regulatory elements (Mariati et al., 2012).

Nevertheless, increased transgene copy number and transcriptional and translational activities do not necessarily improve the amount of secreted recombinant proteins, where very high rDNA copy number and/or mRNA levels have been shown not to correlate to the increase in protein output in both stable (Barnes et al., 2004; Schröder et al., 1999) and transient productions (Ku et al., 2008) as well as *in vivo* (Takahashi et al., 2011). Improvements through rDNA/mRNA strategy will ultimately reach a plateau beyond which no increases productivity can be achieved if no concurrent improvements in post-translational processes are made. Moreover, such strategy can easily fail for some particular difficult-to-express (DTE) proteins, for which even moderate levels of protein expression cannot be attained. In such cases, the host cells are often incapable of handling the folding or processing of the recombinant proteins, thus triggering cellular stress responses and apoptotic effects leading to reduced cell growth and/or productivity (Le Fourn et al., 2014; Pybus et al., 2014a).

The secretion of proteins by mammalian cells is a complex pathway. It involves translocation of nascent polypeptide from the cytosol into the endoplasmic reticulum (ER) where they fold with various chaperones and foldases, followed by post-translational modifications in the Golgi complex (e.g. glycosylation) before being transported to their final destination. Fundamental understanding of how mammalian cells cope with high levels of rDNA expression are critical to further enhance the recombinant protein titres (as a measure of *q*P), and particularly to the development of new, ever increasing difficult to

produce recombinant proteins. In this regard, several recent reports (e.g., Le Fourn et al., 2014; Mason et al., 2012; Pybus et al., 2014a) have provided a detailed description of cellular mechanisms of MAb production, which in turn provided strategies in designing better expression systems during the production process. On the other hand, comparatively little has been done to investigate the production of Fc-fusion proteins at cellular level, where studies on these proteins primarily focus on the aggregation mechanisms (e.g., Fast et al., 2009; Shukla et al., 2007; Strand et al., 2013).

Therefore, the mechanism for a difficult-to-express Fc-fusion protein production is presented here to better understand the potential cellular constraints in manufacturing this growing class of therapeutic proteins. In this work, we used a model Sp35:Fc fusion protein (proprietary of Biogen Idec, USA; Figure 4.1) to study the transient expression of DTE Fc-fusion protein in CHO cells. The recombinant protein was developed to provide means of treating diseases, disorders and injuries involving demyelination and dysmyelination (e.g. multiple sclerosis), by the administration of an Sp35 antagonist, whereas the potential advantages of the Fc fusion include solubility, *in vivo* stability, and multivalency (Mi et al., 2013). The leucine-rich repeat (LRR) motif functions by interacting with specific receptor complex found on neurons to constitute a functional receptor for myelin-derived inhibitors, which in turn suppress myelination and prevent the repair of damaged axons. Therefore, Sp35-derived products could have therapeutic applications for central nervous system injury (Mi et al., 2004; Pepinsky et al., 2014).

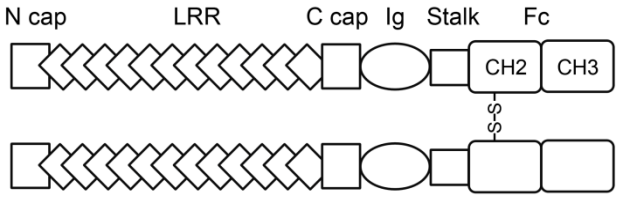

Figure 4.1: Schematic representation of homodimeric Sp35:Fc fusion protein. Each Sp35 molecule (also designated as LINGO-1 in the literature) consists 12 leucine-rich repeat (LRR) motifs flanked by N- and C-terminal capping domains, 1 immunoglobulin (Ig) domain and a stalk, and is fused to the Fc domain (CH2 and CH3) of an IgG1 via a non-covalent bond. The Fc-domain is linked via a disulphide bond in the hinge region. Adapted from Mi et al. (2004).

The Sp35:Fc fusion protein is secreted as a homodimeric molecule comprising two identical polypeptide chains covalently linked through interchain disulphides in the hinge region, with each chain consists of a therapeutic Sp35 molecule fused to the Fc domain of

human IgG1 via a non-covalent bond (Figure 4.1). The recombinant protein, however, is particularly difficult to produce with very low productivity/titre and significant product aggregates. Throughout this thesis, the "monomer" for Sp35:Fc protein is defined as one of the two identical polypeptide chains. Each chain has 10 potential *N*-linked and 3 potential *O*-linked glycosylation sites. Additionally, each Sp35 molecule contains 12 cysteine amino acids that could potentially form mismatched disulphide bonds (Scott Estes, personal communication).

Here we examine the impact of elevated Sp35:Fc DNA loads on the transient protein expression by analysing the recombinant mRNA copy number, intracellular polypeptide content, secreted proteins, as well as the potential repercussions of protein aggregation. We observe that the host cells have a very restricted cellular capability in manufacturing the recombinant protein even at low transgene copy numbers, and that there is a specific threshold level that the cells are able to process nascent polypeptides in the ER without significantly invokes the unfolded protein response (UPR). By elucidating the core cellular mechanisms responsible for differential expression at varying rDNA loads, we hope to gain broader insight of how CHO cells regulate DTE protein production, which can then be utilised to design engineering strategies for improved bioproduction process.

4.2 Materials and Methods

4.2.1 Measurement of mRNA copy numbers

1×10^6 cells were harvested by centrifugation at 200×g for 5 min and cell pellets were immediately resuspended in 300 µL of RNAprotect® Cell Reagent (Qiagen) and stored at –20°C. RNA was purified using the RNeasy® Plus Mini Kit (Qiagen) and genomic DNA (gDNA) was eliminated by using the Ambion® TURBO DNA-free™ Kit (Applied Biosystems, Warrington, UK) according to the manufacturer's instructions in both cases. gDNA-free RNA was converted to cDNA in a 20 µL reaction by using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and an Applied Biosystems Veriti™ 96-well Thermal Cycler according to the manufacturer's instructions. The cDNA was quantified by quantitative real-time polymerase chain reaction (qRT-PCR) using an Applied Biosystems 7500 Fast Real-Time PCR system and the calibration curve method.

Standard curves were generated using known quantities of the linearised rDNA plasmids as a template (Figure 4.2; Pfaffl et al., 2002). The PCR reactions were performed in a final volume of 20 µL using TaqMan® Universal Master Mix II with UNG (Applied

Biosystems), 1 µL of cDNA, 500 nM of forward and reverse primer each and 200 nM of probe. Primers and probes were designed using Applied Biosystems Primer Express® 3.0 software and TAMRA and FAM dyes were used as the quencher and reporter dyes, respectively. Primers and probes (Eurofins MWG Operon, Ebersberg, Germany) for amplification and quantification of the Sp35:Fc cDNA and SEAP cDNA were as follows: Sp35:Fc forward primer 5'-TGCCTGGTCAAAGGCTTCTATC-3', reverse primer 5'-GTTCTCCGGCTGCCCATT-3' and probe 5'FAM-CCGTGGAGTGGGAGAG-3'TAM; and SEAP forward primer 5'-CCGCTTTAACCAGTGCAACA-3', reverse primer 5'-CCCGATTCATCACGGAGATG-3' and probe 5'FAM-ACACGCGGCAACG-3'TAM. The PCR thermal cycle profile was as follows: 50°C for 2 min; 95°C for 10 min; 40 cycles of 95°C for 15 s and 60°C for 1 min.

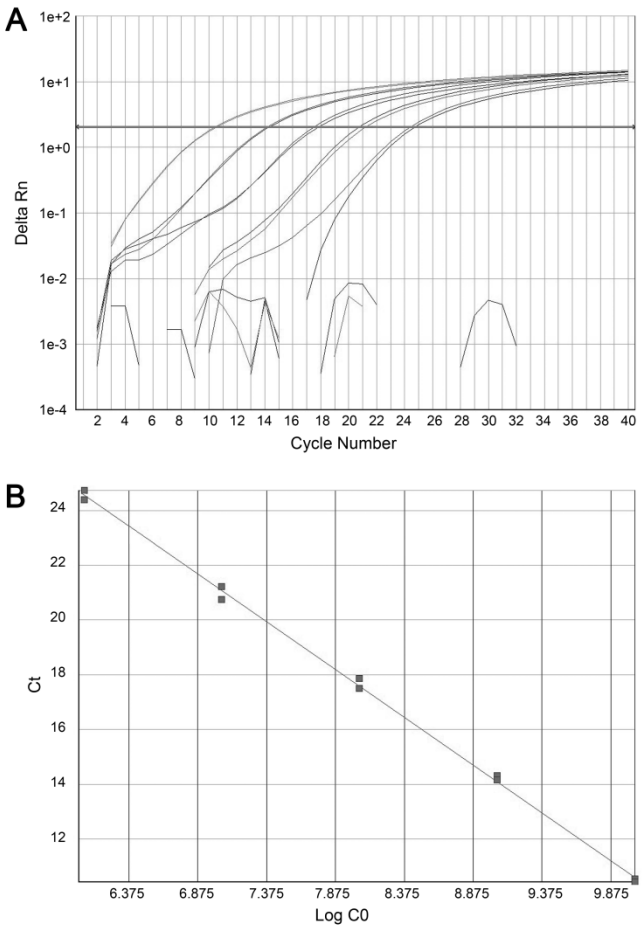

Figure 4.2: The recombinant plasmids served as an external homologous DNA standard of known copy number to generate a standard curve for quantitative real-time PCR. (A) An example of ΔRn vs. cycle number profile for 10-fold dilution series ran in duplicate. **(B)** The corresponding standard curve generated for determination of mRNA copy numbers.

4.2.2 Endo H digestion of Sp35:Fc protein polypeptides

Sp35:Fc protein was purified from cell culture supernatant and cell lysates (prepared as described above) using NAb™ Protein A Spin columns (Thermo Scientific) and concentrated by ultrafiltration using Pierce® Concentrator 20 kDa cutoff membrane (Thermo Scientific) according to the manufacturer's instructions in both cases. Protein samples were denatured by heating the samples in the supplied glycoprotein denaturing buffer at 100°C for 10 min. 0.5–1 µg of intracellular Sp35:Fc was digested with 500–1,000 NEB units of Endo H_f (New England Biolabs, Hitchin, UK) in 40 µL total reaction volume using the supplied G5 reaction buffer for 1 h at 37°C. Digestion was also performed without prior denaturation. RNase B and secreted Sp35:Fc protein were used as positive and negative control, respectively, for enzyme digestions. Samples were resolved by SDS-PAGE under reducing conditions and subjected to immunoblotting prior to imaging and analysis as described Chapter 3, Section 3.6. Resolved RNase B was stained with SimplyBlue™ SafeStain (Life Technologies) according to the manufacturer's instructions.

4.2.3 Analysis of Sp35:Fc polypeptide intracellular degradation

Cell lysates were prepared as described above using RIPA buffer supplemented with 10 mM NEM and Protease Inhibitor Cocktail Set III (Merck Chemicals, Nottingham, UK). Ubiquitinated proteins were isolated using a Ubiquitinated Protein Enrichment Kit (Merck Chemicals) according to the manufacturer's instructions. Briefly, cell lysate protein was incubated with polyubiquitin affinity beads at 4°C for 4 h, washed and ubiquitinated proteins extracted by boiling the beads in Laemmli sample buffer (Sigma-Aldrich) for 5 min. After centrifugation at 10,000×g for 1 min, ubiquitinated proteins in the supernatant and a serial dilution of Sp35:Fc standard were resolved on non-reducing SDS-PAGE and subjected to immunoblotting prior to imaging and analysis (Chapter 3, Section 3.6).

4.2.4 Measurement of UPR proteins

Cell lysates were prepared as described above using RIPA buffer supplemented with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and stored at –80°C. Electrophoresis was carried out on 4–12% NuPAGE Bis-Tris gels (Life Technologies). Resolved protein was blotted and the membrane was blocked with 5% BSA/TBS-T for 1 h at room temperature, washed three times with 1X TBS-T (5 min each)

and then probed for 1 h at room temperature with anti-BiP, PDI, eIF2α, p-eIF2α, GAPDH, unspliced XBP1 (all 1:2000; Cell Signaling) or unspliced XBP1 (1:1000; Abcam, Cambridge, UK) primary antibodies. After three washings, the membrane was incubated with HRP-conjugated goat anti-rabbit IgG secondary antibody (1:2000; Cell Signaling) for 1 h at room temperature, followed by three more washings to remove unbound antibodies prior to imaging and analysis as described in Chapter 3, Section 3.6.3.

4.3 Results

4.3.1 Sp35:Fc protein expression saturates at high gene copy numbers

An important prerequisite for the study was to ensure that the cells were not influenced by the transfection procedure *per se*. Transfections were therefore performed by means of the Nucleofector system using the same batch of cells and with the same total plasmid DNA load, where the latter was achieved by varying the ratio of rDNA : empty vector. The Nucleofector system has been shown to give very high cell viability and transfection efficiency and uniform plasmid uptake compared to other systems such as Lipofection (e.g. Davies et al., 2013; Zeitelhofer et al., 2007). Cell viabilities attained post-transfection were 90±2% with >99% transfection efficiency as analysed by flow cytometry using GFP (Figure 4.3). To study the intrinsic ability of the cells to manufacture recombinant Sp35:Fc, the cells were transfected with varying quantities of rDNA ranging from 0 to 3 µg rDNA per 4.5×10^6 cells and Sp35:Fc expression levels were quantified 48 h after transfection.

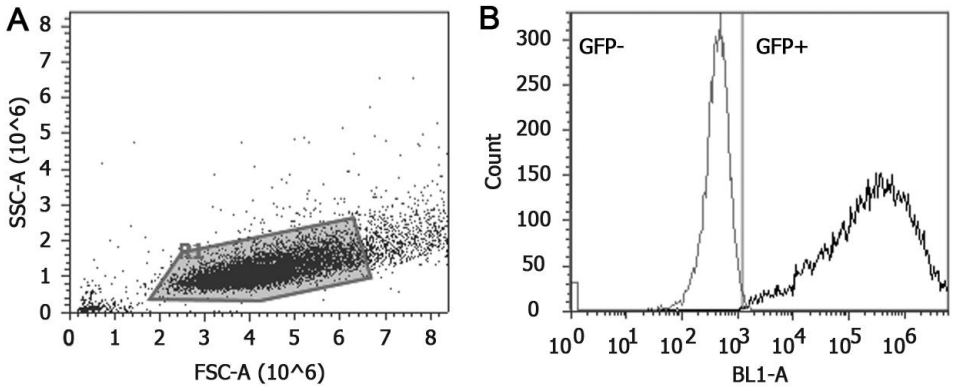

Figure 4.3: Determination of transfection efficiency using intracellular GFP marker and flow cytometry. Cells were harvested 48 h post-transfection. (A) Cell population gated for granularity and size. **(B)** The percentage of GFP-positive cells of the transfected population (black line) was quantified against mock-transfected negative control cells (grey line).

Analysis of the *q*P revealed that the expression of Sp35:Fc did not show a linear response to increasing rDNA load—at high levels of gene copy number, no further increase in the expression level of Sp35:Fc was observed (Figure 4.4A). To substantiate this saturation phenomenon, transfection was also performed using SEAP which is widely used to study promoter activity and gene expression as a model ETE glycosylated protein. Comparison of the two recombinant proteins showed a clear difference in recombinant protein expression in which the Sp35:Fc productivity gradually became saturated as the rDNA load increased, whereas SEAP productivity increased linearly to a much higher rate (Figure 4.4B). In both cases, the mRNA copy numbers increased proportionally to rDNA load (Figure 4.4C), indicating that no restriction of gene expression occurred in the system at the level of transcription. However, the influence of transcription efficiency, size of plasmid and mRNA stability can be expected to introduce some variation between the levels of Sp35:Fc mRNA and SEAP mRNA.

As shown in Figure 4.4A, there was also an inverse relationship between cell specific growth rate (μ) and *q*P, with the highest Sp35:Fc-producing cells had a 28% lower μ compared to the non-producing mock-transfected cells. This however, was less apparent in the case of SEAP-producing cells even at high *q*P (Figure 4.4B), implying that the Sp35:Fc synthesis imposed a substantial metabolic burden onto the cells. For both Sp35:Fc and SEAP-producing cells, the analysis of cell diameter and cell protein biomass by image analysis and biochemical assay, respectively, showed that there were no significant variations in average cell size and total intracellular protein content per cell ($P > 0.05$; data not shown), suggesting that there was neither change in cell morphology nor inhibition of cell biomass accumulation.

4.3.2 Reduction of folding/assembly and secretion rates and induction of disulphide-bridged aggregates

Quantitative Western blot analysis was carried out to assess intracellular and extracellular Sp35:Fc polypeptides per unit cell at different rDNA expressions. To inhibit artefactual disulphide bond scrambling, protein samples (culture supernatant and cell lysates) were treated with the alkylating agent NEM during sample buffer preparation (Taylor et al., 2006). Assignment of bands to particular Sp35:Fc protein species was achieved by comparison of the apparent molecular weight following non-reducing SDS-PAGE. Tris-acetate gels were calibrated with HiMark and Novex Sharp molecular mass standards in

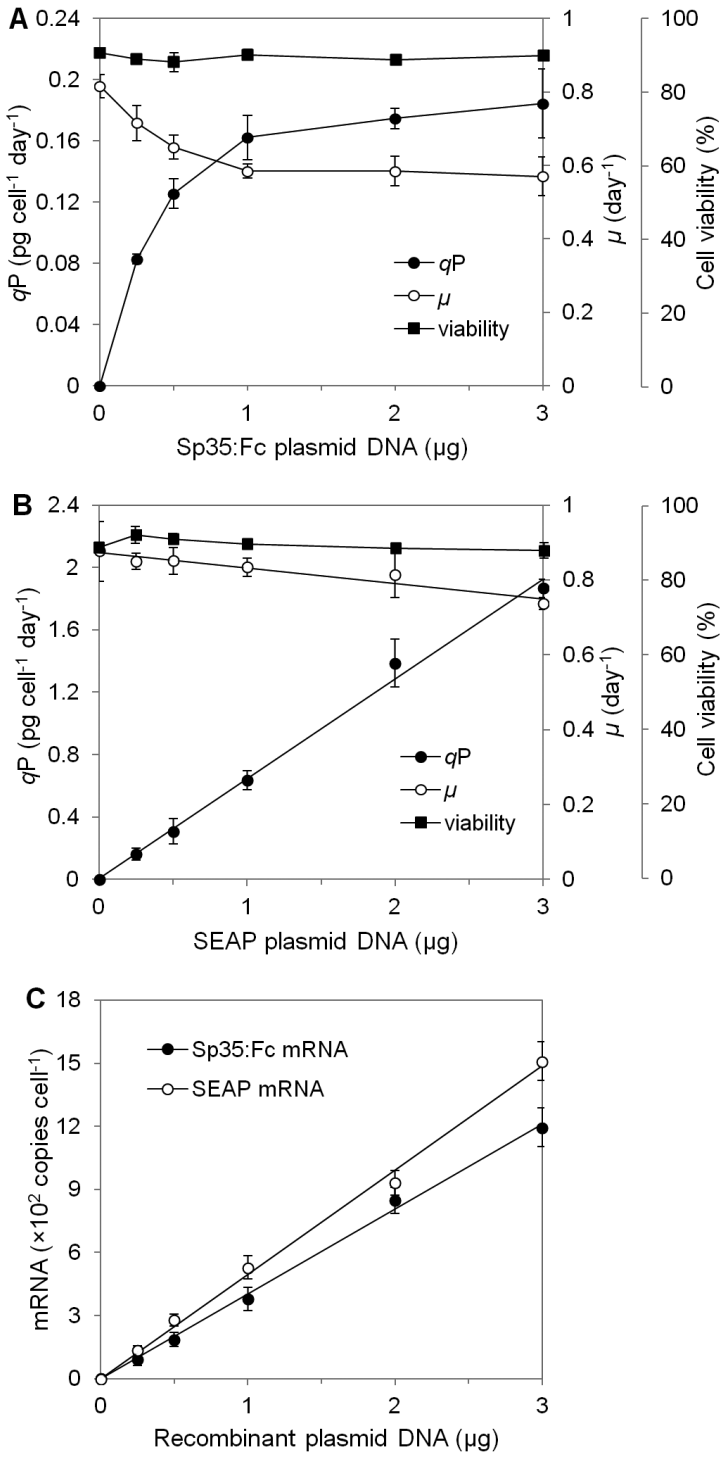

Figure 4.4: Recombinant protein production kinetics in CHO cells at different transgene copy numbers. (A) 4.5×10^6 CHO-S cells were transiently co-transfected with Sp35:Fc plasmid DNA and empty vector at varying ratios (3 μg total plasmid DNA). Cell-specific productivity (*q*P) and growth rate (μ) were measured at 48 h after transfection while post-transfection cell viability was measured at 2 h after transfection. **(B)** Similar transfection procedure and analysis were performed using SEAP plasmid DNA. **(C)** Copies per cell of Sp35:Fc and SEAP mRNAs at 48 h post-transfection were determined by absolute quantification RT-PCR. Data shown is mean value of two biological replicates and three technical replicates. Error bars represent the standard deviation.

the 40–500 kDa and 30–260 kDa ranges, respectively, and biotinylated protein ladder (30–200 kDa) for resultant blots. Combination of these calibrations allowed a highly reliable estimation of the molecular weights of very large species (Hannemann et al., 2009). Using this approach the Sp35:Fc protein was observed to have an apparent mass of approximately 100 kDa per monomer and 200 kDa per native dimer under non-reducing conditions (Figure 4.5). Additionally, the intra- and extracellular Sp35:Fc polypeptides resolved were found to contain tetramer (400 kDa) and (likely) hexamer aggregates (600 kDa). Other aggregate species were not observed, specifically trimer and pentamer, as well as higher oligomers (e.g., octamer). Under reduced condition, all bands were reduced to 100 kDa, indicating that the fully folded Sp35:Fc dimer as well as the aggregates were covalently bonded via intermolecular disulphide bridges. To rule out that Sp35:Fc self-aggregated extracellularly in the culture medium, cell culture supernatant was harvested 24 h post-transfection, incubated at 37°C for an additional 24–48 h and immediately resolved by non-reducing SDS-PAGE. No increase in protein aggregates was observed during the additional incubation period (data not shown).

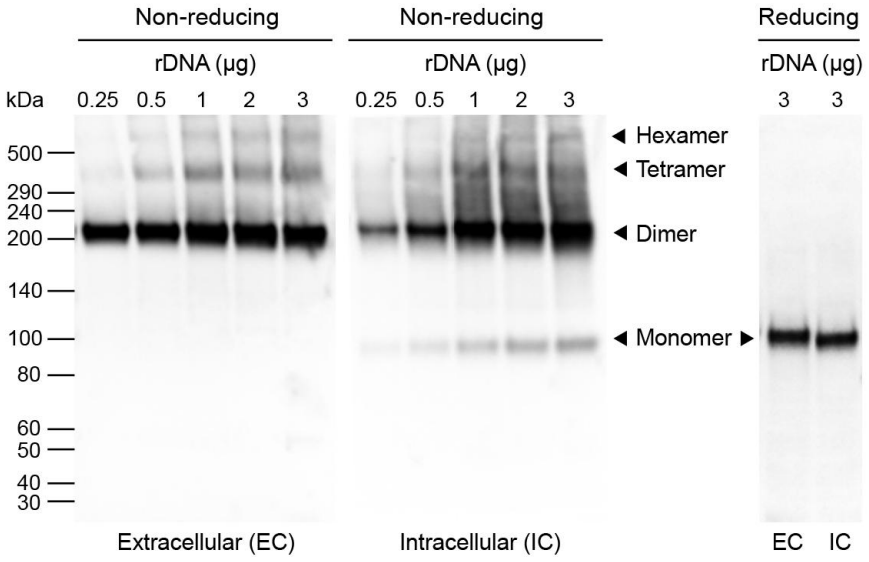

Figure 4.5: Western blot analysis of extracellular and intracellular Sp35:Fc polypeptides at different rDNA transfections. Sp35:Fc has an apparent mass of 100 kDa per monomer and 200 kDa per native dimer and may form 400 kDa tetramer and 600 kDa hexamer disulphide-bonded aggregates. Other aggregate species were not detected. Culture supernatant and cell lysate samples were prepared 48 h post-transfection and resolved by non-reducing SDS-PAGE in the presence of the alkylating agent NEM to inhibit disulphide bond reduction and scrambling. Under reduced condition, all bands were reduced to 100 kDa monomer. The figure shows a representative Western blot of duplicate culture flasks analysed in triplicate.

The amount of each Sp35:Fc polypeptide species in Figure 4.5 was determined by comparing the relative band intensity obtained with anti-IgG Fc γ specific primary antibody. Thin 1.0 mm 3–8% polyacrylamide gels were used in this study to ensure the highest gel-to-membrane transfer efficiency of the high molecular weight species. Western blot was calibrated using purified Sp35:Fc (native dimer) over a range of concentrations run on the same gel as an external standard. Purified Sp35:Fc was also added to non-Sp35:Fc producing CHO-S cell lysates as an internal standard to verify that no disulphide reduction or rearrangement occurred during sample processing. Under the conditions employed, no artefactual loss of dimer (i.e. formation of monomer from dimer) or formation of aggregates was observed (data not shown).

From the quantitative Western blot analysis, the cell-specific production rate of each extracellular protein species was calculated using Equation 3.6 (Figure 4.6A), whilst analysis of the intracellular species were carried out in terms of polypeptides copies per unit cell (Figure 4.6B). With respect to the former, determination of *q*P using this method was highly comparable to orthogonal analysis of Sp35:Fc productivity by ELISA in Figure 4.4A. Both the extracellular and intracellular quantification analyses (Figure 4.6A and B) revealed that the amount of protein aggregates per cell became significant at high rDNA loads. The composition of extracellular tetramer aggregate in Figure 4.6A for example rose from less than 0.5% w/w of the total Sp35:Fc at 0.5 µg rDNA to about 5% w/w at 3 µg rDNA load. There were also marked increases in the amount aggregates at 1 µg rDNA load which concurred with the saturation of Sp35:Fc production.

Moreover, the 2-fold increase in cellular content of Sp35:Fc dimer between 0.5 µg and 1 µg rDNA loads (Figure 4.6B) yielded in only one-third increase in dimer secretion rate (Figure 4.6A) which implies a limited protein transport capacity. From the intracellular analysis (Figure 4.6B), it is also obvious that the cells differed considerably in their relative proportion of Sp35:Fc intermediates when transfected with different amount of rDNA. The amount of assembled intracellular dimer appeared to become saturated at 1–3 µg rDNA even though the amount of nascent monomer increased linearly—indicative of variation between cell translation and folding/assembly processes. With regard to the monomer species, the amount increased from about 1.4% w/w of total intracellular Sp35:Fc content (at 0.25–1 µg rDNA) to 3.8% w/w (at 3 µg rDNA) which indicates an accumulation of unfolded monomer polypeptides in the endoplasmic reticulum (ER).

To quantify the relative proportion of intracellular Sp35:Fc in different parts of the secretory pathway, Sp35:Fc polypeptides were isolated from cell extracts using Protein A,

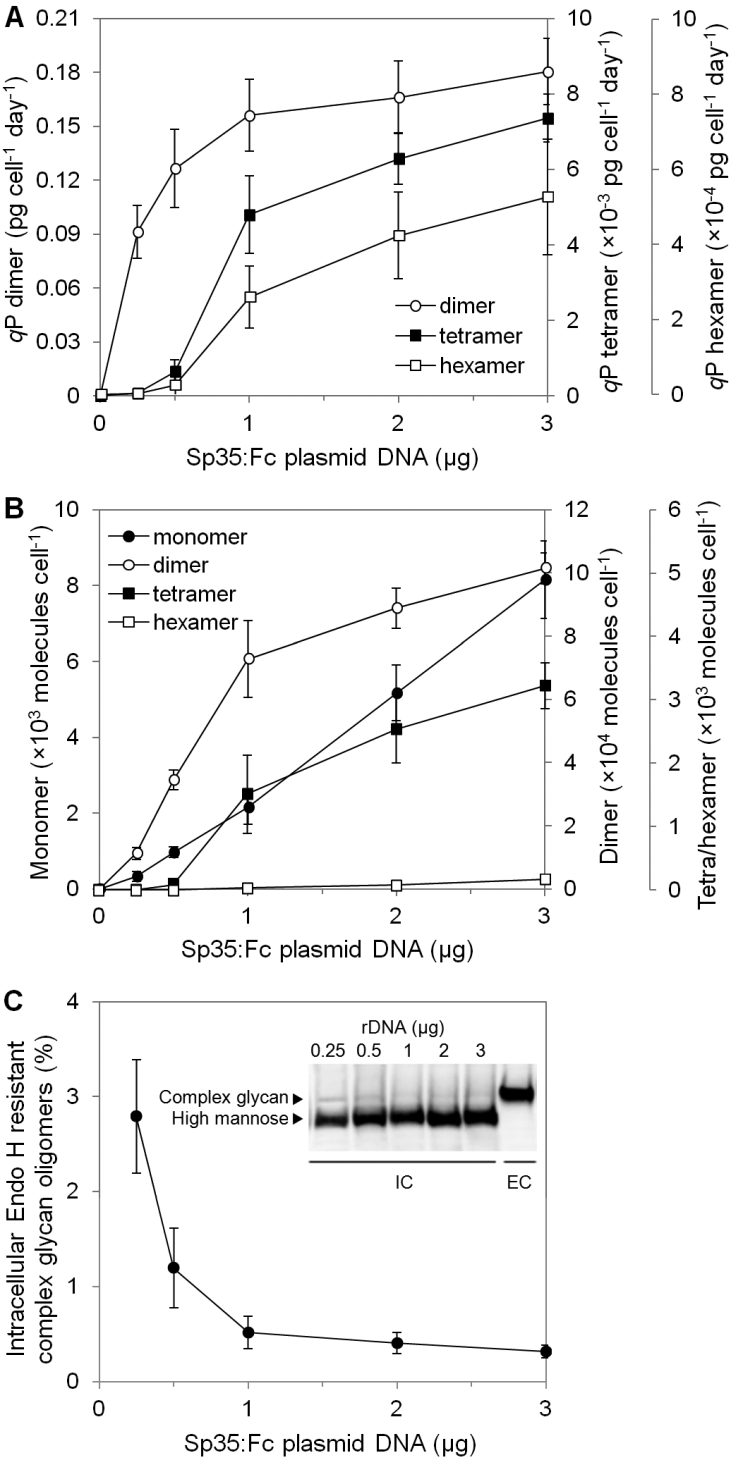

Figure 4.6: Elevation of recombinant gene expression does not result in linear increase in Sp35:Fc native dimer and promotes formation of aggregates and retention of the recombinant protein in the ER. (A) Cell-specific productivity (*q*P) was calculated from the level of respective Sp35:Fc species in the supernatant determined by quantitative Western blotting and the viable cell density. **(B)** Intracellular Sp35:Fc polypeptide copy numbers per cell determined from whole cell lysates by quantitative Western blotting. **(C)** The proportion of fully folded intracelullar (IC) Sp35:Fc species in the latter section of the secretory system decreases as rDNA load increases. Secreted extracellular (EC) Sp35:Fc was used as a negative control. Data shown is representative of duplicate culture flasks analysed in triplicate.

quantitated by ELISA and subjected to treatment with Endoglycosidase H (Endo H) enzyme. Endo H selectively cleaves *N*-glycan side-chains of the high-mannose form that present in the ER and intermediate compartment (between the ER and *cis*-Golgi), whereas glycoproteins that are transported beyond the medial Golgi generally possess *N*-glycans that are resistant to Endo H digestion. Analysis of the treated proteins by reducing SDS-PAGE showed that the vast majority (>97%) of the fully disulphide bonded Sp35:Fc within the cells was associated with high-mannose *N*-glycans and was located in the early secretory compartments (ER/intermediate compartment; Figure 4.6C). These data were analysed in reduced condition as Endo H sensitive and resistant high molecular weight proteins (dimer and other oligomers) could not be separated adequately using non-reducing SDS-PAGE. Analysis of non-reducing SDS-PAGE however showed that the monomer polypeptide was entirely Endo H sensitive (data not shown), implying that this species did not leave the early secretory pathway.

Figure 4.6C also shows that the proportion of high mannose to complex glycans was greatly increased at 1 µg rDNA onwards to nearly 100%, signifying a molecular crowding in the ER/intermediate compartment secretory pathway and reduction in the rate of ER-to-Golgi transport of Sp35:Fc. This molecular crowding coincided with the saturation of intracellular assembled dimer and marked increase in protein aggregation (Figure 4.6B), suggesting that the concentration of Sp35:Fc dimer within the ER might influence the extent to which Sp35:Fc can be found in disulphide-bonded aggregates. Combining all observations made above, we infer that Sp35:Fc protein production bottlenecks were due to the combination of (i) limited folding/assembly capacity, which in turn catalysed aggregate formation, and (ii) inefficient ER export rate.

4.3.3 Accumulated unfolded proteins in cells induce UPR signalling pathways and ERAD

Since there were saturation of fully-folded Sp35:Fc and retention of the protein in the ER, as well as significant amounts of intracellular aggregates, we compared the extent of UPR induction in the cells with different amounts of rDNA. Perturbations of folding process (e.g. accumulation of nascent polypeptides) activate the UPR pathways in an attempt to reinstate the homeostasis between folding load imposed on the ER and its folding capacity. In response to ER stress, protein ER kinase (PERK) phosphorylates eIF2α which inhibits general protein synthesis, whilst IRE1 splices XBP1 mRNA which is then translated

into an active, strong transcription factor to increase the expression of molecular chaperones and induce the ERAD pathway. ATF6 transcription factor is also triggered which produces downstream targets whose functions overlap with those of IRE1 (Chapter 1, Figure 1.7). When a set of UPR related proteins were examined by Western blotting, all the proteins tested were upregulated in Sp35:Fc-producing cells compared to the non-producing as well as SEAP producing cells (Figure 4.7A).

Firstly, the PERK pathway activation was analysed by quantifying the relative abundance of phosphorylated eIF2α (p-eIF2α) over total eIF2α. Cells transfected with Sp35:Fc DNA showed increased phosphorylation of eIF2α with the first band observed at 0.5 µg rDNA load (Figure 4.7A). This suggests that at 0.5 µg rDNA, there was a low-level UPR response to cope with the increased protein folding demand. However, the cell growth data in Figure 4.4A also suggested that evoking a UPR might have undesirable impacts upon the cells—we observed a significant reduction in cell biomass production rate (i.e., lower μ) in the 0.5 µg rDNA transfectants when compared with those transfected with 0.25 µg, potentially arising from reduced global translation. Additionally, the progressive PERK pathway activation (Figure 4.7B) suggests a gradual translational attenuation of Sp35:Fc mRNA (see Chapter 5 for quantitative analysis of the translational kinetics). The eIF2α phosphorylation was a consequence of PERK activation due to intracellular stress and did not occur via the GCN2 kinase due to lack of specific media component (Zhang et al., 2002) as no/insignificant bands were observed at low Sp35:Fc DNA load and SEAP-producing cells (Figure 4.7A).

Secondly, the IRE1 pathway activation was analysed by measurement of the proportion of active spliced XBP1 (XBP1s) and inactive unspliced XBP1 (XBP1u) abundance. Similar to p-eIF2α, cells transfected with Sp35:Fc DNA showed increased splicing of XBP1 with the first band observed at 0.5 µg rDNA load (Figure 4.7A), and the marked increase at 1 µg rDNA (Figure 4.7B) corresponds to the saturation Sp35:Fc qP (Figure 4.4A). Importantly, we note that the XBP1s induction was strongly correlated to the upregulations of BiP (see below; Pearson's product moment correlation coefficient, PPMCC $r = 0.985$, $P < 0.0001$) and PDI (see below; PPMCC $r = 0.957$, $P < 0.001$), as well as to the Sp35:Fc polypeptide ubiquitination targeted for ERAD (see below; PPMCC $r = 0.997$, $P < 0.0001$). Together, these data imply that the overexpression of Sp35:Fc transactivated downstream UPR signalling in an attempt to restore the ER homeostasis by inducing chaperone synthesis and ERAD.

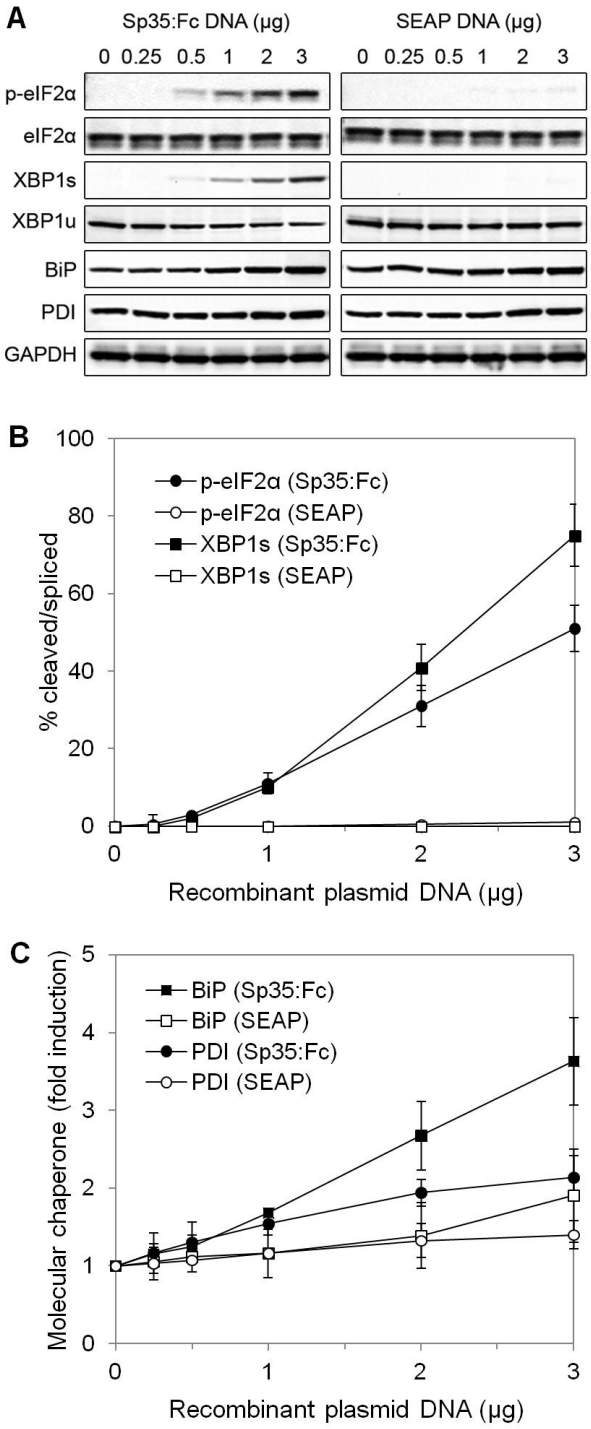

Figure 4.7: Measurements of UPR induction in Sp35:Fc-producing and SEAP-producing CHO cells. Whole cell extracts were prepared 48 h post-transfection and subjected to Western blotting and probed with antibodies detecting UPR-related proteins. **(A)** Representative immunoblots of PDI (57 kDa), BiP (78 kDa), phosphorylated eIF2α (p-eIF2α; 38 kDa), total eIF2α (38 kDa), spliced XBP1 (XBP1s; 55 kDa), unspliced XBP1 (XBP1u; 27 kDa) and GAPDH (37 kDa) in Sp35:Fc and SEAP-producing cells at varying rDNA loads. GAPDH was used as a housekeeping protein. **(B)** Quantitation of p-eIF2α and XBP1s levels. Expression of total eIF2α was detected to serve as p-eIF2α loading controls. Percent XBP1s was calculated based on its respective XBP1u. **(C)** Relative quantitation of BiP and PDI levels.

Furthermore, we assessed the relative abundance of BiP and PDI, known to be induced by the UPR (Chakrabarti et al., 2011). BiP is the major ER chaperone and the regulator of the activation of three arms of ER stress transducers (PERK, IRE1 and ATF6) whereas PDI is an ER-based enzyme that catalyses the formation of disulphide bonds in secretory proteins. BiP was found to be highly expressed in the cells transfected with Sp35:Fc DNA compared to SEAP DNA, and was 3.6-fold more abundant in cells transfected with 3 µg Sp35:Fc DNA than in mock-transfected control cells (Figure 4.7A and C). The rapid increase from 1 µg rDNA (Figure 4.7C) coincided with the saturation in fully-assembled Sp35:Fc and increase in aggregates (Figure 4.6B). On the other hand, immunoblot analysis of PDI showed a slight expression saturation at high Sp35:Fc DNA load, with only up to a 2.1-fold increase at 3 µg rDNA (Figure 4.7A and C). This data suggests that PDI could potentially be a limiting factor that was responsible for the reduced Sp35:Fc folding/assembly activity.

The degradation of a protein via the ubiquitin-proteasome pathway involves the tagging of the protein substrate by covalent attachment of ubiquitin molecules, followed by degradation of the ubiquitinated protein to small peptides by the 26S proteasome complex (Lecker et al., 2006). Therefore, we tested if the Sp35:Fc overexpression resulted in the appearance of tagged Sp35:Fc protein targeted for degradation. Sp35:Fc protein associated with ubiquitin was examined by specific affinity adsorption of ubiquitinated proteins and blotting the immunoprecipitates with anti-IgG Fc γ fragment antibody (Figure 4.8 inset). This analysis demonstrated that both unfolded monomer and fully assembled Sp35:Fc were targeted for degradation, although the former was only apparent at 1–3 µg rDNA loads (Figure 4.8).

As no Sp35:Fc monomer polypeptides were visualised at 0.25 and 0.5 µg rDNA loads even when higher concentration sample was used (data not shown), we assumed that the rate of Sp35:Fc degradation at these rDNA loads were negligible. In contrast, although no ubiquitinated Sp35:Fc aggregates were observed, it was likely that the aggregates were still targeted for degradation but below the detection level as the proportions of the high molecular weight aggregates were very small (<5%). Additionally, as Sp35:Fc monomer did not leave the early secretory compartment as tested by Endo H, the proteins targeted for degradation via the ubiquitin pathway is thought to be transported from the ER back into cytoplasm via ERAD retro-translocation mechanism and degraded by the proteasome (Ron and Walter, 2007).

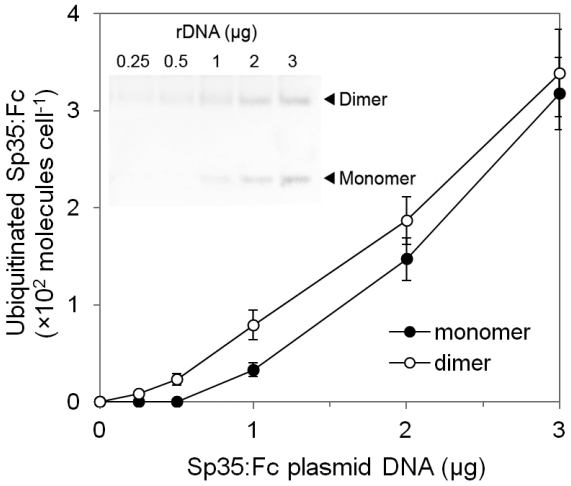

Figure 4.8: Overexpression of Sp35:Fc in CHO cells led to ER-associated degradation. Whole cell extracts were prepared 48 h post-transfection and Sp35:Fc polypeptides targeted for degradation were determined by ubiquitin pull-down and immunoblotting using anti-human IgG Fc fragment antibody. The insert shows a representative Western blot showing the observed trend in Sp35:Fc polypeptide ubiquitination. Data shown is representative of two biological and two technical replicates. Error bars represent the standard deviation.

4.4 Discussion

Mammalian cells, particularly CHO cells, have proven to be invaluable for stable, high-productions of recombinant therapeutic proteins which require various co- and post-translational modification reactions for folding, function, stability and/or subcellular targeting (Walsh, 2010). However, generation of recombinantly produced Fc-fusion derivatives can be problematic in which the production is often limited at a post-translational level. The main cause for that is thought to lie in the different and partially incompatible folding and secretion requirements of protein of interest and Fc domain entities in mammalian cells (Haas et al., 2012; Lee et al., 2007). To engineer mammalian cells to better manufacture Fc-fusion proteins or other DTE proteins with high specific productivity, it is essential to identify the rate-limiting step(s) in the protein production at elevated gene expression levels.

Advancements in gene expression technology have successfully increased recombinant DNA copy number and transcriptional activity in mammalian expression systems (Kalwy et al., 2006). However, in practical terms cell-specific production rate can only be proportional to recombinant mRNA abundance to a limited extent as observed with many (easy-to-express) recombinant proteins including SEAP (Carpentier et al., 2007),

antithrombin III (Schröder et al., 1999) and MAb (Barnes et al., 2004). Similarly, our preliminary data showed the linear increase in SEAP *q*P observed in this study quickly broke down when the cells were transfected with more than 3 µg rDNA (Figure 4.9A), although we note that the significant decrease in cell viability could also contributed to the plateau. Importantly, we found that this constraint was exacerbated in the case of DTE Sp35:Fc protein where increasing expression of using multiple gene copies did not correspondingly result in improved specific productivity even at low rDNA loads. Although a linear relationship between the amount of Sp35:Fc mRNA (Figure 4.4C) and the amount of intracellular nascent Sp35:Fc monomer (Figure 4.6B) exists and extends over the whole range of gene expression level, further analysis via mathematical approach in Chapter 5 revealed that this linear relationship is essentially preserved by the translational attenuation mechanism.

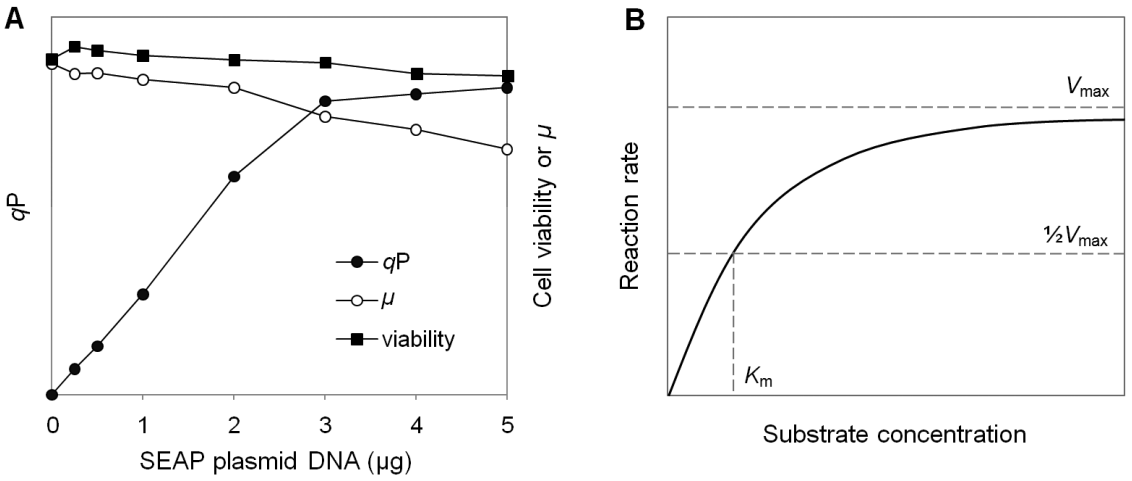

Figure 4.9: Recombinant protein productions including ETE proteins display "Michaelis-Menten kinetics" with saturation of *q*P at a certain point. (A) Overexpression of SEAP above 3 µg rDNA in this study would rapidly saturate the SEAP productivity, although this could (partly) be due to the decreases in cell viability post-transfection and *µ* (due to higher total DNA loads). **(B)** The TGE diagnostic assay utilised in this study is analogous to the Michaelis-Menten model of enzyme kinetics, where a cell's (cf. enzyme's) specific productivity (cf. reaction rate) is measured at varying substrate (recombinant gene) concentrations.

The TGE diagnostic assay presented in this work can be considered analogous to the Michaelis-Menten model of enzyme kinetics (Figure 4.9), where *q*P (cf. reaction rate) is measured at varying rDNA (substrate) amounts. We anticipate that different recombinant proteins synthesised using the same cell line (cf. enzyme) would have different characteristic kinetics profiles. The gradient of initial linear phase at low rDNA load, where

*q*P is proportional to recombinant gene (or mRNA) amount, allows comparison of the intrinsic ease (cf. an enzyme's K_m for a substrate) of expression of a given product (e.g. translation and secretion). If a saturation phase is evident as in the case of Sp35:Fc, then cellular synthetic processes inflicts a finite constraint, therefore limiting the cellular capacity of the cell (cf. enzyme V_{max}). Importantly, the other primary output parameter (cell specific growth rate) identifies feedback mechanisms that may fundamentally limit the level of transient and stable expression of the recombinant protein. With respect to the stable expression, only stable clones with (undesirably) low transcriptional output would proliferate rapid enough to persist competitively in stable transfectant pools.

Sp35:Fc is a relatively large and structurally complex glycoprotein. Elevation of expression level overloads the folding/assembly capacity of the host cells and triggers aggregation and accumulation of the recombinant protein in the ER. With respect to the former, studies have shown that the rate of multimeric MAb folding and assembly is relatively slow (Goto and Hamaguchi, 1981; Lilie et al., 1994) as well as protein and cell line specific (O'Callaghan et al., 2010; Pybus et al., 2014a). It is therefore not surprising that the folding/assembly process of an Fc-fusion protein easily becomes a production bottleneck for which the protein comprise two distinct protein groups (i.e., a therapeutic protein and the Fc domain) and require coordinated assembly of two polypeptide chains. Intracellular aggregation, especially disulphide-bonded, is also a common feature of Fc-fusion proteins (Fast et al., 2009; Lee et al., 2007; Strand et al., 2013) where the aberrant covalent bonds can be caused by thiol-mediated disulphide shuffling or disulphide bond formation between free thiol groups. Since separate work is needed to further elucidate the Sp35:Fc aggregation pathway, we can only speculate that the Sp35:Fc tetramer and hexamer formations were due to aggregation of misfolded proteins that escape the cellular quality-control mechanisms, which is common in eukaryotic systems (Dobson, 2003). The ubiquitination of fully-folded protein was a good indicator that misfolded fully-assembled Sp35:Fc dimer was present in the ER (Chakrabarti et al., 2011). Moreover, we did not observe any trimer and pentamer aggregates, suggesting that the aggregation process was an interaction between (mis)folded, assembled Sp35:Fc proteins.

Folding and assembly reactions in the ER require a considerable cellular energy and involve the concerted action of a variety of molecular chaperones and foldases (Dobson, 2003). With regard to the latter, specific reaction steps may restrict the folding rate of particular proteins more than others. For example, oxidation reactions catalysed by PDI, a folding enzyme that catalyses the formation and breakage of disulphide bonds between thiol groups, is one the main cell engineering targets. As the PDI abundance in our system

did not correspond to the increase of nascent monomer polypeptide at high rDNA loads, overexpression of PDI may improve the *q*P (via improved folding) and possibly curb the rate of aggregation. To achieve efficient protein folding in the ER, PDI must be rapidly re-oxidized, a mechanism that is largely accomplished by ER oxidoreductin-1-like α (ERO1Lα) and its ERO1Lβ homologue. The latter is of physiological interest because it is induced by the UPR to regulate oxidative protein folding under ER stress where it can form, besides ERO1Lβ-PDI dimers, homodimers and mixed heterodimers with ERO1Lα (Dias-Gunasekara et al., 2005; Pagani et al., 2000). Therefore, the overexpression of ERO1 can also potentially create an altered oxidising environment suitable for Sp35:Fc folding (Mezghrani et al., 2001).

However, there was a negative correlation between the PDI abundance and rate of ER-to-Golgi transport of Sp35:Fc where the rate decreases at higher level of PDI. This data is very similar to data reported on the influence of PDI on retention of TNFR:Fc fusion protein in the ER (Davis et al., 2000), although the retention could also be due to the abnormal disulphide-linked oligomerisation (Lobito et al., 2006). Indeed, the outcomes of several attempts to improve the rate of recombinant protein production by overexpressing PDI in mammalian cells have been mixed and inconclusive (Borth et al., 2005; Davis et al., 2000; Hayes et al., 2010; Kitchin and Flickinger, 1995). Other genetic upregulations of discrete chaperones in mammalian cell hosts to improve *q*P have also not been particularly successful. For example, BiP overexpression in CHO cells has been reported to generically decrease the secretion of recombinant proteins it associates with (Dorner and Kaufman, 1994; Dorner et al., 1988; Dorner et al., 1992). Therefore, cell engineering strategy for improved Sp35:Fc production likely requires engineering of multiple targets, either by co-expressing two or more molecular chaperones, or a specific transcription factor such as XBP1 that can simultaneously up-regulate a range of secretory pathway genes (Lee et al., 2003).

It is noteworthy that protein folding and glycosylation are interconnected, where the latter acts as a quality control mechanism and aids protein folding. For example, the glycocomponent of interferon-γ has been implicated in facilitating folding and dimerisation of the protein (Sareneva et al., 1994). Therefore we also hypothesise that the limitation in Sp35:Fc folding/assembly was due to improper glycosylation. This is particularly pertinent given the level of glycosylation of Sp35:Fc, where inability of the cells to properly glycosylate any of the sites could leave the polypeptides susceptible to misfolding and aggregation. In this regard, the distribution of glycan structures (microheterogeneity) can be regulated by modulating culture conditions including nutrient content, temperature, pH,

oxygen and ammonia (Butler, 2006). Moreover, it was feasible that the inefficiency of the Sp35:Fc protein folding occurred further upstream, i.e. during translocation. Proper processing of complex, DTE proteins may necessitate an extended translational arrest if the kinetics of ER docking is slow. Particular combinations of IgG variable domain and signal sequences, or possibly complex compositions of fusion protein, could create unfavourable structures of nascent polypeptides that in turn lead to improper functioning of the SRP complex or lack of signal sequence removal (Le Fourn et al., 2014). To improve ER docking and the translocation of DTE MAb polypeptides, Le Fourn et al. (2014) modulated the translation arrest kinetic via overexpression of exogenous SRP14 component, which eventually restored the efficiency of signal sequence processing.

A further consideration is that the glycans may play a role in upholding the folded structure by minimising the conformational freedom of the polypeptide backbone (Petrescu et al., 2004). However, like *N*-glycosylation, disulphide bond formation is co-translational in the ER and may interfere with the accessibility of some amino acids for glycan transfer. This may result in variable occupancy of a specific site, leading to macroheterogeneity of glycoforms (Allen et al., 1995). Reduction in cell growth rate (by lowering culture temperature, addition of butyrate, etc) has been demonstrated to alter the levels of site occupancy in tissue plasminogen activator and interferon-γ (Andersen et al., 2000; Nyberg et al., 1999). This is also consistent with the notion that the glycosylation efficiency (and therefore protein folding) improves at a reduced rate of protein translation. *N*-glycosylation takes place co-translationally in the ER in which the precursor glycan is exposed to the active site of the oligosaccharyltransferase enzyme over a short period of time. This means lowering of the rate of polypeptide elongation would increase the exposure time (Butler, 2006). With respect to this, Shelikoff et al. (1994) showed that when the elongation rate of prolactin polypeptides was inhibited by cycloheximide, the site occupancy was increased, thereby enhancing the glycosylation of the translocated proteins.

As expected for transfectants with relatively high Sp35:Fc DNA, UPR was constitutively activated during the cell cultivation in response to ER stress. In this scenario, the UPR has three primary functions: (i) to restore normal function of the cell by halting protein translation, (ii) to increase the synthesis of molecular chaperones involved in protein folding and secretion and (iii) to degrade excess unfolded proteins (and misfolded proteins) via ERAD mechanism thereby diminish the folding demand. In certain cases, however, the activation of the UPR may be insufficient to overcome ER stress and the prolonged accumulation of unfolded and misfolded proteins in the ER may have toxic effects, eventually leading to apoptosis (Chakrabarti et al., 2011). The comparisons of

UPR signalling branch activation reveal intrinsic features of UPR stress sensor induction in response to varying degree of ER stress. The complexity of cell cellular mechanisms, however, may necessitate empirical modelling approach to systematically elucidate differences in the cell functions under different conditions (Trusina et al., 2008).

Chapter 5

Systematic Understanding of a Difficult-to-Express Fusion Protein Production in CHO Cells via *In Silico* Analysis

This chapter presents a mathematical model for the synthesis and secretion processes of Sp35:Fc fusion protein in CHO cells. The aims are to systematically and quantitatively describe the cellular kinetics, and evaluate the impact of each cellular process to improve the production of this difficult-to-express protein. Overall, the in silico analysis reveals cell functions in maintaining cellular homeostasis, and importantly predicts the relative importance of different engineering targets.

5.1 Introduction

Systems and synthetic biology have significantly improved our understanding of the recombinant protein synthetic process and how to improve/control it in both stable and transient gene expression systems (O'Callaghan and James, 2008). Whilst progression in molecular biology and analytical techniques have provided imperative information on intricate cellular systems, mathematical modelling has been recognised as a rational approach to systematise the empirical data with the objective of identifying the key bioreactions, reactants and process parameters (Naderi et al., 2011). Although the application of mathematical models are still by and large limited in the development and production of biologics (mainly due to the complexity and unpredictability of host cell systems; Ho et al., 2012), the advancement of computational research have successfully led to the identification of growth limiting factors (deZengotita et al., 2000) and optimisation of medium design and feeding strategies (Dhir et al., 2000; Xie and Wang, 1994). More recent works have also demonstrated that mathematical tools are highly valuable for rapid identification and/or testing of rational engineering targets (e.g. Ho et al., 2006; O'Callaghan et al., 2010; Pybus et al., 2014a).

Mechanistic models can be built using a priori knowledge such as chemical, physical or biological laws and/or a posteriori knowledge based on experimental or empirical evidence. Such knowledge dictates the model structure that normally contains adjustable parameters which may or may not have physical meaning. Structured modelling for instance incorporates biological knowledge by dividing the cells into cellular processes and compartments that are physically or chemically distinct. Typically, one desires to include all explicit model components and precisely determine the parameter values as well as to validate the model statistically. Nevertheless, collecting the data that are needed to construct and validate a biological model, especially in the case of mammalian cell systems, are often resource intensive and time consuming (Franceschini and Macchietto, 2008; Kontoravdi et al., 2007). It is therefore a challenge to have a "reduced" model while maintaining an accurate description of the components/reactions that is required to provide reliable results (e.g., an unfolded protein response (UPR) model; Trusina et al., 2008).

In order to systematically understand the production mechanism of the Sp35:Fc fusion protein used in this study, we employ an empirical modelling approach to quantitatively describe the cellular constraints on recombinant Sp35:Fc synthesis and secretion. The main benefits of this mathematical approach lie in its ability to elucidate the behaviour of a multicomponent system and calculate the relative importance of specific

cellular reactions on overall flux (i.e. *q*P). This information can be subsequently utilised to identify/predict corresponding cellular processes for cell or process engineering. For instance, while the experimental data presented in Chapter 4 (Figures 4.4–4.8) indicate a connection between cell growth, Sp35:Fc synthesis and secretion and UPR induction by elevated expression, we have no systematic information of their relative influence—it is reasonable to assume that UPR induction by DTE Sp35:Fc can occur via several routes, which may include combinations of Sp35:Fc folding/assembly kinetics, aggregate formations and transport/secretion rates. Consequently, it is difficult to identify specific cellular engineering targets that could result in significant improvement of DTE Sp35:Fc production.

The structured model of Sp35:Fc secretory pathway presented in this study stems from the antibody synthetic model originally proposed by Bibila and Flickinger (1991, 1992), which also formed the basis of previous MAb models built in this laboratory (Davies et al., 2011; McLeod et al., 2011; O'Callaghan et al., 2010; Pybus et al., 2014a). Based on the previous experimental observations in Chapter 4, a simplified cellular structure for homodimeric Fc-fusion protein production is introduced with respect to the process of synthesis, folding and secretion in order to develop a modelling framework that could be used in identifying major cellular rate limiting steps (Figure 5.1). However, we are reluctant in our approach to extending our model to include UPR-related proteins such as eIF2α and XBP1 and their effects (e.g. upregulation of ER chaperones; see below).

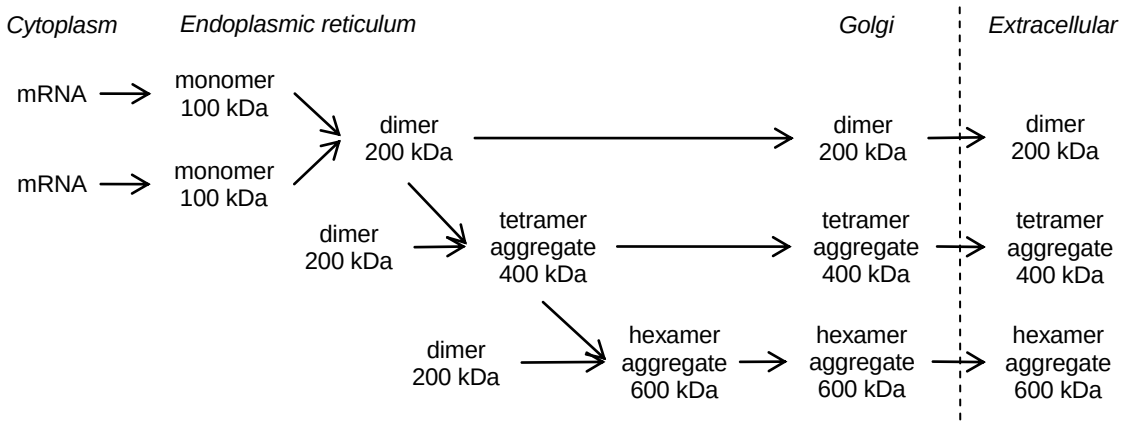

Figure 5.1: Proposed Sp35:Fc fusion protein synthesis, folding and aggregation pathway in CHO cells that forms the structured model of Sp35:Fc production. The translation of Sp35:Fc mRNA in the cytoplasm would produce a nascent Sp35:Fc monomer polypeptide which is then transported to the endoplasmic reticulum. Two monomer polypeptides fold and assemble into a functional dimer before being transported to the Golgi complex followed by secretion into the extracellular medium. A dimer protein may also form a tetramer aggregate by cross-linking with another dimer, or a hexamer aggregate by linking to a tetramer.

Indeed, the UPR comprises a complex series of network that dynamically vary. For example, recent studies have identified a translation suppression network that is independent of the phosphorylated eIF2α signalling (Guan et al., 2014). Even though it is possible to introduce a reduced UPR system in our model (e.g., a minimal UPR regulatory loop to capture the link between recombinant polypeptide level, cell growth rate and ER folding capacity; Pybus et al., 2014a), such empirical model is also incapable of predicting the effects of engineering specific UPR components that can be particularly useful for DTE protein productions. The guiding principle in our model development is to minimise the number of parameters needed for estimation from the available empirical/experimental measurements (Jaqaman and Danuser, 2006). This reduces the need for model validation and removes excessive degrees of freedom in our mathematical description of the key intracellular processes of Sp35:Fc synthesis and regulation. Nevertheless, our model does capture the major UPR events observed, namely the reduction in cell growth and ERAD pathway (via protein ubiquitination).

Figure 5.2 summarises the major steps of model development framework of this study. On the whole, the Sp35:Fc model developed consists of 24 differential and algebraic equations containing 26 parameters. The parameters were fitted to the empirical data obtained in Chapter 4 for five different CHO transfectants (containing different rDNA loads), followed by local sensitivity analysis to evaluate the impact of specific cellular mechanisms on Sp35:Fc productivity at different transgene copy numbers. This computational approach enables systematic understanding of cellular regulation in producing Sp35:Fc, and at the same time provides a rational and predictive cell engineering platform to increase Sp35:Fc protein productivity based on *in silico* engineering of discrete cellular functions.

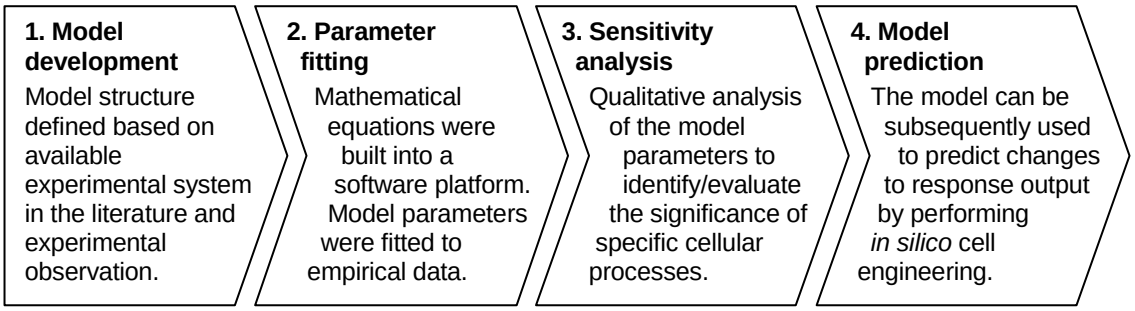

Figure 5.2: Model development framework for Sp35:Fc fusion protein production system.

5.2 Mathematical Model

5.2.1 Synthetic model development

A mathematical model consisting of a set of ordinary differential equations was developed based on a previously published model for MAb synthesis in this laboratory (McLeod et al., 2011; O'Callaghan et al., 2010). In order to reflect the Sp35:Fc synthesis mechanism, several structural modifications were made on the previous MAb model by (i) removing the unrelated light-chain polypeptide equations and expressions, and (ii) adding protein aggregate formations and secretion pathways based on the experimental observations. The model topology is shown in Figure 5.3. We acknowledged that protein synthesis is a complex mechanism that could not be fully captured in our model. For example, to characterise the actual translational frequency in a cell, one also need to determine the protein turnover rate (half-life) in addition to the pure production and ubiquitination levels. Therefore, this simplified model was merely an estimation of the main cellular activities to allow systematic comparisons of the different transfectants.

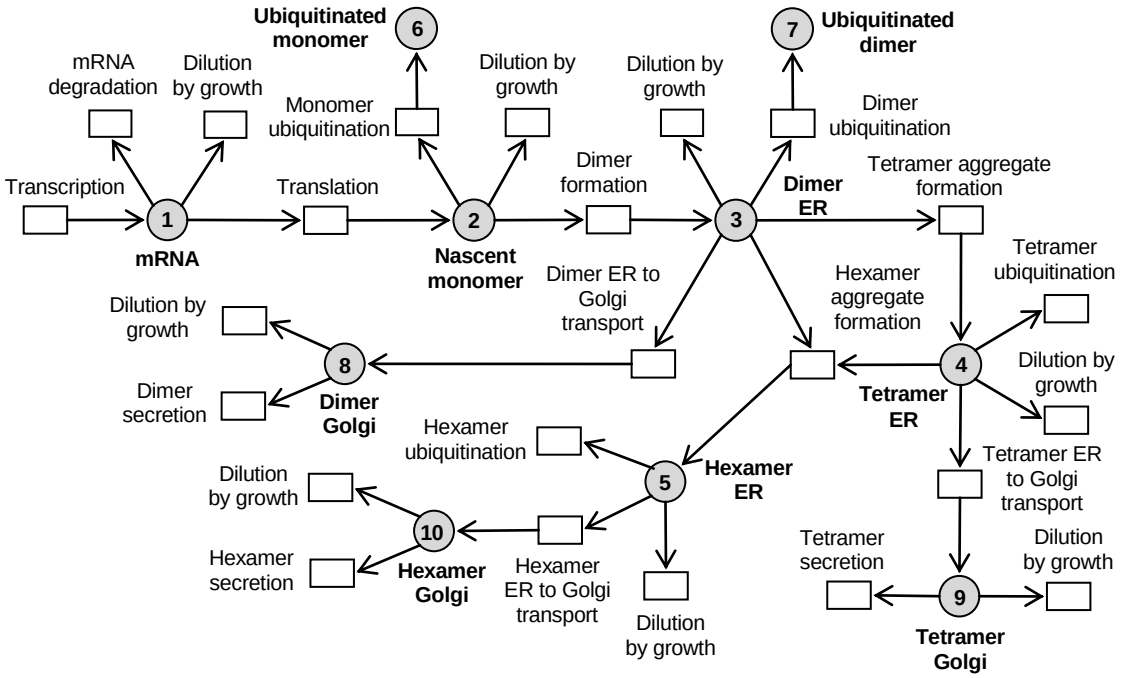

Figure 5.3: Schematic representation of the mathematical model of Sp35:Fc biosynthesis, aggregation, degradation and secretion pathway. Rectangles represent reactions (e.g. transcription) and grey circles represent species (e.g. mRNA). The numbers indicate the associated model's continuity equation described in the text. Dimer and aggregate formations were assumed to be irreversible events.

The structured model of the complete Sp35:Fc protein synthesis and secretory pathway is presented below. It describes the process of recombinant protein synthesis and secretion in the nucleus, ER and the Golgi complex for a homodimeric Fc-fusion protein-producing cell. The zero, first or second-order rate equation for a chemical reaction was applied in order to obtain the reaction (formation) rate based on the concentration of the reactants. For example, the formation rate of nascent Sp35:Fc monomer molecules was determined according to the formation kinetics, T_P and the mRNA concentrations, [m]. A description of the mathematical notations can be found in the Nomenclature section.

Firstly, the accumulation of Sp35:Fc mRNA molecules within a cell can be represented as follows:

$$\frac{d[m]}{dt} = N_{\text{gene}}\, S_m - k_{\text{deg}}\,[m] - \mu[m] \qquad (1)$$

where

$$k_{\text{deg}} = \frac{\ln 2}{t_{1/2,m}} \qquad (1a)$$

The intracellular balance of Sp35:Fc monomer, dimer, tetramer and hexamer in the ER are represented by the following equations:

$$\frac{d[\text{P}]}{dt} = T_{\text{P}}[m] - 2\,R_{\text{P}2} - k_{\text{ubq,P}}[\text{P}] - \mu[\text{P}] \qquad (2)$$

$$\frac{d[\text{P}_2]_{\text{ER}}}{dt} = R_{\text{P}2} - 2\,R_{\text{P}4} - R_{\text{P}6} - k_{\text{ER,P}2}[\text{P}_2]_{\text{ER}} - k_{\text{ubq,P}2}[\text{P}_2]_{\text{ER}} - \mu[\text{P}_2]_{\text{ER}} \qquad (3)$$

$$\frac{d[\text{P}_4]_{\text{ER}}}{dt} = R_{\text{P}4} - R_{\text{P}6} - k_{\text{ER,P}4}[\text{P}_4]_{\text{ER}} - k_{\text{ubq,P}4}[\text{P}_4]_{\text{ER}} - \mu[\text{P}_4]_{\text{ER}} \qquad (4)$$

$$\frac{d[\text{P}_6]_{\text{ER}}}{dt} = R_{\text{P}6} - k_{\text{ER,P}6}[\text{P}_6]_{\text{ER}} - k_{\text{ubq,P}6}[\text{P}_6]_{\text{ER}} - \mu[\text{P}_6]_{\text{ER}} \qquad (5)$$

where R is the rates of consumption of protein polypeptides in the assembly or aggregation process determined using second-order kinetics;

$$R_{\text{P}2} = k_{\text{fa}}[\text{P}][\text{P}] \qquad (3a)$$

$$R_{\text{P}4} = k_{\text{agg,P}4}[\text{P}_2][\text{P}_2] \qquad (4a)$$

$$R_{\text{P}6} = k_{\text{agg,P}6}[\text{P}_4][\text{P}_2] \qquad (5a)$$

and

$$k_{\text{ER,P}2} = \frac{\ln 2}{t_{1/2,[\text{P}2]\text{ER}}} \qquad (3b)$$

$$k_{\mathrm{ER,P4}} = \frac{\ln 2}{t_{1/2,[\mathrm{P4}]\mathrm{ER}}} \tag{4b}$$

$$k_{\mathrm{ER,P6}} = \frac{\ln 2}{t_{1/2,[\mathrm{P6}]\mathrm{ER}}} \tag{5b}$$

The ubiquitinated Sp35:Fc monomer and dimer in the ER are described as follows:

$$\frac{d[\mathrm{P}]_{\mathrm{ubq}}}{dt} = k_{\mathrm{ubq,P}}[\mathrm{P}] - \mu[\mathrm{P}]_{\mathrm{ubq}} \tag{6}$$

$$\frac{d[\mathrm{P_2}]_{\mathrm{ubq}}}{dt} = k_{\mathrm{ubq,P2}}[\mathrm{P_2}]_{\mathrm{ER}} - \mu[\mathrm{P_2}]_{\mathrm{ubq}} \tag{7}$$

where

$$k_{\mathrm{ubq,P}} = \frac{\ln 2}{t_{1/2,[\mathrm{P}]}} \tag{6a}$$

$$k_{\mathrm{ubq,P2}} = \frac{\ln 2}{t_{1/2,[\mathrm{P2}]}} \tag{7b}$$

The fully formed dimeric Sp35:Fc and tetramer and hexamer aggregates are then transported to the Golgi apparatus.

$$\frac{d[\mathrm{P_2}]_{\mathrm{G}}}{dt} = k_{\mathrm{ER,P2}}[\mathrm{P_2}]_{\mathrm{G}} - k_{\mathrm{G,P2}}[\mathrm{P_2}]_{\mathrm{G}} - \mu[\mathrm{P_2}]_{\mathrm{G}} \tag{8}$$

$$\frac{d[\mathrm{P_4}]_{\mathrm{G}}}{dt} = k_{\mathrm{ER,P4}}[\mathrm{P_4}]_{\mathrm{G}} - k_{\mathrm{G,P4}}[\mathrm{P_4}]_{\mathrm{G}} - \mu[\mathrm{P_4}]_{\mathrm{G}} \tag{9}$$

$$\frac{d[\mathrm{P_6}]_{\mathrm{G}}}{dt} = k_{\mathrm{ER,P6}}[\mathrm{P_6}]_{\mathrm{G}} - k_{\mathrm{G,P6}}[\mathrm{P_6}]_{\mathrm{G}} - \mu[\mathrm{P_6}]_{\mathrm{G}} \tag{10}$$

where

$$k_{\mathrm{G,P2}} = \frac{\ln 2}{t_{1/2,[\mathrm{P2}]\mathrm{G}}} \tag{8a}$$

$$k_{\mathrm{G,P4}} = \frac{\ln 2}{t_{1/2,[\mathrm{P4}]\mathrm{G}}} \tag{9a}$$

$$k_{\mathrm{G,P6}} = \frac{\ln 2}{t_{1/2,[\mathrm{P6}]\mathrm{G}}} \tag{10a}$$

Finally, biomass-specific secretion rates of Sp35:Fc native dimer and tetramer/hexamer aggregates are described by the following equations:

$$q\mathrm{P}_{\mathrm{dimer}} = \lambda_{\mathrm{P2}} k_{\mathrm{G,P2}}[\mathrm{P_2}]_{\mathrm{G}} \tag{8b}$$

$$q\mathrm{P}_{\mathrm{tetramer}} = 2\,\lambda_{\mathrm{P2}} k_{\mathrm{G,P4}}[\mathrm{P_4}]_{\mathrm{G}} \tag{9b}$$

$$q\mathrm{P}_{\mathrm{hexamer}} = 3\,\lambda_{\mathrm{P2}} k_{\mathrm{G,P6}}[\mathrm{P_6}]_{\mathrm{G}} \tag{10b}$$

5.2.2 Parameter fitting

The model was built and solved using Systems Biology Toolbox 2 (SBToolBox2; Schmidt and Jirstrand, 2006) for MATLAB software (MathWorks, Natick, MA). The full model syntax is presented in Appendix B. For each transfection condition the model of Sp35:Fc synthesis was solved using a steady-state assumption. All model parameters were fitted to the experimental data (Chapter 4) using a particle swarm optimisation method to yield the best-fit model parameters. The method utilises a stochastic pattern search algorithm to minimise the sum of squared errors between measurement (empirical data) and simulation outputs. A detailed description of the underlying algorithm is described in Vaz and Vicente (2007). A successful fit was defined to have a sum of squared errors <0.5.

5.2.3 Sensitivity analysis

A local sensitivity analysis was performed on the fitted model to calculate sensitivity coefficients for each cellular reaction/process involved in Sp35:Fc production. For each parameter, p_i, the response coefficient $R^{q\mathrm{P}}$ was numerically calculated using 1% perturbations as follows:

$$R_i^{q\mathrm{P}} = \frac{\left(q\mathrm{P}|_{1.01p_i} - q\mathrm{P}|_{0.99p_i}\right)}{\left(q\mathrm{P}|_{1.01p_i} + q\mathrm{P}|_{0.99p_i}\right)0.01} \qquad (5.1)$$

The response coefficient was calculated using a programme written within MATLAB software (Appendix C).

5.3 Results

5.3.1 Quantitative description of cellular processes

To further investigate which trafficking step is the main rate-limiting in the model Sp35:Fc biosynthesis and secretion, we utilised an empirically-based mathematical modelling technique to systematically analyse the influence of specific cellular processes on *q*P. For each rDNA load transfection an empirically derived mathematical model of Sp35:Fc production was built using the model topology outlined above (Figure 5.3) based on the previous work in this laboratory for MAb synthesis (McLeod et al., 2011; O'Callaghan et al., 2010). The parameters of each model were fitted to the empirical data using statistical

best-fit parameter sets determined using the particle swarm algorithm. A successful fit was defined to have a sum of squared residuals <0.5, where the residual is the absolute difference between the model calculated value in logarithmic space and the empirically derived level. This analysis produced a transfectant-specific set of estimated model parameter values representing the statistical best fit to the biological data.

The mean value of the sum of squared residuals was 0.127 with a standard deviation of 7.8% of the mean, indicating that the model-calculated species (e.g. mRNA) levels are highly comparable to the empirically derived data points. The fitted parameter values are tabulated in Table 5.1. The following assumptions and/or qualifications regarding the model were made. Firstly, it was assumed that the tetramer aggregate was an irreversible product of two dimer molecules and the hexamer aggregate must necessarily pass through a tetramer state before merging with another dimer. Secondly, as ubiquitinated tetramer and hexamer aggregates were below minimum detection threshold, dimer ubiquitination rates were used as the minimal degradation rates. The rate constants for ubiquitinated protein tetramer and hexamer in this study could be set to their nominal values as the proteins present in minute amount and did not affect the *q*P. Lastly, as Endo H sensitive and resistant high molecular weight Sp35:Fc could not be resolved using non-reducing SDS-PAGE and had to be analysed in the reduced form, the proportions of Sp35:Fc dimer, tetramer and hexamer in the ER and Golgi were assumed to follow the overall percentage of fully disulphide bonded oligomers in the ER/Golgi shown in Figure 4.6B of Chapter 4.

This set of model-derived parameter values (Table 5.1) is generally comparable to the previous reports from this laboratory using mathematical approaches to improve our understanding of the control of MAb production (Davies et al., 2011; McLeod et al., 2011; O'Callaghan et al., 2010; Pybus et al., 2014a) and does not deviate from the biologically plausible values. For instance, the transport rates (k_{ER} and k_G) are close to the typical rates for ER-to-Golgi and Golgi-to-medium transports in constitutive secretory pathways which have been reported to be 0.2–1.4 h^{-1} and 2–25 h^{-1}, respectively (Bibila and Flickinger, 1992; Dahm et al., 2001). Nevertheless, model-derived parameters are influenced by a wide variety of factors including the recombinant protein product, host cell line and cell culture process which can be expected to generate considerable model variation. This means mathematical models often require recombinant protein product, cell line and cell culture process-specific empirical measurements to optimise their usefulness as a tool to inform cellular constraints on productivity and/or cell engineering strategies.

Table 5.1: Best-fit parameter values for transfectants with different Sp35:Fc DNA loads. The nomenclature used is described in the Nomenclature section**.**

Parameter	Sp35:Fc plasmid DNA per 4.5×10^6 cells				
	0.25 µg	0.5 µg	1 µg	2 µg	3 µg
Cell growth					
μ (h^{-1})	0.0282	0.0259	0.0232	0.0226	0.0226
Transcription and translation					
S_m ((mRNA copy) $gene^{-1}$ h^{-1})	1.33×10^{-3}	1.33×10^{-3}	1.33×10^{-3}	1.33×10^{-3}	1.33×10^{-3}
k_{deg} (h^{-1})	0.058	0.058	0.058	0.058	0.058
T_P (molecule $(mRNA\ copy)^{-1}$ h^{-1})	253.9	177.1	110.9	59.5	42.0
Folding/assembly and aggregation					
k_{fa} (cell $molecule^{-1}$ h^{-1})	1.07×10^{-1}	1.79×10^{-2}	4.87×10^{-3}	9.32×10^{-4}	4.06×10^{-4}
$k_{agg,P4}$ (cell $molecule^{-1}$ h^{-1})	4.18×10^{-8}	3.70×10^{-8}	6.45×10^{-8}	6.03×10^{-8}	6.02×10^{-8}
$k_{agg,P6}$ (cell $molecule^{-1}$ h^{-1})	4.04×10^{-9}	1.04×10^{-8}	7.80×10^{-8}	8.03×10^{-8}	8.04×10^{-8}
Degradation					
$k_{ubq,P}$ (h^{-1})	–	–	3.63×10^{-4}	6.60×10^{-4}	9.17×10^{-4}
$k_{ubq,P2}$ (h^{-1})	2.33×10^{-5}	2.07×10^{-5}	2.53×10^{-5}	4.78×10^{-5}	7.92×10^{-5}
Transport and secretion					
$k_{ER,P2}$ (h^{-1})	1.044	0.468	0.267	0.233	0.224
$k_{ER,P4}$ (h^{-1})	1.234	0.503	0.188	0.156	0.137
$k_{ER,P6}$ (h^{-1})	1.288	0.557	0.274	0.200	0.128
$k_{G,P2}$ (h^{-1})	36.25	38.36	42.31	42.76	42.52
$k_{G,P4}$ (h^{-1})	43.24	41.53	38.15	37.94	43.32
$k_{G,P6}$ (h^{-1})	44.58	44.72	44.72	45.15	39.84

Overall, the model-derived parameters exhibit clear trends across the transfection conditions (Table 5.1). Whilst the cells managed to maintain the Sp35:Fc mRNA transcription rate S_m at high rDNA loads thus resulting in a proportional increase in the mRNA content (Chapter 4, Figure 4.4C), they were unable to retain the rate of post-transcriptional processes, specifically translation (T_P), folding/assembly (k_{fa}) and ER-to-Golgi transport (k_{ER}). Conversely, the rate of aggregate formation (k_{agg}) and protein degradation (k_{ubq}) were intensified at 1 µg rDNA load onwards. Therefore it was obvious that the cellular process rates were transfectant-specific depending on the recombinant gene copy abundance. Cells with 0.5 µg rDNA load characterise the most efficient cell factory via a fast/proper recombinant mRNA translation, folding and secretion in

conjunction with rapid cell proliferation rate, resulting in a relatively high product titre and low aggregate content.

5.3.2 Sensitivity analysis determines the dependence of Fc-fusion protein production on its cellular parameters

In order to understand how individual model parameters affect the *q*P (of native dimer), we performed metabolic control analysis (MCA) to obtain a quantitative description of substrate flux in response to changes in system parameters. MCA is a type of local sensitivity analysis that can be applied to protein synthesis and secretion pathways around a given steady state (Dimelow and Wilkinson, 2009; O'Callaghan et al., 2010). The empirical data obtained in Chapter 4 imply that the polypeptide synthesis processes (i.e. transcription/translation) would not have a significant effect at high rDNA loads, whereas post-translational activities would become limiting. We calculated response coefficients (RCs) for each cellular process involved in Sp35:Fc production which define the extent to which *q*P changes relative to a 1% perturbation of a discrete cellular process. This analysis is shown in Figure 5.4, where all parameters vary in a biologically meaningful manner.

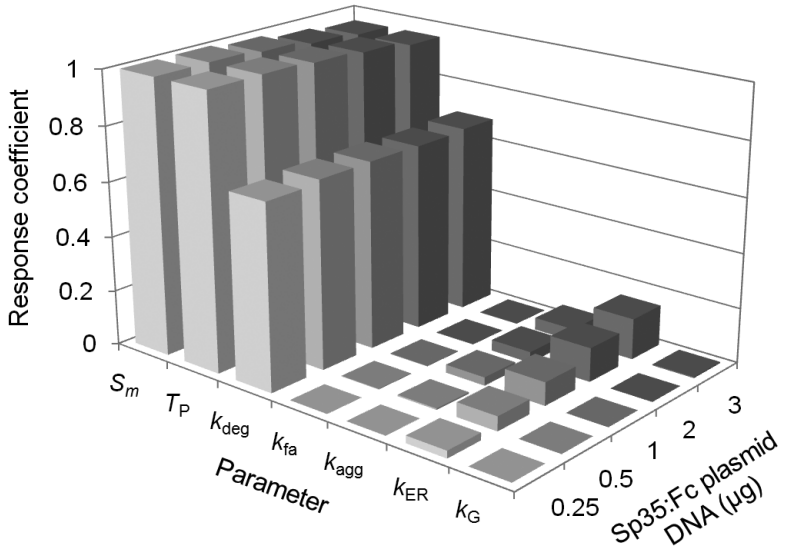

Figure 5.4: Parameter local sensitivity analysis demonstrating the *q*P response coefficient in each transfectant resulting from a 1% change in the rate constant of each synthetic process. Response coefficients were derived from a 1% increase in transcription rate (S_m), translation rate (T_P), folding/assembly rate (k_{fa}), ER-to-Golgi transport rate (k_{ER}) and Golgi-to-medium secretion rate ($k_{G,P2}$) and from a 1% decrease in mRNA degradation rate (k_{deg}) and aggregation rate (k_{agg}).

Overall, the intracellular transport rates of the Sp35:Fc molecule from the ER to the Golgi complex (k_{ER}) and subsequently to the extracellular medium (k_G) have little impact on the *q*P as reflected by the low response coefficients. This is in contrast to the protein synthesis processes, namely transcription (S_m) and translation (T_P). Indeed, transcription and translation parameters have been reported to display a universal, full control over recombinant protein production (e.g. Ho et al., 2006; O'Callaghan et al., 2010). Moreover, the mRNA degradation rate (k_{deg}), although not directly measured in this study, also showed significant influence on *q*P as reported by other studies (O'Callaghan et al., 2010). The difference in the response sensitivities may be explained by the fact that an increase in the rate of intracellular transport will lead to a depletion of the intracellular recombinant protein pool that is mainly regulated through the processes of transcription and translation, i.e. the increased transport rate cannot be sustained without a parallel increase in transcription and translation rates (Ho et al., 2006). Consequently, the effects of transport rates k_{ER} and k_G on increasing Sp35:Fc productivity tend to be significantly lower than those of the transcription (S_m) and translation (T_P) parameters.

However, the sensitivity analysis reveals that the ER-to-Golgi secretory rate constant (k_{ER}) became more sensitive at higher rDNA loads which exhibits a specific limitation in this process (Figure 5.4). This is consistent with the finding that high gene expression resulted in the retention of intracellular Sp35:Fc in the ER which eventually led to limited secretion rate. The consequence is also evident as observed earlier where an increase in intracellular dimer did not result in a corresponding increase in dimer secretion rate (Chapter 4, Figure 4.6A and B). Additionally, the aggregation rate constant (k_{agg}) also displayed sensitiveness, albeit at insignificant levels (RC ≤ 0.05), when the rate of aggregate formation competed with ER export machinery for the supply of newly folded, export ready Sp35:Fc in the ER lumen. In this case, the yield of native protein was determined by the rates of the competing first-order ER-Golgi secretion and second-order aggregation reactions. The formation of aggregates, however, did not make a considerable impact on the *q*P due to the fact that they made up only up to 5% w/w of the total secreted Sp35:Fc protein yield.

Unexpected results were observed with the transcription (S_m) and translation (T_P) processes where the high sensitivities were maintained at high rDNA loads (Figure 5.4). These results indicate, on the contrary to the empirical data (Chapter 4), that *q*P could be increased at high rDNA loads by these two processes. Equally, the response of the folding/assembly rate constant (k_{fa}) unpredictably appeared to be of negligible importance

even though the rate became notably low at 3 µg rDNA transfection (Table 5.1)—implying that the assembly process was not a rate-limiting step as hypothesised. With regard to k_{fa}, *in silico* retro-engineering test showed that the response coefficient became significant (RC > 0.1) only when the lowest reaction rate in Table 5.1 was further decreased from 10^{-4} to 10^{-7} cell molecule^{-1} h^{-1}. Such large reduction is biologically possible, where MAb heavy and light chain assembly rates are known to differ over six orders of magnitude (Bibila and Flickinger, 1992). This large variation is largely due to the nature of second-order assembly reaction compared to first-order reaction of transcription, translation, etc. Previous work in this laboratory (O'Callaghan et al., 2010) for instance had shown that GS-CHO cell lines with varying *q*P had assembly rate constant of heavy chain dimer between 10^{-7} and 10^{-1} cell molecule^{-1} h^{-1}. To understand these unexpected results, we performed *in silico* retro-engineering on the cellular parameters (see below).

5.3.3 *In silico* retro-engineering reveals the role of translational attenuation in maintaining cellular processes

In order to investigate as why the drop in dimer assembly rate k_{fa} at high rDNA load was relatively insignificant even though there was an apparent saturation in assembled dimer (Chapter 4, Figure 4.6B), we analysed the cell translational rate (T_P) across the transfectants (Table 5.1). The decrement in translational rate at higher rDNA loads suggests the effect of mRNA translational attenuation in the cells. Increased levels of unfolded and/or misfolded proteins in the ER of all eukaryotes are known to trigger the UPR which in turn activates global translational attenuation (Chakrabarti et al., 2011). Therefore, this translational attenuation enabled the unfolded nascent monomers (Chapter 4, Figure 4.6B) to increase almost proportionally to the mRNA copy numbers (Chapter 4, Figure 4.4C) or else there would be a rapid accumulation of unfolded monomers due to the decrease in dimer assembly rate (note the slightly exponential curve of the monomer species in Figure 4.6B). In other words, the translational attenuation mechanism cushioned the impact of saturated assembled dimer in the ER—resulting in only a relatively small decrease in the assembly rate and at the same time keeping the unfolded monomer at an appropriate level in order to retain the cell's normal cellular function.

If our hypothesis is correct, the mathematical model would predict that without the translational attenuation (i.e., invariable translational rate at different rDNA loads) the folding/assembly rate would decrease drastically as the saturation of intracellular dimer

would be solely due to the folding/assembly factor, and this in turn would result in an adverse accumulation of the unfolded monomer. To verify this, we performed *in silico* cell engineering on the empirical/experimental data (Figure 5.5A) by artificially keeping the translational rate T_P at different rDNA transfections constant as illustrated in Figure 5.5B. The model was simulated to see the changes in the monomer amount and then re-fitted as described earlier to obtain a new folding/assembly rate k_{fa} for each transfectant.

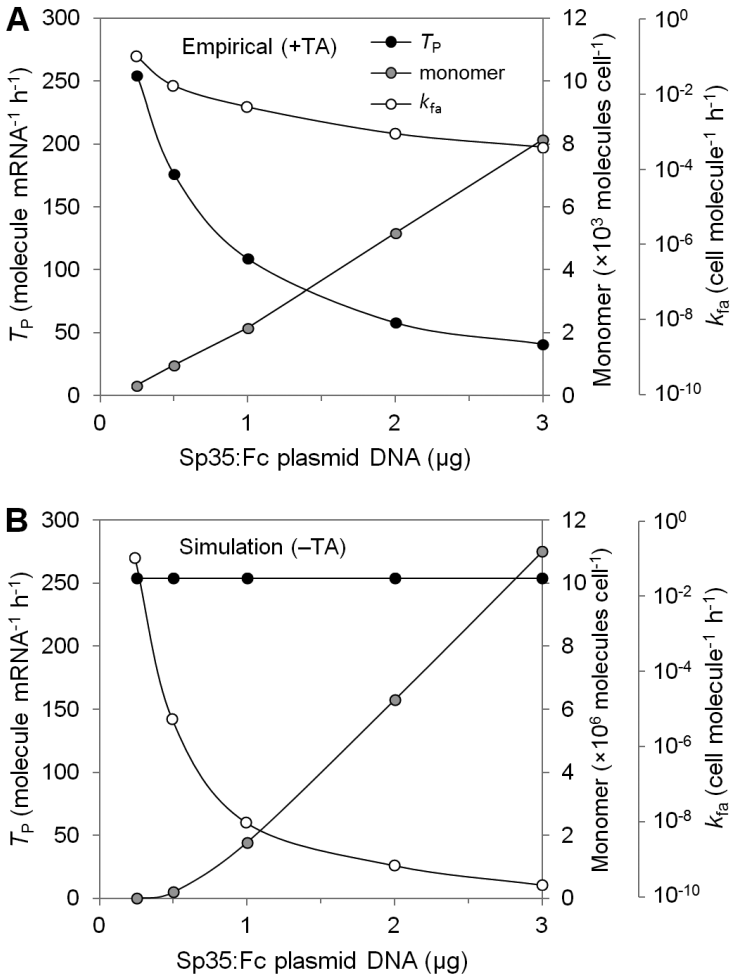

Figure 5.5: *In silico* cell engineering reveals the intrinsic ability of cells to restore homeostasis between folding demand imposed on the ER and its folding capacity. (A) The empirical data showed that a translational attenuation (TA) mechanism, i.e. a decrease in translational rate (T_P), enabled nascent monomers to increase linearly and resulted in an insignificant decrease in the folding/assembly rate (k_{fa}) from 10^{-1} to 10^{-3}. **(B)** By artificially keeping the translational rate T_P at different rDNA loads constant to mimic the absence of translational attenuation, the accumulation of unfolded monomers increases exponentially. Without the translational attenuation mechanism to prevent the rapid accumulation of unfolded monomers, the folding/assembly rate decreases by nine orders of magnitude from 10^{-1} to 10^{-10}.

In agreement to the hypothesis, the simulation outcome demonstrated that at a constant translational rate (T_P) of 254 molecule mRNA^{-1} h^{-1}, the folding/assembly rate (k_{fa}) decreased dramatically from 10^{-1} to 10^{-10} cell molecule^{-1} h^{-1} as the amount of monomer increased rapidly (Figure 5.5B). The sensitivity analysis also revealed that the response coefficient of folding/assembly rate constant became notably significant (RC up to 0.72; Figure 5.6) indicating that in the absence of translational attenuation the folding/assembly process is now a rate-limiting step in the Sp35:Fc production mechanism. Conversely, the response coefficient of the transcription rate S_m and the translation rate T_P became progressively lower, indicating that the parameters exerted less control at higher rDNA loads. To this end, the model predicts that engineering the folding/assembly could produce a greater return in terms of *q*P—much in agreement to our hypothesis.

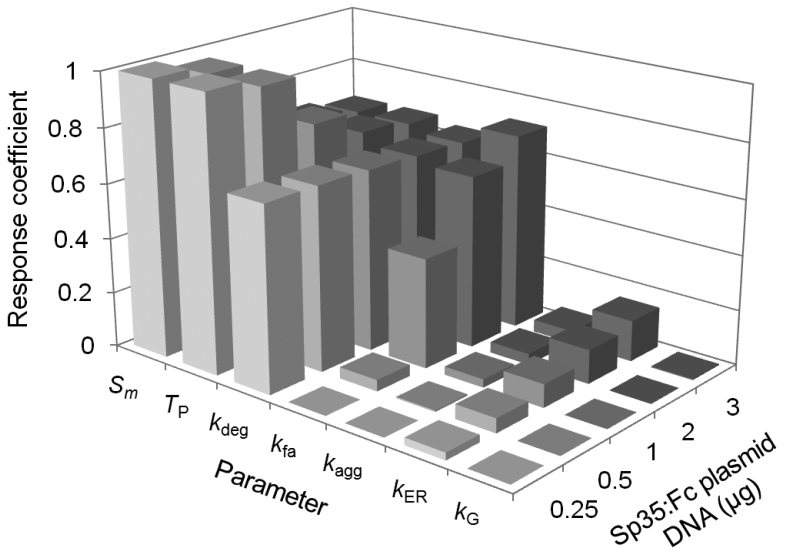

Figure 5.6: The response coefficient pattern changes when the translational attenuation mechanism of the cells is artificially removed via *in silico* engineering. Without cell's translational attenuation mechanism, the folding/assembly rate (k_{fa}) becomes highly sensitive at high rDNA load. Conversely, the sensitivity of transcription rate (S_m) and translation rate (T_P) is reduced implying that the parameters exercise less control at higher rDNA loads.

5.4 Discussion

Two imperative outcomes of the mathematical modelling and sensitivity analysis are (i) the identification of the major cellular production constraint at elevated expression and (ii) a systematic understanding of how cells cope with intracellular stress. With respect to the latter, it is important to note that cellular rates of mRNA translation and protein folding are

functionally linked by complex intracellular signalling pathways such as the UPR which is not readily perceived by empirical data alone. Mathematical modelling therefore allows the organisation of experimental information systematically and coherently, which in turn facilitate the analysis and identification of essential, qualitative features in complex biological systems. Our study highlights that UPR regulates the protein-folding capacity of the ER according to cellular demand. As such, it gives further insights toward understanding how ER stress sensors can confer different responses to achieve optimal cellular function and their impact on DTE protein production rate.

A persistent UPR, however, would be highly undesirable in a stable production cell line as it would reduce protein synthesis and lead to ERAD and ultimately cell apoptosis. Previous studies in this laboratory comparing the proteomes of MAb-producing GS-NS0 cell lines showed that cells with relatively high *q*P had a considerably higher abundance of ER functional proteins including the chaperones BiP, PDI and GRP94. Despite the variations in ER resident proteins, the levels of spliced XBP1 and cleaved ATF6 were similar in all cells lines studies, i.e. the UPR was not induced in cells with high *q*P, suggesting that stable cell lines utilise some cellular strategies to avoid or minimise UPR activation (Dinnis and James, 2005; Dinnis et al., 2006; Smales et al., 2004). Thus we predict that the generation of stably transfected cells with >1 µg Sp35:Fc DNA using the present system will mostly result in instability of Sp35:Fc production (loss of protein production, protein aggregation, cell death, etc), rendering the clone screening and selection processes difficult and with a low probability of success.

Apart from systematic understanding, mathematical models can be utilised for the effective control and optimisation of bioprocess performance. Using our model, we propose that the rational design of cell engineering strategies to improve volumetric Sp35:Fc product yield for this system would be increases in (proper) folding/assembly rate, for example through metabolic and cell engineering approach. This is comparable to the previous study in this laboratory modelling transient difficult-to-express MAbs (Pybus et al., 2014a), and in contrast to easy-to-express MAb production in stable cell lines which emphasised on transcription and translation processes (Davies et al., 2011; McLeod et al., 2011; O'Callaghan et al., 2010) as rational targets to improve productivity. An increase in Sp35:Fc translation rate (and recombinant mRNA abundance) without improving the folding and assembly machinery would exacerbate the accumulation of nascent polypeptides in the ER and destroy cell's translational attenuation control capability, and is likely to shift the bottleneck further downstream in the Sp35:Fc secretion pathway.

Finally, while sensitivity analysis is a valuable tool to evaluate the control of recombinant protein synthesis exerted by specific cellular processes and aid rational engineering strategy, it does not represent the effect *per se*. The model can be further utilised/developed as a predictive tool for cell engineering and/or cell line selection. For example, using the observed range of parameter values across the transfectants (Table 5.1) as biologically possible cellular processing rates within this CHO host cell, we created a theoretical engineered cell line *in silico* using the optimal translation, folding/assembly and secretion rates. The modelled *q*P increased from 0.18 to 1.17 pg cell^{-1} day^{-1} for a 3 µg rDNA load transfectant, a substantial 6.5-fold increase. Certainly, for another Fc-fusion protein product or cell line different limitations on cellular productivity may pertain. We propose that this iterative cell factory reverse-engineering approach could be implemented as a generic framework for Fc-fusion proteins, with the objective of using it as an initial groundwork to better understand the factors underlying Fc-fusion protein synthesis and processing, and subsequently propose predictive, rational engineering strategies to overcome the limiting production factors in mammalian cultivation systems.

Chapter 6

Harnessing Functional Heterogeneity within CHO Cell Populations for Super-Producers of a Difficult-to-Express Fusion Protein

This chapter explores the hypothesis that it is feasible to exploit functional/genetic heterogeneity in the parental CHO cell population to select clonal variants with desirable characteristics. The aims are to identify functional variations that may contribute to enhanced production of difficult-to-express Sp35:Fc protein, and eventually generate new cell lines that are inherently more fit for this purpose.

6.1 Introduction

Mammalian cell populations are highly and inherently heterogeneous. Discrepancies among clonal cells are present at both genetic (Derouazi et al., 2006; Oh et al., 2003) and phenotypic levels (Barnes et al., 2006; Pilbrough et al., 2009). Stockholm et al. (2007) have demonstrated, using both *in silico* and *in vitro* approaches, that the phenotypic heterogeneity in a clonal mammalian cell population can arise due to "extrinsic" and "intrinsic" mechanisms of the cells. Even though the impacts can be unfavourable (e.g. loss of productivity; Lee et al., 1991) and difficult to be dealt with, it is also desirable to find out which components of the observed variation serve a biological function that may be harnessed for cell line development. In this context, genetic heterogeneity allows the selection of cell lines with advantageous phenotypes as a complement to or replacement for cell engineering method.

The potential exists in many cases to select natural variants with a desired feature, either as candidate lines for production or to provide a platform for the analysis of factors contributing to improved characteristics. Whilst cell selections typically focus on productivity and cell growth/density (Pichler et al., 2011; Prentice et al., 2007), novel variant cell lines have also been selected with the ability to consume lactate (Browne and Al-Rubeai, 2011), the ability to grow in glutamine-free medium (Hernández Bort et al., 2010), the ability to be transfected with adenovirus (Condon et al., 2003), resistance to oxidative stress (Keightley et al., 2004) and resistance to high medium osmolarity (Liu et al., 2010). Unlike cell engineering method, this approach does not require extensive knowledge of the genes and cellular mechanisms of the cells which can be a bottleneck in cell development process (Dietmair et al., 2011; Dinnis and James, 2005). However, to our knowledge, there are no reports describing how genetic/phenotypic heterogeneity can be purposely managed or controlled to produce improved mammalian cell hosts specifically for difficult-to-express (DTE) protein production.

In this study, we evaluate the functional performance of clonally derived Chinese hamster ovary (CHO) cell populations to produce DTE Sp35:Fc in transient expression systems. This is theoretically possible as mammalian host cells have been shown to vary considerably in their intrinsic ability to manufacture a given recombinant protein independent of transgene copy numbers (Charaniya et al., 2009; Kennard et al., 2009). Nevertheless, previous studies analysing clonally-derived cell populations transiently or stably expressing MAb and/or green fluorescent protein (GFP) have also demonstrated that a functional trait is subject to stochastic cell-to-cell variation (Merritt and Palsson,

1993; Pilbrough et al., 2009) and not always heritable (Davies et al., 2013; Pichler et al., 2011). With this in mind, we purposely utilise population-averaged measurements taken over many generations (i.e. extended serial subculture) to assess the heritability of key functional traits, in which temporary cell-to-cell phenotypic variability that may arise from micro-environmental differences (Snijder and Pelkmans, 2011) or stochastic gene expression (Raj and van Oudenaarden, 2008) is effectively irrelevant. Somewhat on the contrary, we also hypothesised that the routine cell culture procedures are a logical route to effectively increase genetic and functional variations within and between sub-populations that in turn may increase the probability of selecting new cell lines with superior capabilities to produce DTE Sp35:Fc. Over many generations, the physically isolated cultures evolve randomly/independently and undergo genetic drift to yield their own heterogeneous parental CHO cell populations (Merlo et al., 2006)

The CHO-S cell lineage used in this study had been previously bioreactor adapted as a pool by Biogen Idec and remained that way subsequently (i.e., not sub-cloned) with the idea of providing as much population heterogeneity as possible (Marty Sinacore, personal communication). Taking advantage of this high heterogeneity potential, we isolate 70 clonal variants and subject a subset of 33 untransfected clonal populations to parallel prolonged culture (over 120 generations) to intensify the naturally occurring variations. We evaluate the variation/heritability of key phenotypic traits, namely specific proliferation rate (μ), peak viable cell density VCD, cell size and cell biomass content. By expressing Sp35:Fc and GFP in a sub-subset of 16 clones at 3 different generations, we critically assess the effect of routine cell culture procedures on the clones' functional performance in manufacturing the two distinct recombinant proteins. We effectively identify several CHO-S host cell variants that possess superior cellular machinery for transient Sp35:Fc production as well the general cell characteristics that could be beneficial in producing DTE proteins.

6.2 Materials and Methods

6.2.1 Isolation and long-term culture of CHO-S clones

Single cell clones were isolated by one round of limiting dilution cloning (LDC) of the parental CHO-S cell line (Biogen Idec, Cambridge, MA). Cells were plated out into Greiner CellStar® 96-well plates (Greiner Bio-One, Frickenhausen, Germany) at 0.125 cells per well in CD-CHO medium supplemented with 8 mM L-glutamine (Life Technologies,

Paisley, UK). The probability of clonality by LDC was calculated according to Clarke et al. (2002) using the following equation:

$$P(\text{clonality}) = \frac{P\ (\text{one colony per well})}{P\ (\text{at least one colony per well})} = \frac{sd \times e^{-sd}}{1 - e^{-sd}} \quad (6.1)$$

where *sd* is the seeding density. Individual wells were further subjected to image analysis (CloneSelect™ Imager, Genetix, Hampshire, UK) to ensure selection of clonal populations derived from a single cell. Plates were incubated in a humidified incubator at 37°C in 5% CO_2 and culture medium was replenished on Day 9. On Day 13, static culture of 70 colonies at 40–70% confluency were scaled up to 24-well plate at 5–10% seeding confluency (3 days culture) followed by 6-well plate (3 days culture), and eventually to 125 mL shake flasks (Corning Incorporated, Acton, MA) at which the generation number was reset to zero. Shake-flask cultures were maintained at 37°C in 5% CO_2 under 140 rpm agitation. Cells were sub-cultured every 4 days by centrifuging at 200×g for 5 min to remove the spent medium and re-seeding into fresh medium at a concentration of 2×10^5 cells mL^{-1}. Accumulated generation number was calculated using Equation 6.2;

$$\text{Generation number} = \frac{\ln\left(X_2/X_1\right)}{\ln(2)} = \frac{\mu\left(\Delta t\right)}{\ln(2)} \quad (6.2)$$

where *X* are the viable cell density (VCD) at the first or second time point, μ is the cell proliferation (growth) rate and *t* is the time. Cryopreservation in liquid nitrogen was carried out on the clonally derived cell populations at 15, 45 and 105 generations as described in Section 3.1.2.

6.2.2 Determination of cell size

Average cell diameter based on forward scatter was determined using a Vi-CELL™ Cell Viability Analyser (Beckman Coulter, Brea, CA). Diameter range was determined to be between 6 and 24 µm to exclude cell debris and aggregates (Seewöster and Lehmann 1997). 5,000–8,000 cells at mid-exponential phase were used for the analysis and duplicate readings were taken from two passages. The average cell diameter (*D*) was used to calculate the average cell volume (*V*) by assuming a spherical shape:

$$V = \frac{4}{3}\pi\left(\frac{D}{2}\right)^3 \quad (6.3)$$

6.2.3 Quantification of total cellular protein

Total cellular protein content was determined from whole cell lysates. Triplicate samples of 1×10^6 cells were washed twice in DPBS (Sigma-Aldrich, Poole, UK) then immediately stored at –20°C. Prior to extraction, samples were rapidly thawed on ice and then lysed in 100 µL RIPA buffer (50 mM Tris–HCl, pH 8.0, with 150 mM sodium chloride, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate (Sigma-Aldrich)) supplemented with 1X Protease Inhibitor Cocktail Set III (Merck Chemicals, Nottingham, UK) at 4°C for 10 min. Samples were clarified by centrifugation at 8,000×g for 10 min at 4°C. Soluble protein concentration was determined using Pierce™ BCA Protein Assay Kit (Thermo Scientific, Cramlington, UK) microplate procedure according to the manufacturer's instructions using BSA as a calibrant. Samples were diluted 1:8 in working reagent and concentrations determined using a PowerWave™ XSTM (Bio-Tek, Potton, UK) microplate reader at 570 nm. KC4 3.1 software (Bio-Tek) was used to generate the standard curve and interpolate the protein concentration in cell lysate (Figure 6.1). The average protein biomass content per cell was calculated from the total protein amount and the number of cells lysed.

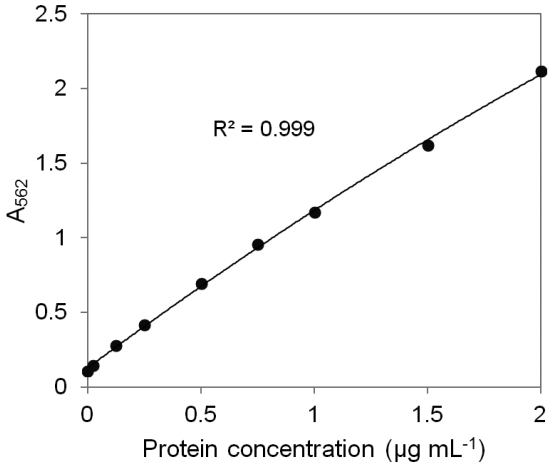

Figure 6.1: Example standard curve generated using BCA protein assay. The curve generated was used to determine the protein concentration in whole cell lysate.

6.2.4 Transient transfections

33 selected clones and the parental cell line were revived from the cell bank and seeded at 3×10^5 cells mL^{-1} in 125 mL Erlenmeyer shake flasks. Prior to transfection, clonally derived

cell populations were maintained for 4 passages (15 generations). All clones utilised CD-CHO medium supplemented with 8 mM L-glutamine at 37°C in 5% CO_2 and 140 rpm agitation. For a small-scale production, transient transfection was conducted at Day 2 of culture and each clone was transfected in triplicate using the Amaxa® Nucleofector® 96-well Shuttle™ protocol for CHO-S cells (Lonza, Basel, Switzerland). Briefly, 1.5×10^6 cells per well were centrifuged at 90×g for 10 min. Cell pellet was resuspended in 20 µL Nucleofector solution SG and transferred onto the Nucleocuvette® plate with 0.33 µg of Sp35:Fc DNA. Following transfection, cells were diluted in 100 µL of culture medium and transferred into a 6-well plate at a seeding density of 1×10^5 cell mL^{-1}. Cells were cultured in CD-CHO supplemented with 8 mM L-glutamine at 3 mL/well working volume and incubated in a static incubator at 37°C, >85% humidity and 5% CO_2 for 48 h prior to Sp35:Fc quantitation. For a larger scale suspension culture production, a subset of 16 clones were co-transfected with 1 µg Sp35:Fc plasmid and 1 µg pmaxGFP® vector (Lonza) using the Amaxa Cell Line Nucleofector Kit V system (Lonza) and cultured in a TubeSpin for 96 h as described in Section 3.3.1.

6.3 Results

6.3.1 Clones isolated from a parental population exhibit significantly different phenotypic characteristics

CHO-S clones were derived from the parental cell line in static microplate culture using a single LDC procedure at a probability of clonality of 0.939. The presence of a single cell-derived colony was identified using automated image analysis of individual wells. After 19 days of static microplate culture, 70 individual subpopulations were transferred to shake flask suspension culture with a strict 4-day sub-culturing regime in which clone-specific proliferation rate (μ) and accumulated generation number were routinely calculated. At 30 generations of suspension culture, the clonal populations exhibited 1.5-fold variation in μ and 2.6-fold variation in peak VCD and there was a significant positive correlation between the two traits (Pearson's product moment correlation coefficient, PPMCC r = 0.849, P < 0.0001; Figure 6.2). As we were parsimonious in our approach to study the heritability/evolution of clone functional phenotypes, a subset of 33 clones with varying μ and peak VCD were selected from this data and subjected to parallel prolonged culture (>120 generations, up to 100 days).

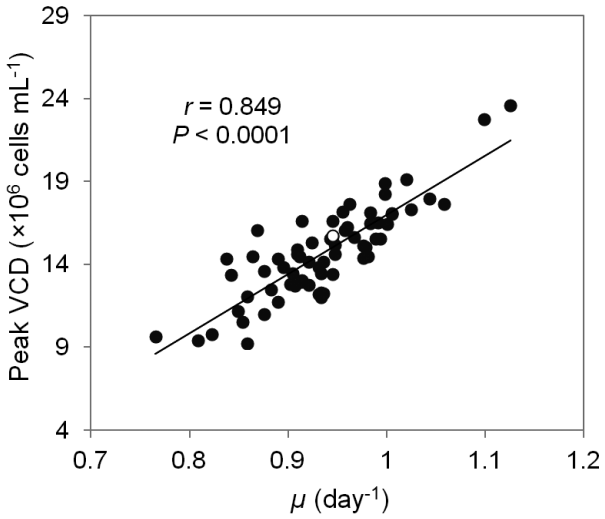

Figure 6.2: CHO-S clonal populations isolated by limiting dilution cloning exhibit large variations in phenotypic traits. At 30 generations of suspension culture, the specific proliferation rate, μ, varied by 1.5-fold while the peak viable cell density (VCD) varied by 2.6-fold. The two traits are also positively correlated (PPMCC r = 0.849, P < 0.0001). The donor CHO-S parental cell population (white symbol) had an average μ and peak VCD of the clones.

The clonally-derived cell populations varied markedly in μ throughout long-term suspension culture (Figure 6.3). This was most obvious at the first passage immediately following transfer from static into suspension (μ ranged from 0.398 to 0.914 day^{-1}, or 2.3 to 5.3 generations), although by the third passage, clone-to-clone variation in μ had narrowed to 0.655–1.037 day^{-1} (3.8–6.0 generations). Additionally, the majority of clones displayed larger fluctuation in μ early in suspension culture and gradually became stable. Despite the general increase and stabilisation in μ during the long-term culture, the isolated clonal variants did not necessarily achieve the same growth fitness under the governing selective pressure (i.e., reach the maximum proliferation rate achievable by other variants). At 120 generations, the μ still varied between 0.951 and 1.176 day^{-1} (up to 24% difference).

In order to systematically quantify and compare the rate at which clone specific μ converged during the long-term suspension culture, the calculated values of μ for each clone during serial sub-culture were fitted to either a linear, second-order polynomial or logarithmic regression model depending upon the line that best fitted each data (mean R^2 = 0.605). For each clone, we then determined the average specific rate of increase in μ (% per generation) between 5 and 120 generations in suspension culture using values for μ derived from the line of best fit and Equation 6.4.

$$\Delta\mu = \left(\left(\frac{\mu_{120}}{\mu_5}\right)^{(1/(120-5))}\right) - 1 \qquad (6.4)$$

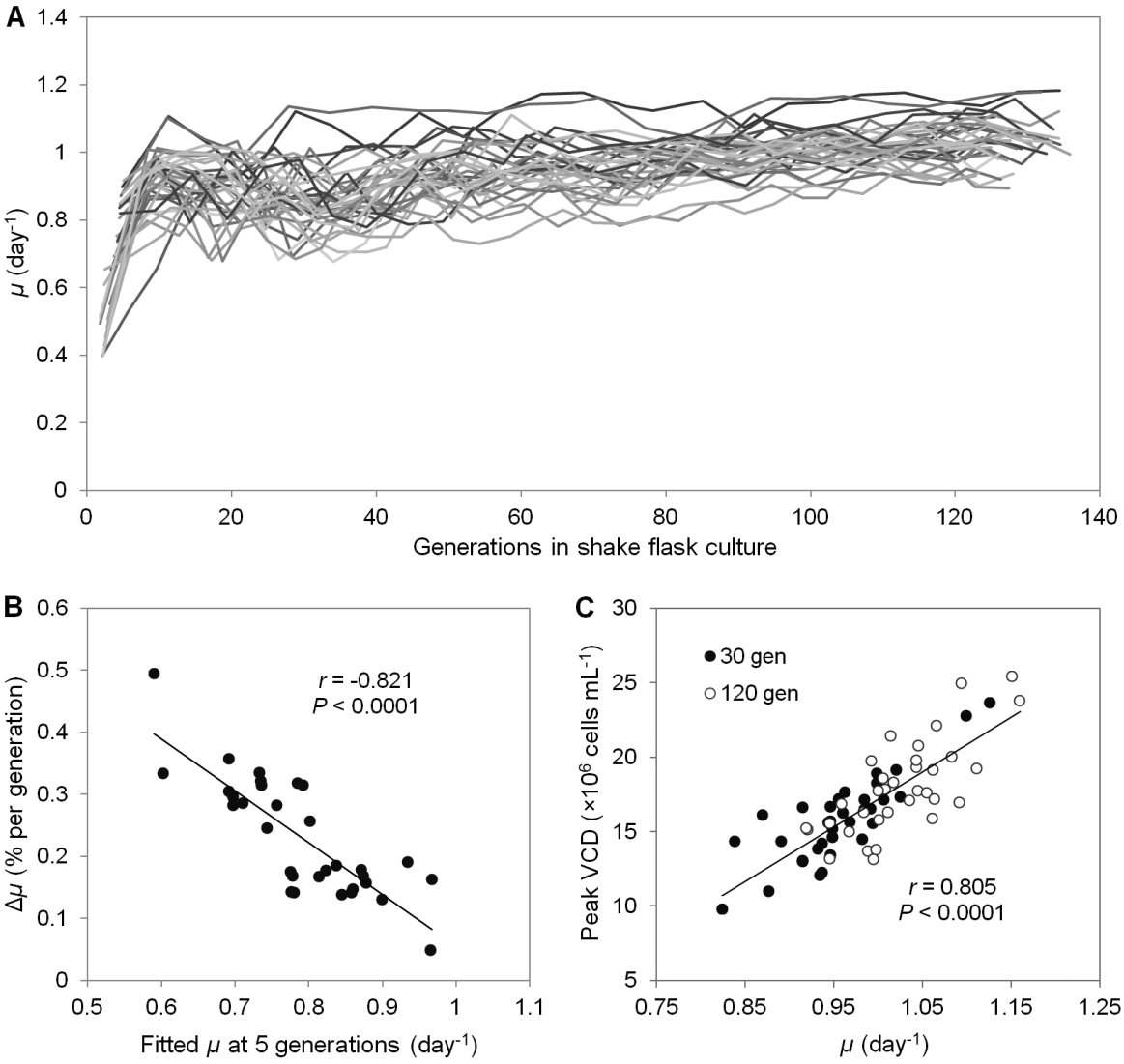

Figure 6.3: CHO-S clones exhibit great variation in specific proliferation rate and peak viable cell density during long-term shake flask culture. (A) Specific proliferation rate (μ) of 33 individual CHO-S clones and the parental cell line during shake flask culture for >120 generations. The clones display a general increase in μ over time. **(B)** For each clonal isolate in **A**, calculated values for μ during serial sub-culture were fitted to the best fit model in each case. Plot shows that the clone-specific rate of change of μ ($\Delta\mu$) is inversely proportional to initial μ (fitted μ at 5 generations). **(C)** Fast-growing clones tend to have higher peak viable cell density (VCD), and as the clones increase their μ between 30 and 120 generations, their peak VCD also increases.

where μ_5 is fitted μ at 5 generations in shake flask culture and μ_{120} is fitted μ at 120 generations in suspension culture. Figure 6.3B shows that the clone-specific rate of change in μ varies greatly between 0.048 and 0.494% per generation (median $\Delta\mu$ = 0.118% per generation), and that there was a strong negative correlation between initial μ (fitted at 5 generations) and the subsequent average rate of increase in μ (PPMCC r =

–0.821; $P < 0.0001$). Together, the data in Figure 6.3A and B corroborate a recent study in this laboratory using CHOK1SV cells (Davies et al., 2013) that whilst clones further from the maximal observed μ at the beginning of serial-sub culture tend to increase their μ more rapidly than those with an intrinsically higher initial μ, clonal variation in growth characteristic persisted throughout prolonged culture.

Interestingly, 12 clones and the parental cell line demonstrated significant increases in peak VCD ($P < 0.001$; mean increase 4.25×10^6 cells mL^{-1}) as they acquired their optimal μ after 120 generation, while 10 clones displayed slight improvements in peak VCD ($P < 0.05$; mean increase 1.69×10^6 cells mL^{-1}). The rest of the clones (12 clones) did not show significant changes ($P > 0.05$). This data is illustrated in Figure 6.3C which shows a general shift in μ and peak VCD from 30 generations to 120 generations. We note that at both points of culture there was a positive correlation between μ and peak VCD (overall PPMCC $r = 0.805$, $P < 0.0001$). As the population-averaged cell culture processes and measurements are devoid of population and temporal noises (Brock et al., 2009), we deduce that the majority of clonally derived populations gradually underwent genetic drift where fast proliferating clonal variants were in sufficient abundance to permit their selection during serial sub-culture procedures and eventually became predominant.

6.3.2 Variations in clone-specific proliferation rate are due to variations in cell biomass and size

To study if the observed variations in clone-specific proliferation rate were correlated to differences in cell biomass and volume control, both parameters were examined using the 33 clonal populations and the parental population that had undergone 60 generations. We reasoned that at this generation number, the cell populations were relatively stable (i.e., less susceptible to stochastic temporary variations in μ) and had minimal phenotypic diversity within a population. Clones were revived from the cell bank at 45 generations, cultured for 4 passages (approximately 15 generations) then sampled simultaneously at mid-exponential phase (Day 4). Average cell spherical volume was determined by image analysis (measured as mean diameter in µm of 5,000–8,000 cells per sample) and total extracted cellular protein was determined by biochemical assay as an estimate of mean cell biomass content. These data are shown in Figure 6.4.

Data from the biochemical assay revealed the variation in cell protein content (1.42-fold) was positively correlated to cell volume (PPMCC $r = 0.718$; $P < 0.0001$, Figure 6.4A)

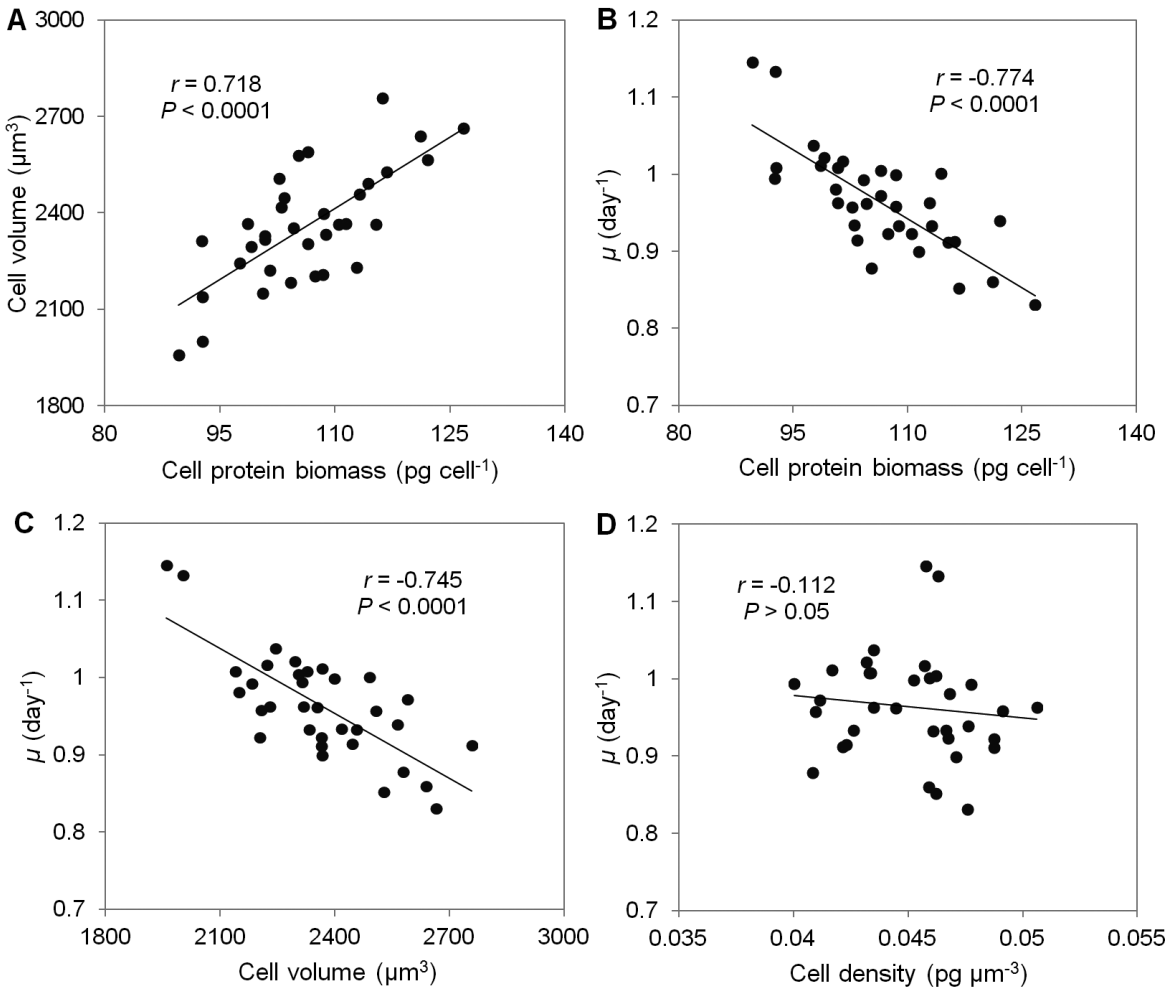

Figure 6.4: Relationship between the specific proliferation rate, protein biomass content, size and density at 60 generations. Clones were sampled simultaneously 96 h post-inoculation. **(A)** Mean cell protein biomass of the clones positively correlates with mean cell size (volume) determined by automated image analysis, and **(B)** negatively correlates to clone-specific proliferation rate, μ. **(C)** μ is also influenced by cell size, but **(D)** not by cell density.

and was negatively correlated to μ (PPMCC r = –0.774; P < 0.0001; Figure 6.4B). Additionally, we note a negative correlation between μ and cell volume (PPMCC r = –0.745; P < 0.0001; Figure 6.4C). In other words, those clones that had large mean biomass and volume essentially required longer time to accumulate their biomass (lower μ) whereas clones with small mean cell biomass and volume required shorter time (higher μ). These data imply, in contrast to Davies et al. (2013), that the biomass accumulation rate of the clones are a largely similar of around 92–115 pg cell^{-1} day^{-1} (1.25-fold variation). Accordingly, the variation in cell density (1.27-fold) was less clear than that of size (1.41-fold) and there was no relationship between the former and μ (PPMCC r = –0.112; P > 0.05; Figure 6.4D). Nevertheless, at the extremes of the observed diversity, the clonal

phenotypes could also be either "biomass efficient" (high cell protein content and high μ) or "biomass inefficient" (low cell protein content and low μ) that could potentially impact biomanufacturing processes. Significant clone-to-clone heterogeneity in size/biomass control was therefore evident.

6.3.3 Impact of cell evolution on recombinant protein productivity and heritability

To narrow down the number of clones to be tested for Sp35:Fc expression capabilities, preliminary small-scale transient transfections were performed using plate-based Nucleofection system on the subset of 33 clones and the parental population (all at 30 generations). Cell line specific recombinant protein production has been demonstrated to be significantly less variable using this transfection method (i.e. Nucleofection) than for lipoplexes (Davies et al., 2013) and polyplexes (Thompson et al., 2012) mediated transfections, as well as other electroporation methods (Zeitelhofer et al., 2007). After 48 h of static culture, the supernatant from each transfectant pool was analysed using ELISA. This analysis shows that the Sp35:Fc production capability of the clones vary by up to 2-fold ($P < 0.0001$; Figure 6.5).

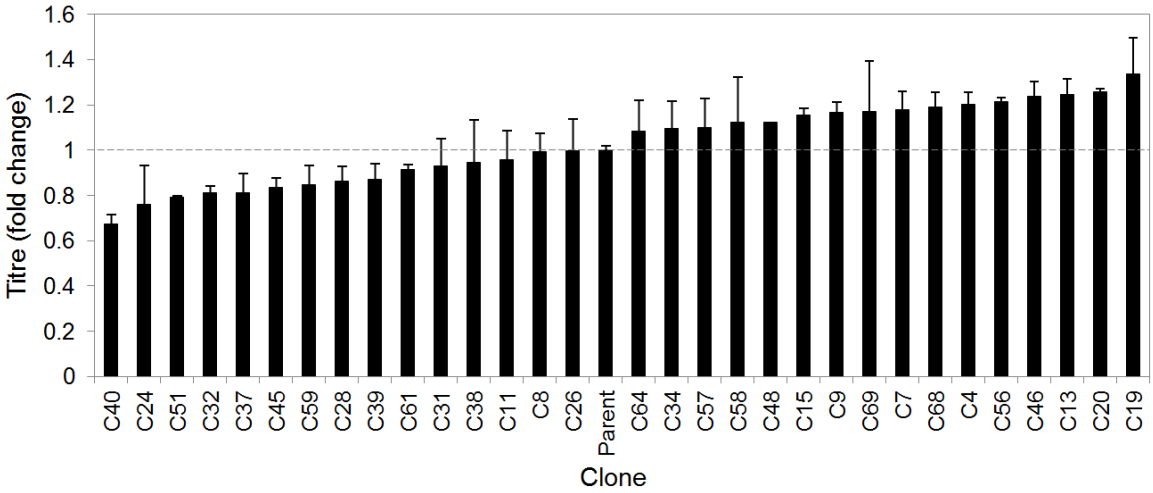

Figure 6.5: CHO-S clones exhibit variation in transient difficult-to-express Sp35:Fc protein production. 33 clones and the parental population at 30 generations were each transfected in triplicate using plate-based Nucleofection system. The extracellular Sp35:Fc titre was measured by ELISA after 48 h of static culture. Clone-specific Sp35:Fc production is shown relative to that of parental CHO-S cell line used to normalise assays. Error bars represent the standard deviation of three technical replicates.

Based on the data in Figures 6.3A and 6.5, a sub-subset of 16 clones with varying proliferation rates and volumetric titres were selected. This is akin to the traditional screening of the highest stable producers in static culture following cloning before analyses in suspension culture. We acknowledge that there are differences in static and suspension cultures (e.g. oxygen transfer rate, cell confluency/density) that could affect cells behaviour and recombinant protein production. Therefore, cells with low productivities/titres in static culture might actually exhibit high production in suspension mode, or vice versa. Despite the differences, studies have shown that there are still good confidence intervals—for example, if one were to obtain the highest producer in suspension culture, the selection of the top 30% of the highest producers in static culture would give a 1% probability of incorrect rejection (Brand et al., 1994). As we were to select 50% of the clones, this small-scale production data gave a higher chance in selecting high, mid and low producers than relying on the cell proliferation data alone.

For each clone (from the sub-subset) as well as the parental population, transient transfections were performed at 30, 60 and 120 generations and cultured in suspension. At the latter generation number, the clonal populations were the most "evolved" with regard to μ and peak VCD. Thus, we hypothesised they exhibited the greatest genetic and functional diversities deriving from genetic drift and population dynamics both between and within physically isolated clonal populations. Additionally, to test the hypothesis that it was possible to obtain clonally-derived cell lines with intrinsic cellular capability to produce specific (DTE) recombinant proteins, the clones were tested for their ability to simultaneously express the difficult-to-express Sp35:Fc and an easy-to-express GFP. Both recombinant genes utilised CMV promoter encoded on different plasmid expression vectors, and transfected at the same DNA load. We reasoned that co-expression of two recombinant proteins with inherently distinct cellular requirements and using the same type of promoter would permit direct evaluation of both generic (e.g., cell transfectability) and protein specific (e.g., intracellular or extracellular secretion, cytosolic or ER folding/assembly) functional capabilities with minimal promoter interference (Huliak et al., 2012).

Transfections were performed using cuvette-based Nucleofection and cultured in suspension for a restricted 96 h period to minimise loss of recombinant plasmids (i.e., the effect of plasmid dilution on protein expression; Codamo et al., 2011). Cell viability of the clones measured 2 h post-transfection was 75.3–85.6% (median 82.9%). The transfection efficiencies determined by measurement of intracellular GFP fluorescence by flow cytometry (example plot shown in Figure 6.6) at 48 h post-transfection revealed that

transfection efficiency of the clones were highly comparable, i.e. between 96.5% and 99.6% (median 98.4%), with relatively homogenous transfected population compared to the parental cell line (Chapter 4, Figure 4.3). However, we observed that the mean and median fluorescence between clones varied by up to 5-fold (Figure 6.6B). Similar results were observed at 96 h post-transfection but with slightly lower clone-specific mean/median fluorescence intensity (data not shown) that can be ascribed to the effect of recombinant plasmid dilution/loss at high cell generation number post-transfection.

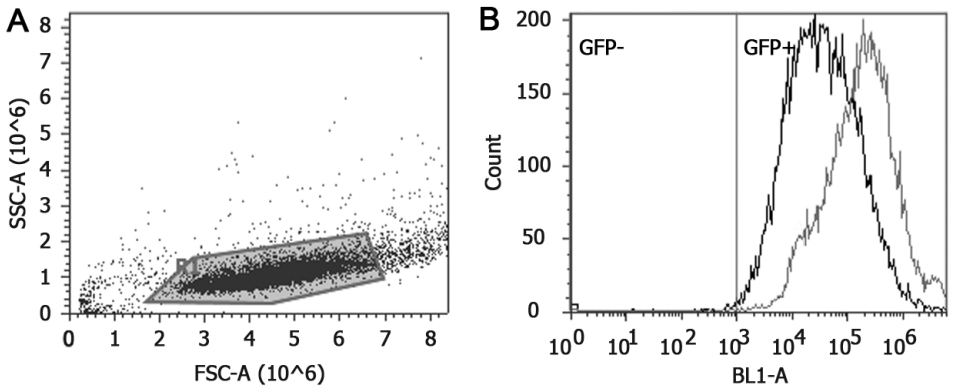

Figure 6.6: Example flow cytometry plots of GFP content of different clones. Cells were co-transfected with an equal amount of Sp35:Fc and pmaxGFP plasmids. GFP fluorescence was quantified 48 h post-transfection against negative control cells that were mock-transfected. Transfection efficiency obtained was 96.5–99.6% and mean/median fluorescence varied by up to 5-fold. **(A)** Cell population gated for granularity and size. **(B)** Example histogram of two clonal populations with low fluorescence (black) and high fluorescence (grey).

Figure 6.7 shows the transient Sp35:Fc production and IVCD for 16 selected clones and the parental population at 30, 60 and 120 generations. With respect to the IVCD, the majority of clones acquired significant increases in μ (effectively the "efficiency" of biomass synthesis) between 30 and 120 generations which was consistent with the long-term culture data (Figure 6.3A), indicating that it was possible to maintain the clones growth characteristic during cryopreservation procedure. Furthermore, it is interesting to highlight that all of maximally "evolved" clones (i.e., cells at 120 generations) except clone 8 exhibited an IVCD below that of the donor CHO-S parental population (Figure 6.7). This could be explained by the population dynamics within the highly heterogeneous donor population, in which fast proliferating clonal variants progressively became predominant resulting in a relatively high population-averaged μ (Davies et al., 2013).

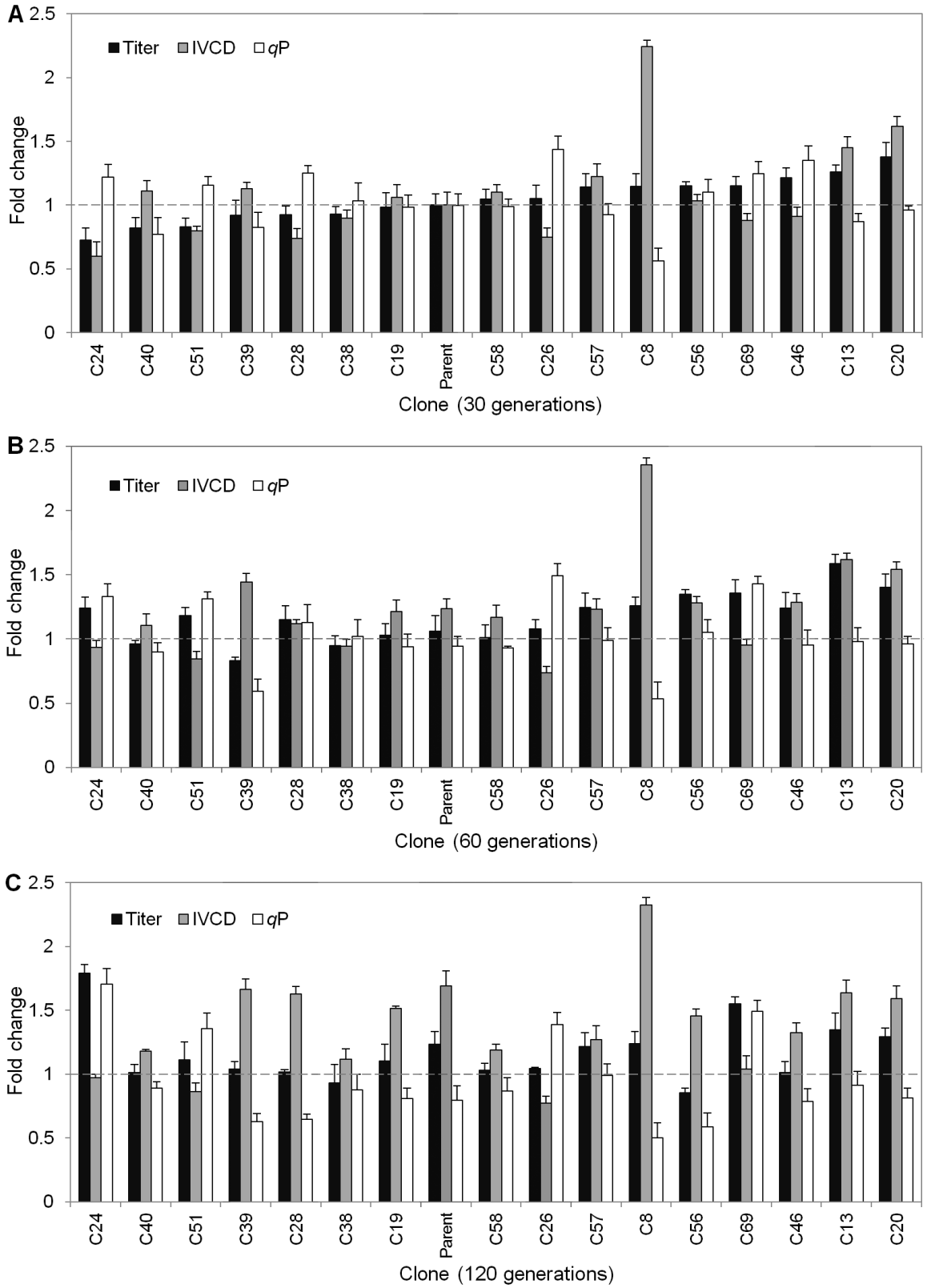

Figure 6.7: "Evolution" of CHO-S clones impact their transient recombinant Sp35:Fc protein production capability. Each clone at 30 generations **(A)**, 60 generations **(B)** and 120 generations **(C)** was transfected in duplicate and maintained in suspension culture for 96 h. Integral of viable cell density (IVCD) was calculated from the daily viable cell density, while secreted Sp35:Fc was measured 96 h post-transfection. Clones are ranked in order of

ascending titre at 30 generations. Clone-specific IVCD, Sp35:Fc titre and *q*P are shown relative to that of parental cells at 30 generations used to normalise assays. Error bars represent the standard deviation of two biological and two technical replicates.

Unlike the generally acquired increment in clone μ during the long-term culture, divergence in the clone-specific ability to manufacture recombinant Sp35:Fc was bi-directional (Figure 6.7). In this case, difference in *q*P changed on a clone-by-clone basis and therefore there was not a similar conserved trend in the change in clone Sp35:Fc process titre. On the extremes, C24, which was the lowest producer at 30 generations, evolved to become the highest producer after 120 generations, whereas C46 and C56 which were among the early promising candidates (at 30 and 60 generations) lost the production capability during long-term culture. Several high-producing clones exhibited heritably improved transient Sp35:Fc production, and this could be achieved via an increase in either IVCD (C13) or *q*P (C20 and C69).

We observed in Figure 6.7 that there was an inverse relationship across the clones where fast proliferating clones had relatively low Sp35:Fc *q*P (e.g. C8, C39) compared to slower growing cells (e.g. C26). Additionally, despite the 25% average increase in IVCD from 30 to 120 generations, the clones' average *q*P decreased by 9% and only one clone (C24) showed significant improvement. Analysis of these data show that there was a significant negative correlation between the *q*P and IVCD (PPMCC $r = –0.784$, $P < 0.0001$; Figure 6.8A). This inverse relationship was not (solely) due to the faster dilution of Sp35:Fc plasmid DNA in fast-proliferating cells as no correlation was observed with median (or mean) GFP fluorescence on any day of culture (best PPMCC $r = 0.126$, $P > 0.05$; Figure 6.8B). Indeed, our early transfection optimisation work using GFP on the parental cell line, as well as other work in this laboratory using different cell lines, demonstrated that GFP fluorescence intensity was always in proportion to the transgene copy numbers (data not shown). Therefore, we hypothesise that the negative correlation between μ and Sp35:Fc *q*P was rather derived from increased host cell competition for energy and synthetic resources for the production of either cell own protein biomass or the DTE protein.

We further analysed the relationship between the specific productivity and the protein biomass of each clone (untransfected clonal populations) at their respective generations. This analysis revealed that *q*P was directly correlated to the cell size/biomass content (PPMCC $r = 0.759$, $P < 0.0001$; Figure 6.8C) but this was not the case with GFP where there was no correlation between cell biomass and median (or mean) fluorescence (best PPMCC $r = 0.060$, $P > 0.05$; Figure 6.8D). Together, these data show that cell lines with the

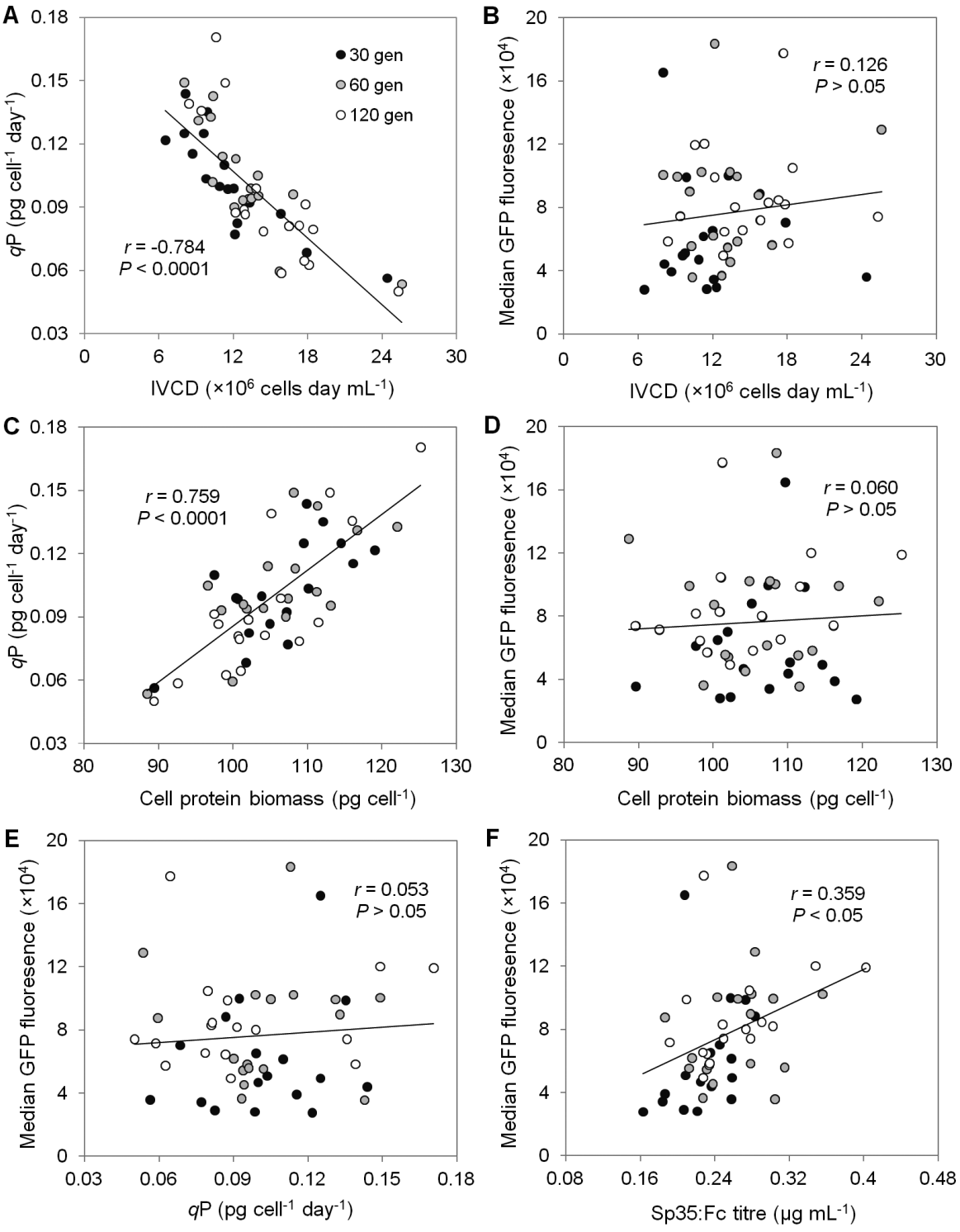

Figure 6.8: Difficult-to-express Sp35:Fc production rate correlates to cell-specific proliferation rate and biomass content but no relationship is observed in the case of GFP. (A) Fast-growing cells tend to have low Sp35:Fc specific productivity (*q*P), **(B)** whilst no such correlation is observed with GFP. **(C)** Cells that are good in manufacturing cell protein biomass (and thus have large size) are good in manufacturing Sp35:Fc, **(D)** but not necessarily GFP. **(E)** No apparent relationship is observed between GFP content and Sp35:Fc productivity, **(F)** although a weak correlation is observed with the volumetric titre. Data shown are for 96 h culture.

capability to accumulate high biomass while maintaining relatively high proliferation rate compared to the parental cell line were very likely to be high producers for DTE Sp35:Fc protein. In contrast, cells that were good in producing GFP were not necessarily good Sp35:Fc producers (Figure 6.8E), even though we did observe a weak correlation between median GFP fluorescence and Sp35:Fc volumetric titre (PPMCC r = 0.352, P < 0.05; Figure 6.8F). However, we acknowledge that the lack of correlation, although unlikely, could also be due to different ratio of GFP plasmid to Sp35:Fc plasmid uptake. We deduce that it is feasible to isolate variant parental clones that exhibit enhanced, heritable trait(s) to produce a particular DTE protein such as Sp35:Fc, or to harness the heterogeneity within the cell populations for this purpose.

6.3.4 Variation in clone-specific productivity does not significantly alter the aggregate amount

Non-reducing SDS-PAGE and immunoblotting for Sp35:Fc showed the presence of high molecular weight complexes (i.e., 400 kDa tetramer and 600 kDa hexamer aggregates) in all clonal cell lines (at different generations) tested as shown in Figure 6.9A. Under reducing condition, these high molecular weight complexes were reduced to 100 kDa monomer (data not shown) as previously observed with the parental cell line (Chapter 4, Figure 4.5), which indicated that the aggregates were disulphide bonded.

Despite the significant clone-to-clone variations in qP (Figure 6.7), the quantitative analysis showed that all clones at different generations produced comparable amount of aggregates (Figure 6.9B). Only two clones, C26 and C69, demonstrated appreciably (and consistently) lower aggregate levels compared to the CHO-S parental cell line at 30 generations (P < 0.1). Even though we were still able to establish a negative correlation between qP and aggregate amount (PPMCC r = -0.571, P < 0.0001; Figure 6.10), the large measurement noise/error produced from the immunoblotting technique might diminish the statistical confidence. An orthogonal method for aggregate quantification, preferably a more robust and sensitive one such as analytical ultracentrifugation or dynamic light scattering (Arakawa et al., 2007), is required to cross-check the Western blot data. However, it is worth noting that the two clones with noticeably lower aggregates (C26 and C69) had relatively high qP compared to other clones and the parental cell line (Figure 6.7). We contemplate that these two clonal cell lines had better ER processing facilities for the production of DTE proteins.

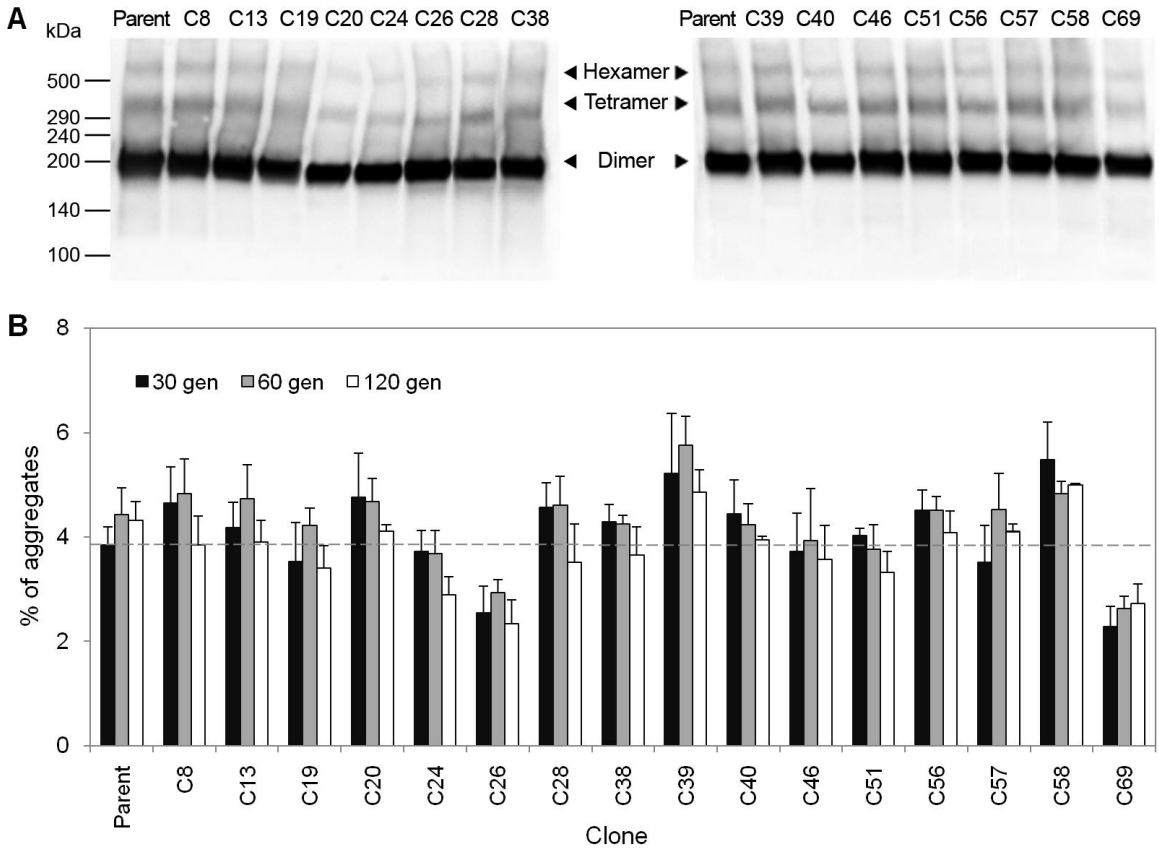

Figure 6.9: CHO-S clones and the parental cell line at different generations did not differ significantly in the amount of Sp35:Fc aggregates produced. Cell culture supernatants from the parental cell line and 16 clones at 30, 60 and 120 generations were harvested 96 h post-transfection and subjected to quantitative Western blotting. **(A)** Representative Western blot image showing Sp35:Fc native dimer (200 kDa) and aggregates (400 and 600 kDa) produced by the parental cell line and clones at 60 generations. Under reducing condition, all bands were reduced to 100 kDa (data now shown). **(B)** Relative quantification reveals that most clones produce comparable amount of aggregates relative to the parental cell line. C26 and C69 demonstrated appreciably and consistently lower aggregate levels compared to the parental cell line at 30 generations ($P < 0.1$). Error bars represent the standard deviation of two biological and two technical replicates.

6.4 Discussion

This work demonstrates that it is practically possible to harness the intrinsic variability of CHO cells in heterogeneous parental populations to acquire clonal derivatives that exhibit heritable variations in key functional attributes (e.g. μ, biomass accumulation) that favour DTE protein productions. The most basic observations were related to the adaptation response of clonally-derived host cell populations in their μ during routine sub-culture. Importantly, rapidly proliferating CHO-S clones with enhanced Sp35:Fc production rate compared to the donor parental cell line could also be obtained. The clones steadily

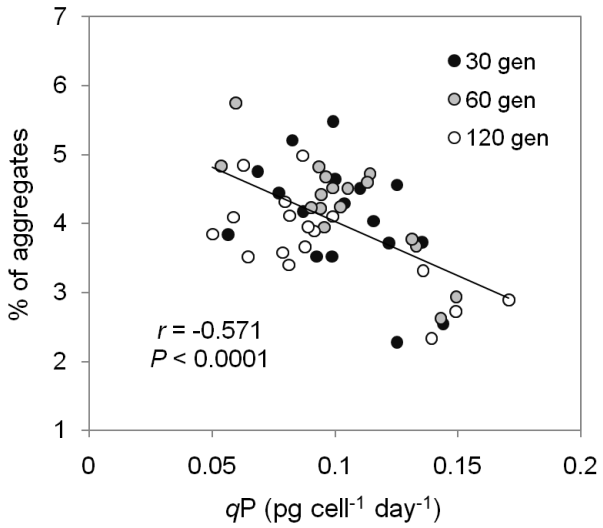

Figure 6.10: Clones that have a high capacity to manufacture difficult-to-express Sp35:Fc may produce less aggregates. There is a weak but highly significant negative correlation between clone specific Sp35:Fc productivity and the amount of aggregates produced across the clones at different generations.

increased their capability to efficiently acquire cell biomass (i.e., increase μ) towards their clone specific maximum as well as to achieve high peak cell density, although this was not the case for all clonally derived populations. We also observed a number of clones that did not respond to the imposed selective pressure, exhibiting a relatively constant, sub-maximal μ and peak VCD throughout the study (e.g., C26 and C51), and two clones displayed a similar maximal μ and peak VCD throughout the long-term culture, constantly in excess of the CHO-S parental average. Underpinning this averaged evolution progression were diverse clone-specific adaptive phenotypes derived from genome mutation and genetic drift (Davies et al., 2013).

The exact origins and molecular mechanisms underpinning genetic mutation or instability in mammalian cell lines are not yet fully understood. The most relevant and convincing evidence arises from the genetic instability models of cancer cell progression (O'Callaghan and James, 2008). Genetic instability is a characteristic of most cancers in which cancer cells display highly rearranged karyotypes characterised by translocations, deletions, inversions, abnormally banding chromosome regions, and extrachromosomal double minutes (Ma et al., 1993; Ruiz and Wahl, 1990). These, in turn, give rise to new (mutated) cells that tolerate some degree of malfunction in mitotic homeostasis and have survival advantages over their normal progenitors (Boland et al., 2009). Epigenetic changes may also be relevant in which they can collaborate with genetic changes to

perturb multiple key pathways in ways that promote carcinogenesis (Jones and Baylin, 2007). For example, DNA hypomethylation has been associated with chromosomal instability (Wilson et al., 2007) whereas methylation of CpG island promoters is known to cause silencing of key regulatory genes, which leads to an increased frequency of mutations (Jones and Baylin, 2007).

In the CHO-S donor parental cell population used in this study, such altered genotypes/phenotypes were apparently present at a frequency high enough to yield, after sorting 70 individual cells, at least 3 different stably improved subclones (C13, C20 and C69) with regard to μ and qP. Using these new clones, 1.4–1.6-fold increases in volumetric titre were achievable, whilst further increase (up to 1.8-fold) could be obtained using the "highly evolved" C24 clone at 120 generations. Even though we were unable to identify host cell lines with significantly lower Sp35:Fc aggregates, this was not unexpected considering the fact that biopharmaceutical companies normally screen several hundred to several thousand clones to obtain a desirable production cell line. The fact that some of our top producers at 30 generations (e.g. C46 and C56) had a temporarily stable production characteristic was also not surprising—we had observed this in other experiments in this laboratory which dealt with phenotypic/functional properties that vary over long term culture (Davies et al., 2013; Fernandez-Martell et al., unpublished data). However, our long-term culture data suggest that the rate of acquirement of genetic/functional diversity in some clonal populations was sufficient to respond to the imposed selective pressure to yield high Sp35:Fc producers. In view of this, we hypothesise that subsequent cloning of Clone 24 at 120 generations, which displayed the substantial improvement in both growth rate and productivity, will yield sub-clonal populations with significantly improved transient Sp35:Fc production.

It was observed that the CHO-S sub-populations exhibited substantial clone-specific difference in size and cell protein biomass that is directly correlated to μ. This data is in agreement with a mathematical analysis showing that mammalian cells exhibit variable control in mitotic timing mechanisms where cell proliferation rate is size-dependent throughout the cell cycle (Tzur et al., 2009). Additionally, difference in biomass accumulation rate obviously impacts mammalian cell-based manufacturing processes that utilise cellular protein biomass to create recombinant proteins. Recent studies have shown that biomass accumulation rate impacts the ability of clonal variants to perform specific *N*-glycan reactions (Davies et al., 2013), whilst deliberate engineering of cell biomass synthesis, either via engineering amino acid transport (Tabuchi et al., 2010) or cell signalling systems (Dreesen and Fussenegger, 2011), has been positively associated with

increased productivity. In the case of DTE Sp35:Fc, we found that it was possible to directly compare the host cell's own commitment to biomass synthesis to that of the product, i.e. clones with high unit biomass per cell and large size had more efficient regulation of Sp35:Fc production (and potentially with minimal aggregates). Although it may be argued that the variation in clone-specific proliferation rate may influence the Sp35:Fc productivity (due to variation in plasmid dilution rate in the cell), another study in this laboratory by Syddall et al. (unpublished data) also demonstrated that large/high biomass CHOK1SV cells were better producers compared to their smaller CHOK1V counterpart that had similar μ.

We speculate that the large CHO-S clones with high biomass have superior protein folding capacities in the ER, which in turn can be associated with large ER size and high mitochondrial mass (Bi et al., 2004). This leads to better processing and secretion of recombinant proteins that are relatively difficult to fold, resulting in higher productivity and better cell survival. Hu et al. (2013) showed that CHOK1 cells, which had relatively large ER size and mitochondrial mass, enabled stable cell line development for MAbs that are difficult to express in DUXB11-derived cells. Indeed, protein folding process consumes considerable energy (Dobson, 2003) which is supplied by the mitochondria, and therefore high mitochondrial mass may be advantageous for the production of DTE proteins (Bravo et al., 2011; Simmen et al., 2010). On the other hand, the significance of ER size can be recognised from the unfolded protein response (UPR), which drives expansion of the ER to boost its protein folding capacity and induce ER-associated degradation (ERAD) to degrade misfolded proteins (Bommiasamy et al., 2009; Sriburi et al., 2007). Overexpression of the active form of X-box binding protein 1 (XBP1s) in CHOK1 cell lines has been shown to expand the ER and the Golgi that lead to an increase in overall production capacity (Tigges and Fussenegger, 2006).

In cell line development, GFP fluorescence represents a useful surrogate correlated to levels of the protein of interest. Often, there is a high correlation between clone production of GFP and (easy-to-express) protein of interest including MAb (Davies et al., 2013; Kim et al., 2012) and growth factors (Freimark et al., 2010; Meng et al., 2000), thus enabling easy, rapid identifications of high producers. Yet, this could not be applied in the case of Sp35:Fc and possibly other difficult-to-express proteins. We hypothesise that this is due to the activation of the UPR during DTE protein expression that altered the cellular processes. Instead, we found that higher probability of success could be achieved by screening fast-growing cells (high μ) with large size or high biomass content. This data provides an important understanding for CHO host cell choice to express difficult to

produce recombinant proteins, even though advanced studies at cellular levels (transcriptomics, proteomics, etc; Charaniya et al., 2009; Dinnis et al., 2006) are required to reveal the underpinning factors. Increasingly complex recombinant protein products necessitate host cells with dedicated cell factories to achieve both high IVCD and *q*P. We anticipate that such cell lines would require an extensive "directed evolution" approach (Hernández Bort et al., 2010; Prentice et al., 2007) rather than mere exploitation of natural variation/mutation within the host cell populations.

Chapter 7

Synthetic Amplifier Circuit using ER Stress Element for Difficult-to-Express Protein Expression

This chapter describes a novel vector engineering approach specifically beneficial for difficult-to-express (DTE) protein production. The objective is to construct an expression vector that can circumvent the impact of overexpression of DTE protein on cell growth whilst achieving high productivity. By incorporating ER stress elements into a vector and manipulating endogenous and/or exogenous UPR transactivors, we show that it is feasible to create a synthetic amplifier circuit that could lead to improved DTE protein production.

Acknowledgements

The CMV-ATF6c vector (Appendix A) was constructed by Dr Stefan Schlatter. The ERSE-SV40 vectors driving Sp35:Fc, SEAP or ATF6c, as well as SV40 vectors driving Sp35:Fc or ATF6c (Appendix A) were constructed by Dr Nathan West. Recombinant SEAP assay for this chapter was performed by Dr Nathan West.

7.1 Introduction

Viable cell density (VCD) and cell specific productivity (*q*P) are two key factors in a biomanufacturing process using mammalian cells. As a molecular basis to the latter, efficient promoter/enhancer systems are a prerequisite for achieving high recombinant protein expression levels, especially for characteristically “easy-to-express” (ETE) proteins where transcriptional rates have been quantitatively shown to exert high levels of control over production (Ho et al., 2006; McLeod et al., 2011; O'Callaghan et al., 2010). Accordingly, expression is generally directed by strong constitutive promoter/enhancer combinations to maximise recombinant gene transcription levels.

Today, the human cytomegalovirus (CMV) promoter is utilised to drive expression of many biopharmaceutical products (Xia et al., 2006). It is a complex genetic element evolved by the virus to allow it to infect many different mammalian host cells including Chinese hamster ovary (CHO) cells (Stinski and Isomura, 2008). Yet, given its use for more than 25 years, the search for new expression promoting elements has become of major importance (Makrides, 1999). For example, powerful expression systems have been constructed from the promoter region of the Chinese hamster elongation factor-1α (EF-1α) gene and its 5' and 3' flanking sequences (Running Deer and Allison, 2004). Novel, strong promoters have also been synthetically derived (Brown et al., 2014; Schlabach et al., 2010) in which they offer a potentially attractive advantage of removing repeat elements from the vectors that could contribute to long-term expression instability via promoter methylation and gene deletion events (Kim et al., 2011a).

Apart from the strength of a promoter, the design of a vector also plays an important role to yield high expression levels. For instance, a bi-directional promoter system (Chatellard et al., 2007) has been reported to achieve similar expression levels for multi-chain proteins (antibodies, heterodimeric peptide hormones, etc) compared to conventional promoters without the use of gene amplification steps. Additionally, a single

open reading frame (ORF) expression vector driving both MAb heavy chain and light chain expressions can yield higher expression compared to dual ORF vector (Jostock et al., 2010), even though contrary reports (Davies et al., 2011) have also been published. On the other hand, for difficult-to-express (DTE) proteins that have mis-matched kinetics of transcription, translation and folding reactions, maximising transcription alone is unlikely to be beneficial. Therefore, efficient expression systems also require optimised vector designs for specific proteins and/or mammalian cell hosts.

Recent work in this laboratory by West (2014) has demonstrated that the transcriptional activity of the simian virus 40 (SV40) promoter can be increased when an ER stress response element (ERSE) is incorporated into the promoter's upstream region and is activated (Appendix D). ERSE, possessing a consensus sequence of CCAAT-N_9-CCACG, is present in the proximal promoter regions of many ER stress-responsive genes. This sequence is necessary and adequate for the induction of a number of ER chaperones including BiP, GRP94, and calreticulin during the unfolded protein response (UPR; Yoshida et al., 1998). Under normal conditions, only the general transcription factor NF-Y binds to the CCAAT motif of the ERSE. Under ER stress conditions, the cleaved ATF6 (ATF6c) and spliced XBP1 (XBP1s) bind to the CCACG motif of the ERSE, resulting in the transcriptional activation of the ER chaperones and foldases (Yamamoto et al., 2004; Yoshida et al., 2000). Even though the enhancement effect of ERSE on SV40 promoter was not observed when the ESRE was fused to a human CMV promoter (West, 2014), the relatively weak ERSE-SV40 promoters could be utilised to provide optimised protein-specific transcription activity kinetically coordinated with polypeptide-specific translation or folding rates.

Clearly, some aspects of an endogenous UPR (e.g. translation attenuation, ERAD and apoptosis) and its sustained activation are incompatible with high expression of DTE recombinant proteins. Stable transfectants capable of DTE protein production effectively avoid UPR induction via minimisation of ER load through reducing the rate by which recombinant polypeptides enter the ER including reduced rDNA transcription and/or mRNA translation (Davies et al., 2011; Mason et al., 2012). On the other hand, for any given transient expression of DTE protein, a balance of recombinant polypeptide synthesis activity (i.e., mRNA transcription and translation rates) and host cell UPR activation act in concert to define the sub-optimal productivity. Consequently, cell engineering solutions to improve DTE protein production typically act to minimise UPR induction (Pybus et al., 2014a), whereas conventional gene expression control via gene copy number or promoter

strength (Brown et al., 2014) does not allow discrete control/change over expression during the culture process.

We hypothesised that it was possible to control recombinant protein expression by incorporating ERSE into an expression vector and manipulating endogenous/exogenous UPR transactivation. To achieve this, we constructed several vectors containing the ER-stress responsive promoter (ERSE-SV40) driving Sp35:Fc, SEAP (as a model ETE protein) or the active form of the UPR transactivator ATF6. We then expressed these in CHO cells in various combinations using short-term transient platforms as well as in a fed-batch mode. We tested the general hypothesis that we can create a small synthetic "amplifier/dual activator" circuit specifically beneficial for DTE proteins, where expressed ATF6c both amplifies its own expression via transactivation of ERSE-SV40, and activates Sp35:Fc expression via the same mechanism, whilst generally transactivating cellular ER capacity via endogenous ERSEs (Figure 7.1).

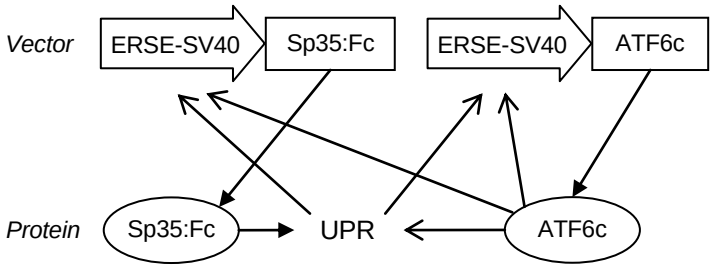

Figure 7.1: Schematic diagram of the synthetic "amplifier" circuit designed for a dynamic rDNA expression. Expression of difficult-to-express Sp35:Fc activates the UPR and cellular ER capacity via endogenous ERSEs. Expressed ATF6c both amplifies its own expression and Sp35:Fc expression via transactivation of ERSE-SV40 whilst the endogenous UPR proteins further amplify the recombinant gene expressions.

7.2 Materials and Methods

7.2.1 Plasmid DNAs

pSEAP2-Control plasmid (Clontech, Mountain View, CA) was used as the SV40 vector backbone for all ERSE-SV40 vectors as well as the SV40-Sp35:Fc and SV40-ATF6c vectors. The ERSE vectors were created by inserting oligonucleotides containing the ERSE sequence upstream of the SV40 promoters using a method described by Jobbagy et al. (2002) for the unidirectional (same orientation) insertion of repeated DNA sequences.

The CMV-ATF6c vector was created by inserting the ATF6c gene, which is based on the mouse (*Mus musculas*) coding DNA sequences, into the pcDNA™3.1(+) vector backbone (Life Technologies, Paisley, UK). GeneArt® pMA DNA vector (Life Technologies) containing Sp35:Fc gene but no promoter was used to equalise DNA load where necessary.

7.2.2 Transient fed-batch production

Transient transfection was conducted using Lipofectamine® LTX with PLUS™ reagent (Life Technologies) as described in Chapter 3, Section 3.3.2. The cell cultures were agitated at 140 rpm and incubated at 37°C under 5% CO_2. Nutrient supplementation started 6 h post-transfection using 10% v/v CHO CD EfficientFeed™ B (Life Technologies), followed by 10% v/v EfficientFeed B each on Days 2, 4, 6 and 8, giving a total of 50% v/v EfficientFeed B throughout culture.

7.3 Results

7.3.1 Amplification of recombinant protein expression via transactivation of the ERSE-SV40 promoter

A total of nine different expression vectors were constructed, harbouring either the CMV promoter, SV40 promoter or an ER-stress responsive promoter that drives either SEAP, Sp35:Fc or the active form of the UPR transactivator ATF6 (Figure 7.2). The ER-stress responsive promoter consisted of nine copies of ERSE units for optimal response, where these ERSEs were fused with an SV40 promoter in the vector construct, giving rise to the ERSE-SV40 promoter (henceforth called 'ERSE promoter'). Additionally, UPR transactivator ATF6c was utilised as it had been shown to give higher induction effect on ERSE than XBP1s (West, 2014; Appendix D). Both SEAP and Sp35:Fc were used to show whether the addition of ERSE had any significant impact on DTE protein expression. CHO-S cells were transfected with the same recombinant plasmid copy number (48,000 copies per cell) using the transient Nucleofection platform in various combinations as shown in Figure 7.3. Cloning plasmids containing no promoter were used to maintain the same total DNA load in all transfection experiments.

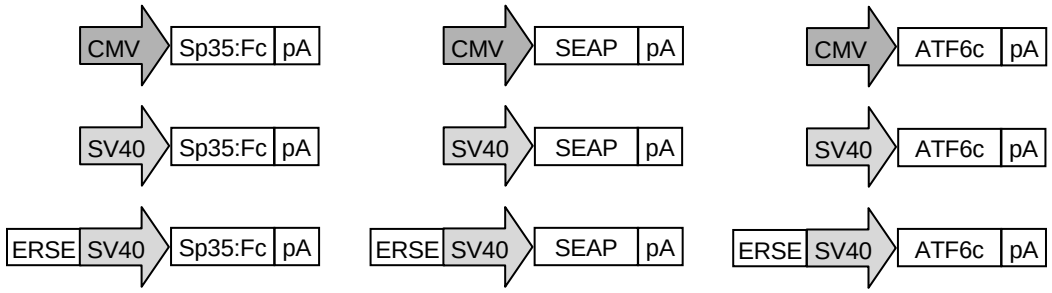

Figure 7.2: Schematic representation of nine different vectors containing CMV promoter, SV40 promoter or an ER-stress responsive promoter (ERSE-SV40) driving either Sp35:Fc, secreted alkaline phosphatase (SEAP) or the active form of the UPR transactivator ATF6 (ATF6c). The ERSE-SV40 promoter was constructed by fusing nine copies of ERSEs with an SV40 promoter.

Figure 7.3A shows the transient expression experiments of the ETE SEAP, normalised to the CMV control culture. On their own, the SV40 and ERSE promoters possess less than one-third the activity of a control construct harbouring the CMV promoter, in agreement with previously published literature that the CMV promoter is stronger than the SV40 promoter in many cell lines especially CHO (Foecking and Hofstetter, 1986; Liu et al., 1997; Qin et al., 2010; Zarrin et al., 1999). Comparison of the control for both SV40-SEAP and ERSE-SEAP showed that their levels of expression were identical, showing that the incorporation of ERSE alone has no effect on recombinant promoter expression from a vector using the SV40 promoter. Co-expression CMV-SEAP vector with CMV-ATF6c vector resulted in 1.8-fold increase in *q*P compared to the CMV control culture. When the CMV-ATF6c vector was replaced with an equal amount of ERSE-ATF6c vector, an increase of 1.4-fold was observed. Even though the effect was halved, this is evidence that recombinant ATF6c is capable of transactivating a vector containing ERSE upstream of a SV40 promoter (considering the fact that SV40 promoter has only one-third activity of that of the CMV). Increasing the ERSE-ATF6c copy number ratio would also further increase the *q*P. In all cases, the cell growth was hardly affected, leading to net 1.3–1.5-fold increases in volumetric titre.

When comparing the SEAP protein expression between the SV40 vector transfection (alone) and the ERSE-SEAP and CMV-ATF6c vectors co-transfection, there was a 3.1-fold difference in *q*P (Figure 7.3A). This highly significant boost in SEAP expression is a good reflection on the positive effect of the inclusion of ERSE sequences upstream of the SV40 promoter and the response to the ATF6c transactivator. However, the strength of SV40-ATF6c and ERSE-ATF6c vectors were too weak to produce noteworthy effects on ERSE-

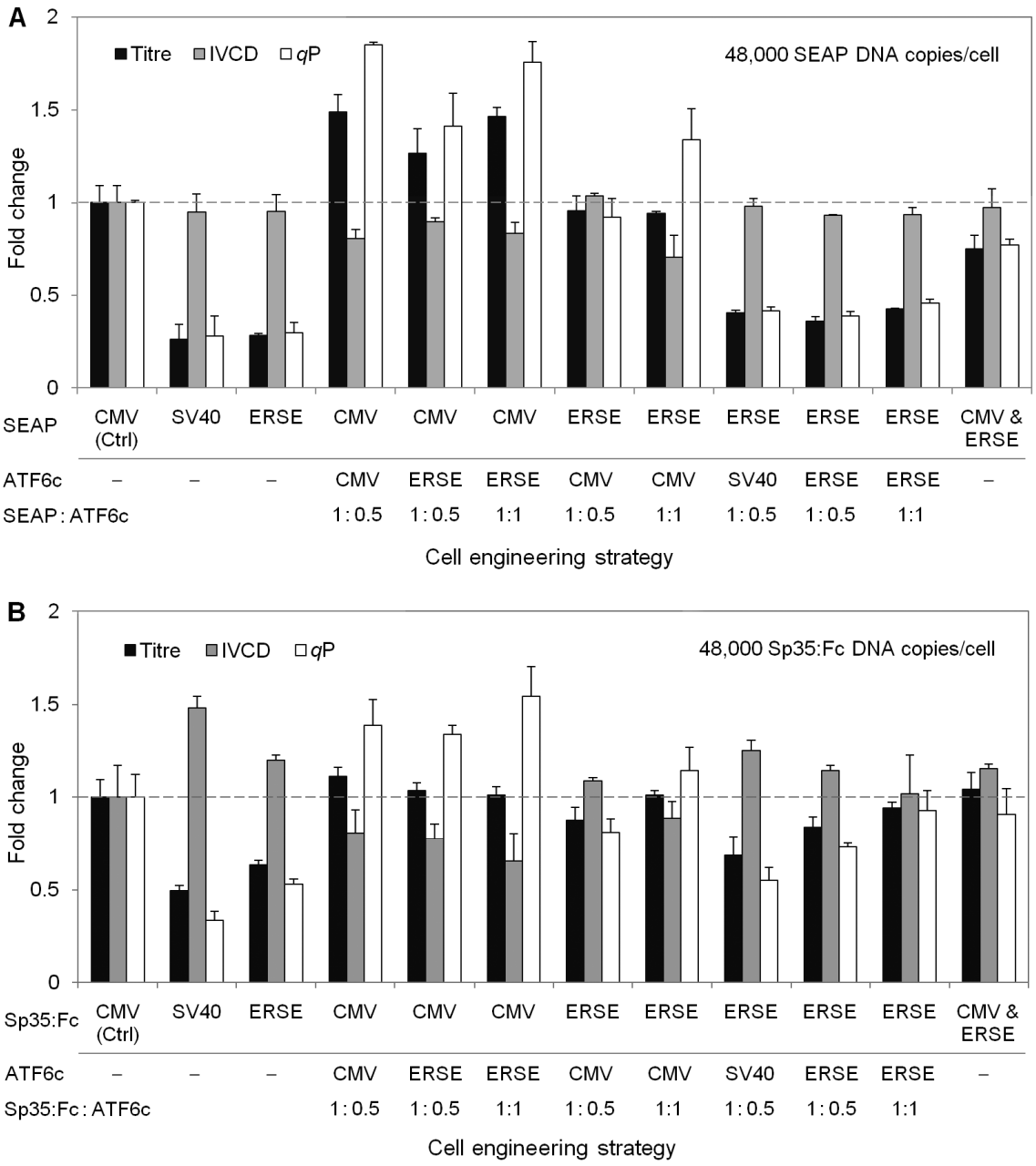

Figure 7.3: Effects of CMV, SV40 and ERSE-SV40 (ERSE) promoters on transient SEAP and Sp35:Fc productions, with and without ATF6c co-expression. The transfections were performed with an equal amount of 48,000 SEAP or Sp35:Fc DNA copies with a cloning vector used to equalise the total DNA load of 3.6 µg per 4.5×10^6 cells. Cell culture supernatant was harvested 96 h post-transfection and analysed by ELISA. **(A)** Transfections were carried out in various formats to systematically understand the effect of the ERSE system on easy-to-express SEAP expression. Each co-transfection was carried out at a specific SEAP to ATF6c rDNA copy number ratio as indicated under each transfection format. **(B)** Similar transfection formats were performed for difficult-to-express Sp35:Fc. All data are expressed as a relative (fold change) of the production exhibited by their respective CMV control promoter without any ATF6c co-expression. Error bars represent standard deviation of two biological and two technical replicates.

SEAP, even when the ERSE-ATF6c were doubled to 1:1 ratio. Indicatively, when a mixture of CMV-SEAP and ERSE-SEAP (1:1 ratio) was used, the SEAP expression was at about 75% of the CMV control culture which was roughly the average of individual CMV promoter and ERSE promoter expression levels.

Figure 7.3B shows the results for the DTE Sp35:Fc protein, in which very different μ (reflected in the IVCD) and *q*P profiles were observed. Firstly, even though the SV40 promoter maintained its relative activity (*q*P) compared to the CMV control, the low-level expression of DTE protein also reduced the metabolic burden and minimised the impact of the UPR on the cells, hence enabling the cells to grow better. However, the presence of low-level endogenous UPR transactivators could still be observed where expression using the ERSE-Sp35:Fc vector alone was able to increase the recombinant protein expression compared to using SV40 vector, although this was offset by a reduction in cell growth. Unlike SEAP, the effects of ATF6c co-expression were less profound on CMV-Sp35:Fc productivity, with improvements of only around 1.3–1.5-fold. We infer that this is related to the pre-existence of endogenous UPR transactivators in the cells that diminished the apparent effect of recombinant ATF6c. Moreover, the cell growth was significantly depressed, resulting in no benefits in terms of volumetric titre. Similarly, the ERSE-Sp35:Fc vectors were generally less responsive to CMV-ATF6c co-expression.

Interestingly, co-expression of either SV40-ATF6c or particularly ERSE-ATF6c with ERSE-Sp35:Fc resulted in relatively high activity of Sp35:Fc productivity (Figure 7.3B), especially when compared to SEAP production using the same strategies (Figure 7.3A). This suggests that a positive feedback loop existed between the ATF6c/endogenous UPR transactivators and ERSE-SV40 promoter activation, and that the system is particularly beneficial for DTE protein production. Furthermore, the mixture of CMV-Sp35:Fc and ERSE-Sp35:Fc (1:1 ratio) was able, to some extent, to induce the ERSE-Sp35:Fc vector that resulted in higher expression activity compared to SEAP. This again demonstrates that a vector containing ERSE upstream of an SV40 promoter is capable of responding to endogenous UPR transactivators. In the hindsight, as Sp35:Fc production was generally less responsive to ATF6c and/or resulted in considerable growth depression, no significant improvements in volumetric titre was observed in all cases. On the other hand, as we did not measure the levels of ATF6c, further work is required to validate that the lack of responses to ATF6c were not due to the inability of the cells to overexpress the UPR transactivator. Similarly, additional work is required to validate the synthetic circuit by measuring the levels of recombinant mRNA and UPR induction in each case.

7.3.2 The activity of the ERSE-SV40 promoter is dependent on UPR response/transactivator levels

In order to further understand the relationship between the ERSE system and endogenous UPR induction during DTE protein overexpression, we transfected the cells with 50% higher Sp35:Fc DNA copy number (72,000 copies per cell) to induce a higher UPR response as previously shown in Chapter 4 (Figure 4.7; equivalent to an increase from 2 to 3 µg of Sp35:Fc DNA). We anticipated that the ERSE promoters would become more active at very high Sp35:Fc overexpression that in turn could lead to an increase in volumetric product titre. These results are shown in Figure 7.4.

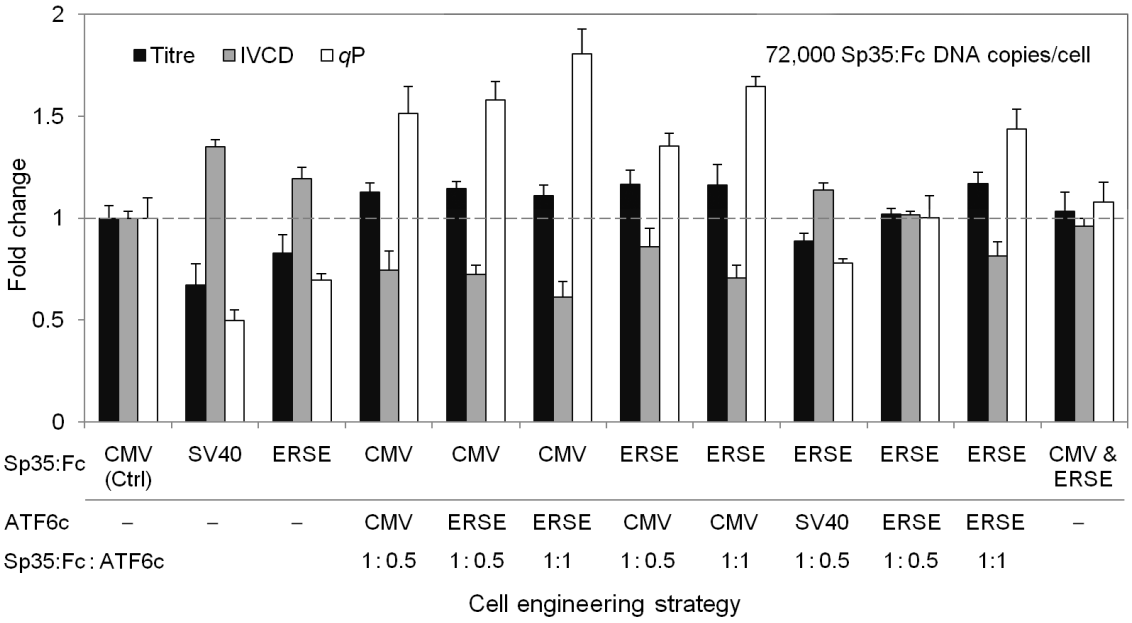

Figure 7.4: Cells were transfected with higher rDNA load (72,000 DNA copies per cell) to invoke more UPR response. A cloning vector used to equalise the total DNA load of 4.5 µg per 4.5×10^6 cells. Cell culture supernatant was harvested 96 h post-transfection and analysed by ELISA. Data are expressed as a relative (fold change) of the production exhibited by CMV control promoter without any ATF6c overexpression. Error bars represent standard deviation of two biological and two technical replicates.

At this high Sp35:Fc DNA copy number (72,000 copies per cell; Figure 7.4), the relative activity (*q*P) of SV40 promoter was approximately half of the CMV control culture, which is higher than the one-third activity achieved when 48,000 copies were used (Figure 7.3B). Together, these data reflect the saturation phenomenon of DTE protein expression using CMV as observed in Chapter 4 (Figure 4.4A). Significant activity of the ERSE

promoter can also be observed when the ERSE-Sp35:Fc vector was used on its own—a positive consequence of the high endogenous UPR transactivator activity (Figure 7.4). Expectedly, the *q*P of the ATF6c-overexpressing cells was considerably higher (1.5–1.9-fold increase compared to the CMV control culture; Figure 7.4) as opposed to the 1.3–1.6-fold increases obtained with 48,000 copies shown earlier (Figure 7.3B). Additionally, the ERSE-ATF6c promoter became slightly more active than the CMV-ATF6c (Figure 7.4), hence granting more benefit compared to cells with lower DNA copy number (Figure 7.3B). Importantly, we observed that the co-expression of ERSE-SP35:Fc with CMV-ATF6c or ERSE-ATF6c resulted in significant enhancement in the ERSE promoter activities and *q*P levels (Figure 7.4) compared to when 48,000 Sp35:Fc copies were used (Figure 7.3B). The combination of CMV-Sp35:Fc and ERSE-Sp35:Fc (1:1 ratio) also appeared to yield similar results to the CMV control. Yet again, in all cases, cell growth was substantially oppressed at high Sp35:Fc expression, and this negative relationship resulted in little or no improvement in product volumetric titre compared to the normal CMV culture control. We deduced that the full potential of the synthetic amplifier circuit could only be realised during fed-batch culture.

7.3.3 The dynamic feature of ERSE-SV40 promoter increases DTE fusion protein production in transient fed-batch mode

We hypothesised that transition of CHO cells through an extended Sp35:Fc production process, with the associated dynamic variation in cell physiology and function (e.g., growth rate, UPR induction) may further amplify the relative proportion of hetero/endogenous factors affecting ERSE promoter activity. However, an initial experiment using Nucleofection system showed complete loss of productivity after 7–8 days post-transfection, and therefore transient transfection was performed using lipoplex-mediated transfection where cell culture and productivity could be maintained for about 8–9 days in fed-batch mode. The lipid-based method was preferred over polymer-based (e.g. polyethylenimine) to ensure production variability was directly linked to differences in promoter activity rather than mediator-specific artefacts, where lipoplex-delivered plasmids have been shown to be more efficiently expressed than polyplex-delivered plasmids on the basis of protein expression per plasmid number in the nucleus (Cohen et al., 2009). Transfection efficiency was determined by measurement of intracellular GFP by flow cytometry which showed that transfection efficiency using this method was >96% at 96 h post-transfection (Figure 7.5).

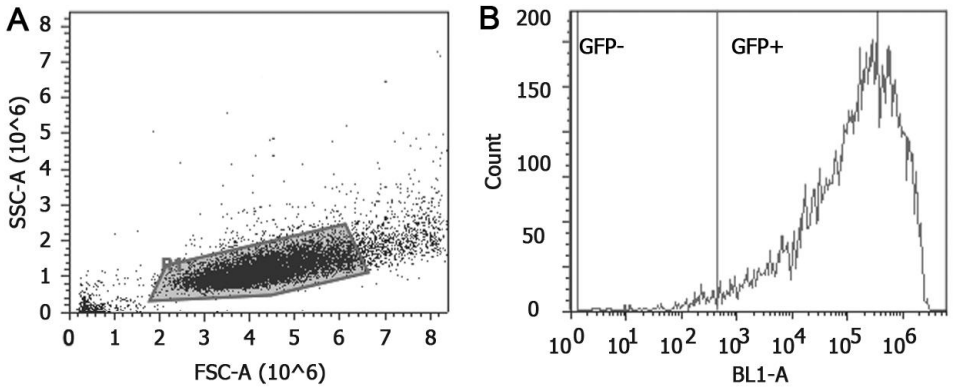

Figure 7.5: Example flow cytometry plots of intracellular GFP content of different clones. 25×10^6 cells were co-transfected with 16.67 µg pmaxGFP plasmid and GFP fluorescence was quantified 96 h post-transfection against negative control cells that were mock-transfected with empty vector. **(A)** Cell population gated for granularity and size. **(B)** Example histogram of cells with fluorescent label with transfection efficiency >96%.

In order to comprehend the dynamic ERSE promoter activity during fed-batch culture, we studied the co-expression of different Sp35:Fc and ATF6c vectors at a 1:0.5 ratio, namely the combinations of (i) CMV-Sp35:Fc promoter and CMV-ATF6c promoter (CMV/CMV), (ii) CMV-Sp35:Fc promoter and ERSE-ATF6c promoter (CMV/ERSE), and (iii) ERSE-Sp35:Fc promoter and ERSE-ATF6c promoter (ERSE/ERSE). The transfected cells were cultured alongside a CMV control culture with no ATF6c co-expression (CMV/–). We observed that the differences in promoter activity observed in the 4 days culture (Figure 7.4) were further amplified in the fed-batch transient system.

Figure 7.6A shows the VCD profile of the different vector engineering strategies. Most notably, the CMV/CMV and CMV/ERSE strategies had lower cell densities in the first 4–5 days due to the effect of Sp35:Fc and ATF6c overexpressions on cell growth. Additionally, the latter strategy appeared to further reduce the cell growth on Day 4 post-transfection, as well as resulted in the lowest peak VCD. We infer this was the effect of the dynamic loop where expressed ATF6c amplifies its own expression over time via transinduction of ERSE-SV40. Despite the negative effect on cell growth, the CMV/ERSE strategy yielded the highest Sp35:Fc volumetric titre, which was 53% higher than that was obtained by the CMV control culture. This was followed by the ERSE/ERSE strategy (38% increase), and the lowest titre (23% increase) was obtained with the conventional CMV/CMV promoters. It is also interesting to point out that the ERSE/ERSE strategy had a relatively low titre level at the beginning of the culture, which was consistent with the 4 days culture data (Figures 7.3B and 7.4), but eventually surpassed the control and

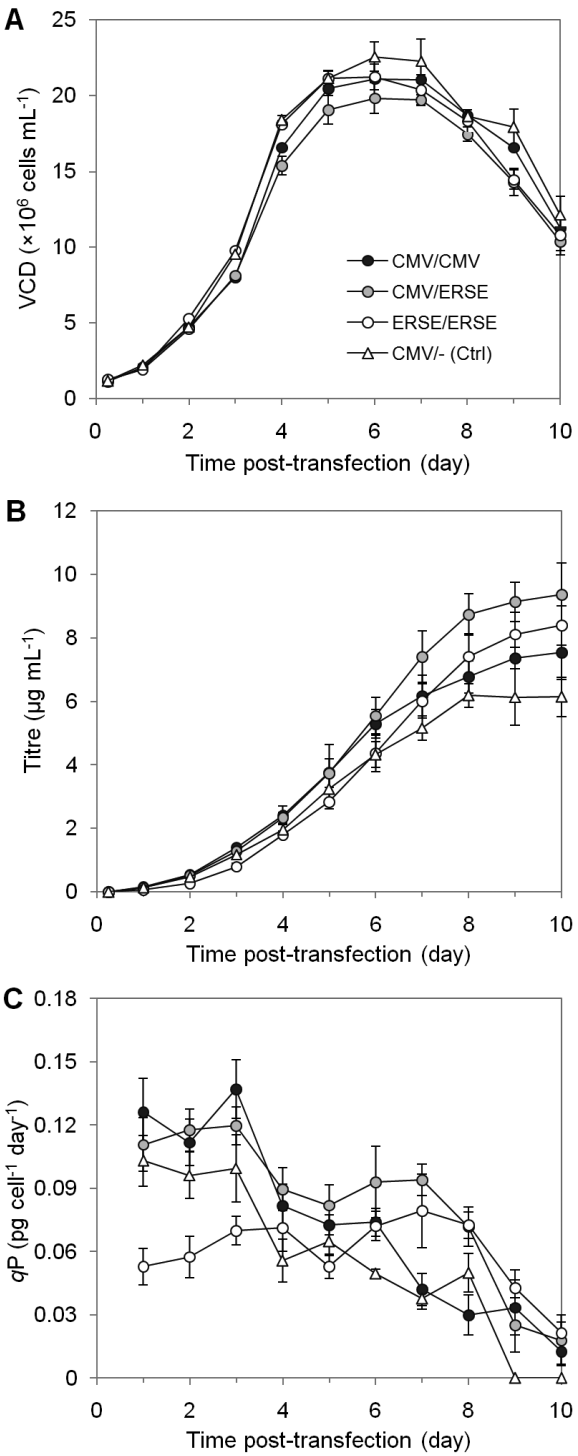

Figure 7.6: ERSE-SV40 (ERSE) promoters exhibit dynamic profile in Sp35:Fc fed-batch transient production process. CHO-S cells were co-transfected with different Sp35:Fc and ATF6c vector combinations (1:0.5 ratio), namely CMV-Sp35:Fc and CMV-ATF6c promoters (CMV/CMV, ●); CMV-Sp35:Fc and ERSE-ATF6c promoters (CMV/ERSE, ◍); and ERSE-Sp35:Fc and ERSE-ATF6c promoters (ERSE/ERSE, ○), and cultured alongside a CMV control culture with no ATF6c co-expression (CMV/–, △). The total DNA load was equalised using a cloning plasmid with no promoter. **(A)** Viable cell density (VCD) profile post-transfection. **(B)** Sp35:Fc titre profile measured by ELISA method. **(C)** Daily cell specific productivity (*q*P) profile based on the VCD and Sp35:Fc titre. The data represent two biological replicates.

CMV/CMV levels. This again demonstrated the dynamic, beneficial feature of ERSE promoters for DTE protein production over long-term culture.

To understand the variations in the different CMV and ERSE promoter combinations, we further analysed the Sp35:Fc expressions in term of daily *q*P (Figure 7.6C). The CMV/CMV strategy showed a continuous decline in *q*P over time post-transfection that can be attributed to the dilution of plasmid DNA into daughter cells after successive rounds of cell division (Codamo et al., 2011). This normal cell engineering strategy, lacking the dynamic feature of ERSE promoter and combined with slow initial growth rate, resulted in only a 23% increase in volumetric product titre (Figure 7.6B) despite the very high initial *q*P. The most optimal strategy appeared to be the CMV/ERSE combination as demonstrated by the consistently high *q*P, yielding the highest improvement in volumetric product titre. Interestingly, the combination of ERSE/ERSE exhibited the most dynamic *q*P profile over the fed-batch culture, in which the *q*P level increased by about 1.5-fold from Day 1 to Day 7. Even though the peak *q*P achieved with this strategy was still considerably lower than the peak *q*P of CMV/CMV and CMV/ERSE strategies, a moderate increase (37%) in volumetric titre was still attainable. It is also worth mentioning, even though the data is not included here, that the combination of ERSE-Sp35:Fc with CMV-ATF6c (ERSE/CMV) in fed-batch culture was not particularly beneficial. The relatively weak ERSE-driven Sp35:Fc expression, combined with the progressive loss of CMV-driven ATF6c expression over time, resulted in an insignificant 11% increase ($P > 0.05$) compared to the CMV control culture (data not shown).

7.4 Discussion

Tunable promoters for controlled protein expression is not a new concept in synthetic and systems biology. Regulated transcription/translation can be achieved through the redesign of endogenous promoters (Alper et al., 2005; Tornøe et al., 2002), the creation of fully synthetic promoters (Brown et al., 2014; Yim et al., 2013) and the insertion of regulatory sequences (Bulter et al., 2004; Stapleton et al., 2012). To date, discrete control of protein expression in eukaryotic systems has been achieved in the form of promoter activation and repression for programmable expression (Blount et al., 2012; Farzadfard et al., 2013; Hurley et al., 2012) as well as an oscillating gene expression system for periodic induction of specific target genes (Tigges et al., 2009). With the further progression of biological research and development, including in the biopharmaceutical industry, the need for controlled expression of recombinant proteins is ever increasing.

This study describes a novel tunable vector system that utilises a UPR transcription factor and its DNA binding element. UPR, activated in the event of ER stress, regulates expression of genes encoding the resident ER molecular chaperones and folding enzymes (Chakrabarti et al., 2011). In view of that, UPR activating elements ATF6 and XBP1 transcription factors have been extensively used in cell engineering to increase the expression of therapeutic recombinant proteins in mammalian cells, albeit with varying degrees of success (e.g., Campos-da-Paz et al., 2008; Ku et al., 2008; Pybus et al., 2014a; Tigges and Fussenegger, 2006). Recently, their DNA binding elements, ERSE and UPRE, have also been introduced into DNA vectors as a non-invasive system for the detection and quantification of endogenous and induced UPR factors in CHO cells (Du et al., 2013). However, these two UPR components (i.e., the transcription factors and their binding elements) have not been utilised together as part of a cell/synthetic biology engineering strategy. We therefore, for the first time, simultaneously used the transcription factor ATF6c and ERSE incorporated SV40 promoter, and tested the synergistic effect of various co-expression formats on easy-to-express (SEAP) and difficult-to-express (Sp35:Fc) protein productions.

For the ETE recombinant protein, we observed that the *q*P and volumetric titre could be increased by co-expressing UPR transactivator ATF6c, and the use of weak SV40 and ERSE-SV40 promoters were not desirable compared to CMV. Under these circumstances ERSE-SV40 promoter offers fewer advantages for ETE proteins as the expression of these proteins is limited primarily by transcription/translation (Chapter 4, Figure 4.4) and its overexpression does not easily invoke a UPR (Chapter 4, Figure 4.7). In contrast, for DTE production, the (originally) weak ERSE-SV40 promoter effectively reduces UPR induction via minimisation of ER “input rates”, which in turn grants the much needed rapid cell growth at the beginning of culture. Additionally, the relatively high level of endogenous UPR transactivators associated with DTE production forms an active, continuous loop with ERSE-SV40 promoter and amplifies its activity—yielding a dynamic (increasing) recombinant gene expression over time. From a mechanistic perspective we conclude that the ERSE-SV40 promoter enables a shift from fast-growing cells to high-producing cells during the transient production of DTE recombinant proteins using extended culture. Further work will validate this mechanism by measuring the levels of recombinant mRNA and UPR induction at different time points of culture.

The dilution of recombinant plasmid DNA after repeated rounds of cell division post-transfection is a major factor limiting yields with transient systems, where the expression level of the transgene per cell inevitably decreases with every generation. Even though it is

possible to (slightly) extend the transgene expression by using non-physical transfection methods such as lipofection and polyfection, this is likely due the different kinetics of gene delivery into host cells (Mehier-Humbert and Guy, 2005), as previous studies showed that the internalisation of cationic lipid/polymer–DNA complexes into cell nucleus is relatively slow (Carpentier et al., 2007; Zabner et al., 1995). To circumvent this issue, a range of episomal expression vectors has been developed that enable plasmid DNA amplification and/or maintenance extrachromosomally using viral genetic elements such as the Epstein-Barr virus nuclear antigen-1 (EBNA-1; Kunaparaju et al., 2005). Even so, the *q*P in many cell lines employing these systems appears to peak between 1 and 3 days post-transfection (Carpentier et al., 2007; Wulhfard et al., 2012) thus still leaving some room for improvement. Our data show that whilst the CMV promoter suffered from progressive loss of productivity, the ERSE-SV40 promoter maintained or even increased the *q*P during the exponential and stationary phases. Therefore we propose the ERSE-SV40 vector as an alternative or complement to episomal-based expression systems.

One obvious drawback of the ERSE system is that it requires high expression level of DTE protein (to induce endogenous UPR) or co-transfection/expression of UPR transactivators, where the former can be difficult to achieve with the relatively weak ERSE-SV40 promoter while the latter is not always desirable (e.g., limiting rDNA load of protein of interest, increased transfection toxicity). Additionally, the current system might not be compatible with stable expression systems where studies have shown that stable cell lines effectively avoid UPR as there were no indication of UPR-transactivator upregulations even in high producing cells (Dinnis et al., 2006; Smales et al., 2004), whilst the generation of stable UPR-transactivator expressing cell lines can be very difficult (Becker et al., 2008). Nevertheless, where co-expression of UPR transactivator is undesirable or intractable, it is possible to attain an adequate expression of a given DTE protein by combining the strong CMV promoter and the dynamic ERSE promoter at specific ratios. Our data on Sp35:Fc demonstrated that the CMV-driven Sp35:Fc expression was sufficient to transactivate the ERSE promoter to achieve at least the same *q*P to the CMV only control during short-term culture. Moreover, for ETE proteins or stable cell lines with low/no endogenous UPR, we surmise that the system limitation can be possibly overcome by invoking the UPR using chemicals such as dithiothreitol (DTT) and tunicamycin, where they can be easily titrated and added into culture medium at the desired time point for optimal results.

We anticipate that the combination of ERSE-Sp35:Fc and ERSE-ATF6c vectors in this study would yield lower, or at least similar, amount of Sp35:Fc aggregates due the relatively low ER stress imposed (i.e., low *q*P) on the host cells throughout fed-batch

culture (Figure 7.6C). Accordingly, even though the combination of CMV-Sp35:Fc and ERSE-ATF6c resulted in the highest improvement in volumetric titre (Figure 7.6B), the repercussion of improved *q*P on aggregate formation needs to be determined as the system might not resolve the cellular bottlenecks associated with Sp35:Fc production (Chapters 4 and 5). Further study would address the impacts of the dynamic nature of ERSE systems on product qualities such as protein aggregates and glycosylation profile. Additionally, it would be interesting to know if XBP1s may yield better results when used on the systems (as a replacement or complement to ATF6c) where this UPR transactivator has been widely demonstrated to improve the secretory capacity of CHO cells (Codamo et al., 2011; Ku et al., 2008; Tigges and Fussenegger, 2006).

The data presented in this study is only the first steps towards developing a mammalian controlled/dynamic expression system for difficult-to-express recombinant proteins and there are various potential applications including controlling the expression of multiple genes for heterologeous proteins as well as increasing ETE protein productions via artificial UPR induction. For example, for DTE monoclonal antibodies (MAbs) where the productivity is affected by the proportion of light to heavy chain polypeptides (Pybus et al., 2014a), the optimal production condition may be self-regulated using an ERSE-SV40 promoter transcribing the light chain polypeptide—under ER stress, the UPR will induce the light chain synthesis to improve the MAb assembly rate, and as ER homeostasis is restored the light chain synthesis is reduced to the preceding level. For other DTE recombinant proteins, the optimal engineering strategy is likely to be unique. Further refinement and expansion of this system will make it a very valuable tool not only for the production of recombinant proteins but many areas of biotechnology.

Chapter 8

Cell and Process Engineering for Improved Production of a Difficult-to-Express Fusion Protein

This chapter investigates the effects of cell and process engineering approaches on difficult-to-express protein production. The aims are to find out which mitigation strategies can be utilised to alleviate the Sp35:Fc protein production bottlenecks, as well as the combined effects of specific strategies. We show that the integrated cell and process engineering strategies were able to improve the Sp35:Fc production by several fold due to restoration of normal cellular function.

Acknowledgements

Plasmid vectors encoding an activated form of UPR transactivator (cleaved ATF6 or spliced XBP1) were constructed by Dr Stefan Schlatter (Appendix A).

8.1 Introduction

Mammalian cells, including Chinese hamster ovary (CHO) cell lines, can be poorly efficient for high production levels of biologics due to limiting translational and/or post-translational mechanisms. The bottlenecks are exacerbated in the case of difficult-to-express (DTE) recombinant protein production using transient gene expression (TGE) systems as the host cells are easily overloaded with transgenes and nascent recombinant polypeptides. Consequently, for many recombinant products, productivity is unpredictably low; even monoclonal antibody (MAb) protein products in the same isotype/sub-class can display variable expression levels due to different translational and post-translational process rates (Pybus et al., 2014a). This can also be expected from artificial fusion proteins where they have not coevolved in which the two (or more) combined components are thought to have different folding and/or secretion requirements (Lee et al., 2007).

Stable transfectants capable of efficient DTE recombinant protein expression avoid unfolded protein response (UPR) induction through reducing the rate of recombinant gene transcription and/or mRNA translation resulting in low stable expression (Mason et al., 2012). Therefore, engineering strategies for improved TGE processes are desirable not only to increase the rapid preclinical supply of DTE recombinant proteins but also can provide strong preliminary data on stable “manufacturability”. Very low transient production titres indicate that considerably more cell development and/or process development may be needed to generate a stable production cell line with desirable attributes. Whilst the bioengineering of host cells may reduce the effort required to select clonal lines with high productivities (Cain et al., 2013), the transient production data itself can be used to inform on how to engineer the host cell line or process platform for enhanced stable performance (Nishimiya et al., 2013).

Various studies have shown that endoplasmic reticulum (ER) chaperones, UPR components, and stress-mediated apoptosis pathway components could be co-expressed to improve the cell survival or to solve bottlenecks and cellular limitations caused by recombinant protein overflow (Dietmair et al., 2011; Schröder, 2008). With respect to the

latter, co-expression of functional proteins involved in protein secretion including binding immunoglobulin protein (BiP), protein disulphide isomerase (PDI) and signal recognition particle 14 (SRP14) had successfully improved the expression of DTE MAbs (Le Fourn et al., 2014; Pybus et al., 2014a). Other attempts included the ectopic expression of the activating transcription factor 6 (ATF6) and X-box binding protein 1 (XBP1), two transcription factors that regulates ER maintenance and expansion, to decrease ER stress and increase protein processing and secretion in CHO cells (Pybus et al., 2014a; Tigges and Fussenegger, 2006).

Furthermore, treatment of mammalian cell cultures with the so-called chemical chaperones to increase expression/secretion of recombinant proteins is a method with a long-standing history. Sodium butyrate, a histone deacetylation (HDAC) inhibitor and arguably the most frequently used chemical chaperone, is known to increase gene transcription by enhancing gene accessibility to transcription factors (Jiang and Sharfstein, 2008). This leads to the upregulation of many genes involved in protein processing, secretion and redox activity (Yee et al., 2008). Interestingly, different chemicals appear to act via distinct mechanisms and therefore can confer different/multiple benefits. There are various reports on the applications of chemical chaperones in reducing protein aggregation and suppressing ER stress (Chapter 2, Table 2.1) that can be particularly beneficial for DTE proteins. With regard to the former, chemical chaperones have the ability to help the correct folding of aggregation-prone proteins by promoting the expression of molecular chaperones, or to maintain their native state by interacting directly with them (Papp and Csermely, 2006; Vagenende et al., 2009).

Despite the extensive reports on cell engineering and chemical chaperone strategies in mammalian cell-based production process, the combinatorial effects of different chemical and/or molecular chaperones have not been adequately explored. This is particularly relevant in mammalian cells which have very complex cellular regulations and may suffer from multiple production bottlenecks. Recombinant protein overexpression has been reported to induce a UPR that can affect global mRNA translation rates that lead to reduced cell growth and biomass content (Pybus et al., 2014a). On the other hand, the beneficial effects of many chemical chaperones on recombinant protein production are often compromised by their cytotoxic effect on cell growth via the activation of cellular apoptosis (Rodriguez et al., 2005). Therefore it is necessary to identify the most effective external manipulations that can overcome/minimise deleterious cellular regulation in optimising both the integral of viable cell density (IVCD) and cell specific production rate (qP) to achieve maximum volumetric product titre.

We have previously shown in Chapter 4 that the Sp35:Fc expression system became saturated above a specific threshold level as cells were not capable of processing more than a certain amount of nascent polypeptides. The overexpression, in turn, resulted in considerable protein aggregation and triggered the UPR to restore cellular homeostasis. Therefore, to overcome the cellular bottlenecks we compare the benefits of (i) co-expressing a variety of molecular chaperones, foldases and UPR transactivators, and (ii) using chemical chaperones/inhibitor and mild hypothermic condition, to improve either the rate/capacity of folding reactions or cell growth as well as to suppress aggregate formation. Various combinations of the engineering strategies were also tested, while the design of experiment (DOE) methodology was employed to study and optimise the interactions between different effectors. We further show that the combined cell and process engineering strategies were able to improve the TGE system by several fold due to restoration of normal cellular function and was particularly beneficial for DTE recombinant protein production.

8.2 Materials and Methods

8.2.1 Expression vectors and chemicals

Molecular chaperone and foldase genes, specifically human BiP, human PDI, human ERO1Lβ and human Cyclophilin B (CypB), were driven by CMV promoters encoded on plasmids from OriGene (Rockville, MD). UPR transactivators (active forms of ATF6 and XBP1) inserted into a pcDNA™3.1(+) (Life Technologies) vector backbone were driven by CMV promoters. Chemical chaperones sodium 4-phenylbutyrate (PBA), betaine, dimethyl sulfoxide (DMSO), glycerol and trimethylamine N-oxide (TMAO) were purchased from Sigma-Aldrich (Poole, UK) with deionised Milli-Q water used as a solvent and diluent. PERK inhibitor was purchased from Merck Chemicals (Nottingham, UK) with DMSO used as a solvent and diluent. Chemicals were added and cells were shifted to 32°C at 3 h post-transfection.

8.2.2 Transient fed-batch production

Transient transfection was conducted using Lipofectamine® LTX with PLUS™ reagent (Life Technologies) as described in Section 3.3.2. The cell cultures were agitated at 140 rpm and incubated at 37°C under 5% CO_2. Nutrient supplementation started 6 h post-

transfection using CHO EfficientFeed™ Kit (Life Technologies) with a total of up to 60% v/v CHO CD EfficientFeed™ A and B throughout culture as follows; 10% v/v each on Days 0, 3, 5, 7, 9 and if necessary, on Day 11.

8.2.3 Design-of-experiment

DOE software Design-Expert® 8.0 (Stat-Ease Inc, Minneapolis, MN) was used for the analysis of responses to chemical/molecular chaperones. The programme provided a mixture matrix, a fitted linear or quadratic mixture model, and a contour plot of the predicted elongation values of the three responses, namely recombinant product titre, IVCD and *q*P. The design analysis was performed using analysis of variance (ANOVA) where it provided the weights of the experimental factors for all of the responses. Each response was described by either a first-order linear model (Equation 8.1) or a second-order polynomial model (Equation 8.2).

$$y = b_0 + \sum_i b_i x_i + \sum_{ij} b_{ij} x_i x_j + \mathcal{E} \tag{8.1}$$

$$y = b_0 + \sum_i b_i x_i + \sum_{ij} b_{ij} x_i x_j + \sum_i b_{ii} x_i^2 + \mathcal{E} \tag{8.2}$$

where *y* is the response, *b* is the model coefficients that are estimated using a least square fit of the model to the experimental results, *x* is the input factor, *i* and *j* are the design variables and *ε* is the error.

8.3 Results

8.3.1 Cell and process engineering strategies can improve Sp35:Fc production via distinct modes of action

As Sp35:Fc folding and assembly reactions were limiting, we attempted to address the hypothesis that co-expression of specific functional proteins could improve ER capacity for Sp35:Fc folding and assembly reactions, increase secretion rate and/or relieve host cells from ER stress, where the major consequence of UPR activation is a decrease in cell growth. Specifically, we engineered the CHO cell protein factory by overexpressing the ER-resident proteins known to be involved in chaperone machineries (BiP, PDI, ERO1Lβ, CypB) and activated forms of UPR transactivators (the cleaved 50 kDa form of ATF6 and

spliced 54 kDa XBP1) known to generically upregulate the expression of ER chaperones/foldases, as well as by inhibiting PERK, which is involved in translation attenuation and cell cycle arrest mechanism (Pybus et al., 2014a; Schröder, 2008). ERO1Lβ was chosen (over ERO1Lα) as this isoform is induced during the UPR to alleviate ER stress (Pagani et al., 2000) and has been reported to enhance antibody production in transient systems rather than stable (Chapter 2, Table 2.2). All functional genes were human cDNAs driven by the human CMV promoter, while PERK inhibition was achieved by using an inhibitor (IC_{50} = 0.4 nM) added 3 h post-transfection at 0.04% v/v of total culture volume. For each transfection, Sp35:Fc gene load was kept constant and empty vector was utilised to equalise total DNA load where necessary. The ERSE vector system developed in Chapter 7 was not utilised so as to obtain a direct readout of the relative effect of functional genes or chemicals without interference from promoter transactivation. Cell viabilities measured 2 h post-transfection were 84.5%–89.9% (mean 87.9%).

Molecular chaperones, foldases and UPR transactivators exerted molecule-specific effects on either Sp35:Fc productivity and/or cell growth (Figure 8.1A). Overexpression of BiP significantly improved the volumetric titre, mediated via an increase in *q*P, where higher quantities of chaperone having a proportionately higher effect (up to 1.6-fold increase). PDI also improved the production in a similar fashion but to a limited extend, and further improvement was obtained when it was co-expressed with ERO1Lβ. At the optimal amount of 20–40% w/w rDNA, CypB substantially increased the volumetric titre (1.4-fold increase), however this was mediated via an increase in IVCD. Expression of either UPR transactivator gave similar effect to BiP expression, with maximum specific productivity observed at 80% w/w rDNA. However, both transactivators functioned to suppress cell growth rate, leading to a considerably lower IVCD (with no effect on cell viability) that in turn resulted in no significant improvement in volumetric titre. On contrast, PERK inhibitor functioned in the way of CypB, where increased in cell growth was observed (dependent upon inhibition level) although this did not lead to an improvement in product titre due to substantial reduction in *q*P.

Additionally, we utilised chemical chaperones and mild hypothermic condition (32°C) to induce protein folding and stabilisation. Whilst this intervention may not be considered directly specific to cellular processes that govern difficult-to-express recombinant protein production rate, we chose to investigate the relative potency of chemical chaperones that may improve the Sp35:Fc folding and assembly interaction and/or offset the consequence of Sp35:Fc induction of the UPR (de Almeida et al., 2007; Roth et al., 2012). We tested 4-phenylbutyrate (PBA), dimethyl sulfoxide (DMSO), glycerol, betaine and trimethylamine N-

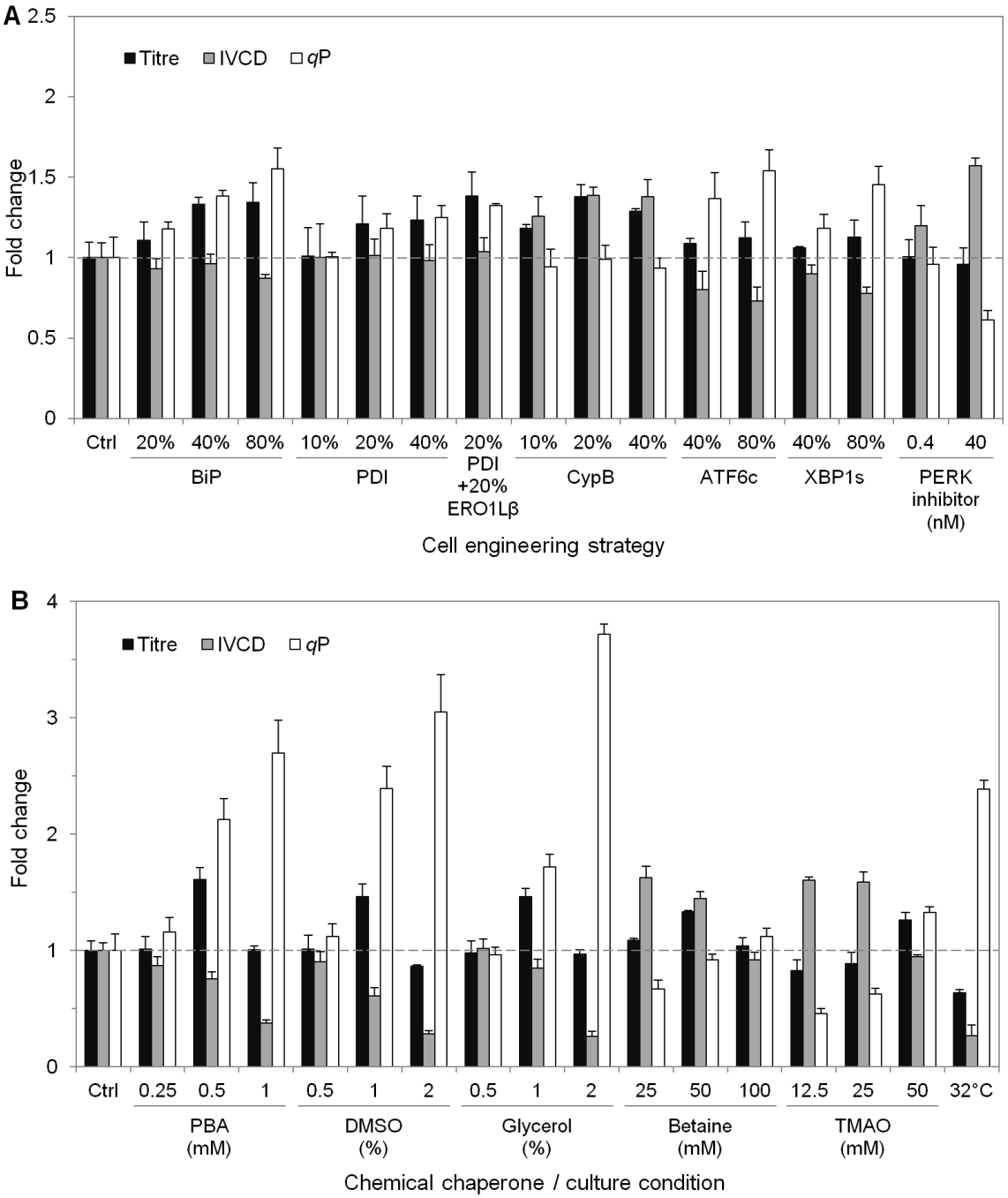

Figure 8.1: Co-expression of molecular chaperones, foldases or UPR transactivators and use of chemical chaperones or hypothermic condition can improve difficult-to-express Sp35:Fc production via two distinct modes. (A) Molecular chaperones, foldases or active form of UPR transactivator gene was co-transfected at 10–80% (w/w) of Sp35:Fc plasmid, while PERK inhibitor (IC_{50} = 0.4 nM) was added 3 h post-transfection. An empty vector was used to equalise total DNA load in all cases. **(B)** Each chemical chaperone was titrated at its active concentration, added 3 h post-transfection. For hypothermic condition, the cells were shifted to 32°C 3 h post-transfection. Data are normalised with respect to control transfection of Sp35:Fc gene alone with equalised total DNA load and cultured under normal condition. Cell culture supernatant was harvested 96 h post-transfection and analysed by ELISA. Error bars represent the standard error of two biological replicates and two technical replicates.

oxide (TMAO) at different concentrations where purified water was used as a solvent and diluent and added at 2% v/v of total culture volume. Concentrations above 2 mM PBA, 3% v/v DMSO, 3% v/v glycerol, 200 mM betaine and 200 mM TMAO resulted in severe growth cessation and cytotoxicity and thus the chemicals were titrated below these concentrations. These data are illustrated in Figure 8.1B.

As shown in Figure 8.1B, the IVCDs of PBA, DMSO and glycerol treated cultures were significantly lower than that of control culture due to depressed cell growth depending on the chemical dosage. The volumetric titres however were largely the same or higher, which imply that the improved Sp35:Fc production in these chemical chaperone-treated cultures were attributed to improved *q*P. In all cases, the productivity appeared to be strongly induced by the chemical concentration, where the highest increase in *q*P was observed at the highest concentration, with up to 3.7-fold increase using 2% v/v glycerol. However, taking the effect of reduction in IVCD, improvement in volumetric titre of product was observed only at a specific (mid) concentration of each chemical. Similar effect was observed with hypothermic condition, although the increase in *q*P was not sufficiently high to counteract the effect of depressed cell growth resulting in a net decrease in volumetric titre. In contrast, the methylamines (betaine and TMAO) appeared to improve cell growth, possibly by protecting host cells from ER stress although at the expense of recombinant protein productivity. This cell growth benefit however was suppressed at 100 mM betaine and 50 mM TMAO concentrations, likely due to high osmolality. Interestingly, the improvement in volumetric titre for betaine was observed at 50 mM with increase in IVCD and sub-optimal *q*P, whereas the improvement for TMAO was observed at 50 mM TMAO which was due to increase in *q*P.

To verify that the effect of CypB, PERK inhibitor, betaine and TMAO on cell growth was due to the recombinant protein production, cells were “mock-transfected” with empty vector and either with CypB gene or addition of PERK inhibitor, betaine or TMAO into culture medium. Results show that the IVCDs were similar in all cases compared to the mock-transfected control culture (data not shown). It is important to note that the increase in volumetric titre for slow growing cells could not be attributed to the effect of slower dilution of plasmid (i.e. higher rDNA copies per cell; discussed briefly in Chapter 7). Higher Sp35:Fc DNA copy number will not increase *q*P due to the expression saturation (Chapter 4; Figure 4.4), and this saturation theoretically allows some dilution of the plasmids without affecting the *q*P. Additionally, without a significant increase in *q*P to counteract the lower IVCD, this would actually result in a net decrease in volumetric titre as observed in the 32°C culture. Further analysis of the engineering strategies in Figure 8.1 showed that there

was a strong negative correlation between their impact on *q*P and on cell growth (Figure 8.2). This data highlights the importance of identifying genes and/or chemicals that can work synergistically to provide an optimal solution, i.e. enable the host cell to achieve both high *q*P and IVCD.

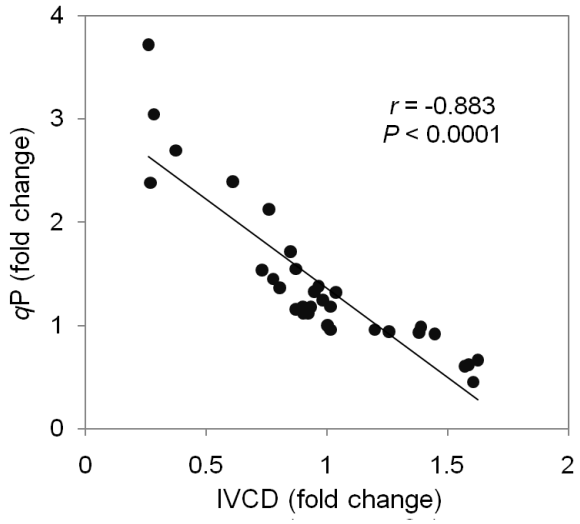

Figure 8.2: Improvement in *q*P using cell and process engineering strategies correlates to repression of cell growth and vice versa. The stimulation of Sp35:Fc productivity is achieved at the expense of cell growth, whilst the stimulation of cell growth would result in a reduction in productivity.

8.3.2 Cell and process engineering strategies may have a combined effect and can be cell line- and protein-specific

Having understood the effect of single molecular gene or chemical chaperone, we studied the effect of combining two cell/process strategies on Sp35:Fc production. These combined engineering strategies resulted in mixed outcomes (Figure 8.3). Firstly, whilst no further improvement was observed in single co-expression of PDI above 20% w/w (Figure 8.1A), further increases of 14% and 29% in volumetric titre were observed when 20% w/w BiP was co-expressed with 20% and 40% w/w PDI, respectively (Figure 8.3). Simultaneous co-expressions of CypB and BiP, where the two proteins acted via different mode of actions, also further enhanced the volumetric titre where a 1.6-fold increase compared to control culture obtained using 20% w/w CypB and 20% w/w BiP.

On the contrary, attempts to re-programme the UPR by co-expressing 40% w/w ATF6c or 40% w/w XBP1s while inhibiting PERK did not significantly enhance the

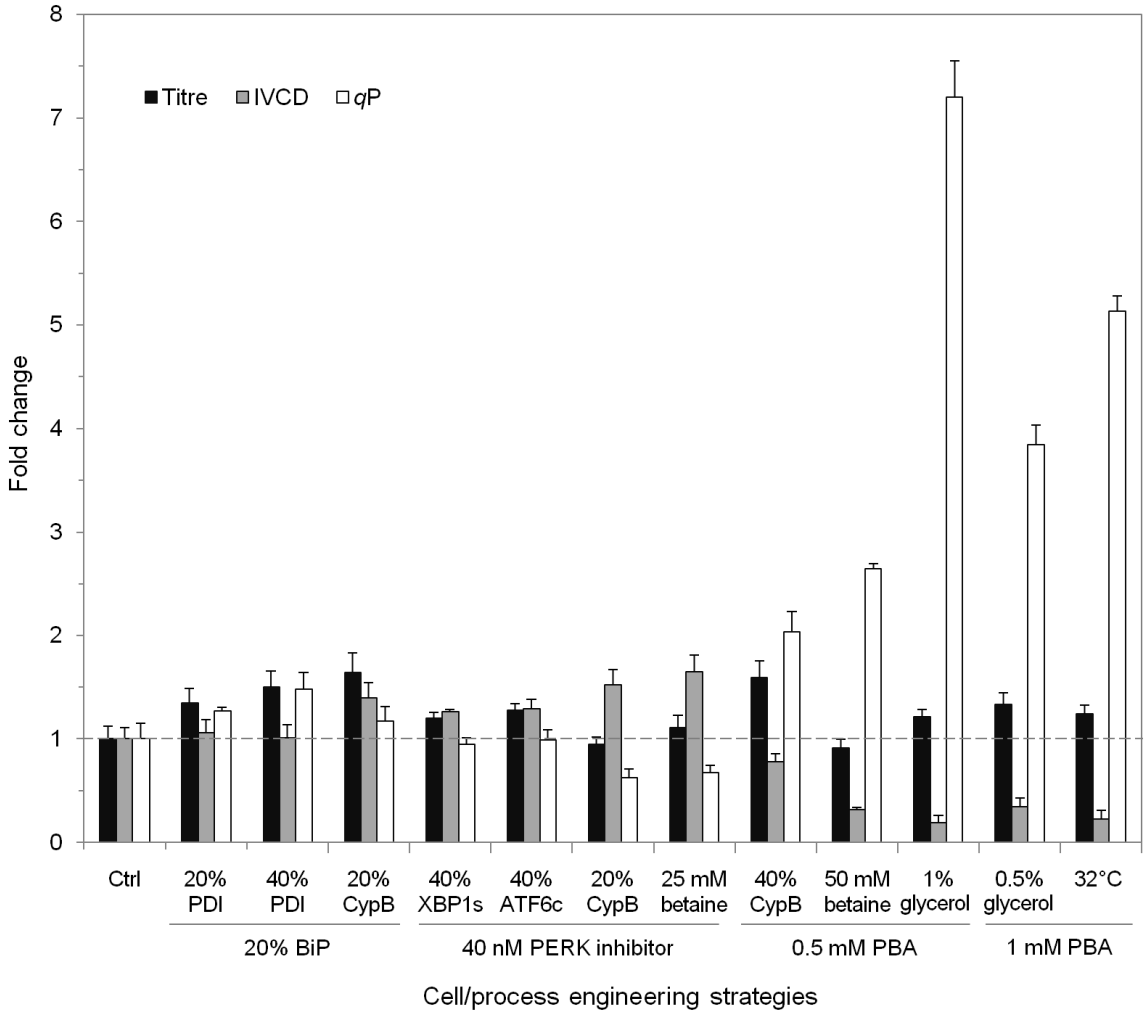

Figure 8.3: Combinations of cell and/or process engineering strategies yielded mixed results. Functional protein plasmids (BiP, PDI, CypB, ATF6c and XBP1s) were co-transfected with Sp35:Fc plasmid at specific functional protein to Sp35:Fc plasmids w/w percentage. Chemicals were added 3 h post-transfection. For hypothermic condition, the cells were shifted to 32°C 3 h post-transfection. Data are normalised with respect to control transfection of Sp35:Fc plasmid alone with equalised total DNA load and cultured under normal condition. Cell culture supernatant was harvested 96 h post-transfection and analysed by ELISA. Error bars represent the standard error of two biological replicates and two technical replicates.

volumetric titre—whilst the IVCD was improved by PERK inhibitor (i.e. by avoiding the translational attenuation mechanism), the improvement in *q*P was not sufficiently high to confer benefit on the production titre (Figure 8.3). Similar results were obtained using 80% w/w ATF6c or 80% w/w XBP1s and 40 nM PERK inhibitor in which the IVCD benefit was offset by a relatively lower *q*P level (data not shown). Moreover, we found out that combining PERK inhibitor and CypB or betaine did not further improve the IVCD, whereas the use of PBA, even at a moderate concentration of 0.5 mM, fully inhibited CypB and

betaine's cell growth benefit (Figure 8.3). In fact, the use of 0.5 mM PBA and 50 mM betaine resulted in lower IVCD compared to 0.5 mM PBA alone although the reason behind this is unclear.

We found that simultaneous utilisation of moderate concentrations of PBA (0.5 mM) and glycerol (1% v/v) in culture substantially increased the *q*P to 7.2-fold compared to control culture, implying that there was a synergistic effect between the two chemicals on Sp35:Fc production (Figure 8.3). This synergistic effect yielded a far higher cell productivity compared to individual and other combined cell/process engineering strategies (Figure 8.4). However, at this combined concentration the cell growth was reduced dramatically (with no effect on cell viability) resulting in only a 1.2-fold increase in volumetric titre. Utilising both 1 mM PBA and 32°C strategies resulted in 5.2-fold increase in *q*P (Figure 8.3) compared to the respective 2.7-fold and 2.4-fold increases when used individually (Figure 8.1B), indicating that both strategies are only additive. Despite having a lower combined specific productivity compared to the dual chemical chaperones strategy, employing this strategy in fed-batch culture may yield better result due to the extended culture duration associated with hypothermic condition.

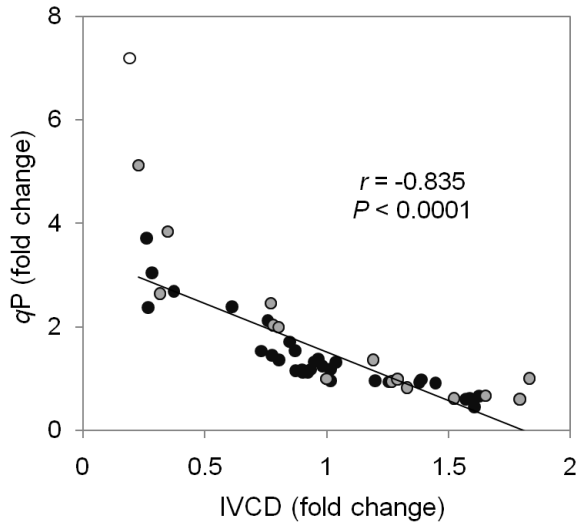

Figure 8.4: Combined engineering strategies can profoundly impact Sp35:Fc production. The combination of PBA and glycerol (open circle) is identified as an outlier that deviates from the general correlation between *q*P and IVCD of individual and combined cell/process engineering strategies (black and grey circles, respectively).

Furthermore, to assess whether the cell and process engineering strategies are cell line-and protein-specific, we tested a number of strategies on two best-performing clones

isolated from the parental population that have distinct characteristics (C13 and C69; Chapter 6), as well as on SEAP-producing cells. These data are shown in Figure 8.5. With respect to the clones, C13 exhibited improved Sp35:Fc production via rapid cell growth without compromising productivity whereas C69 displayed inherently high *q*P with similar growth rate compared to the parental population (Chapter 6, Figure 6.7). Engineering strategies on C13 using CypB, betaine and PBA further enhanced the IVCD or *q*P (Figure 8.5), comparable to the results obtained in the parental population shown in Figure 8.1B.

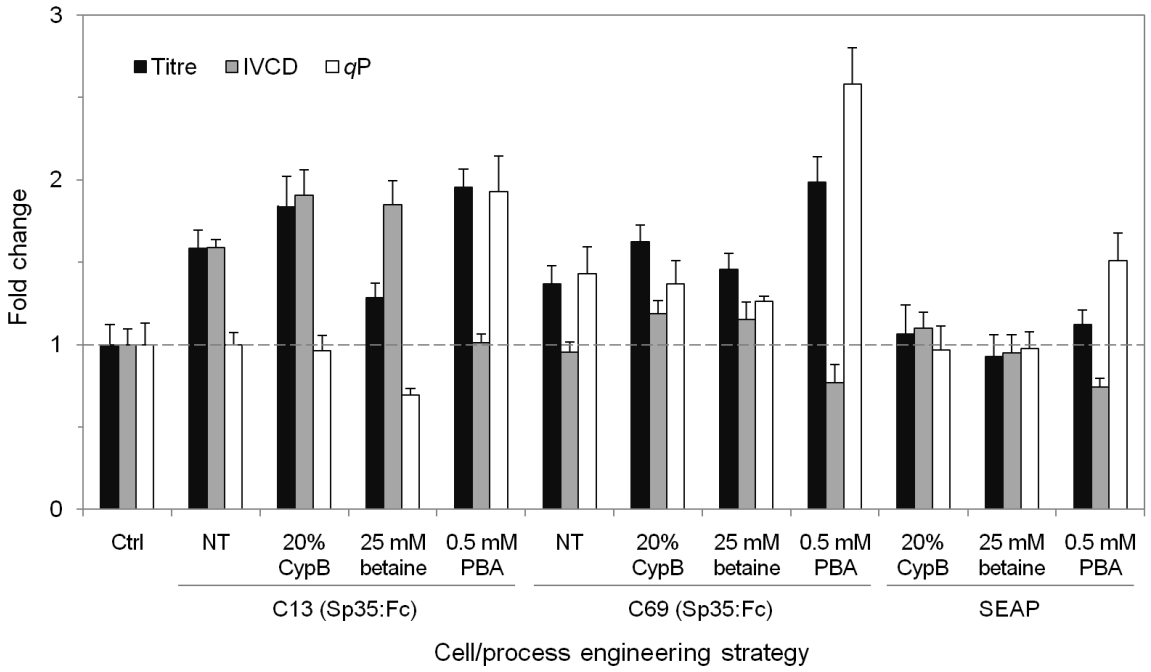

Figure 8.5: Cell and process engineering strategies can be clone and protein specific. In addition to Sp35:Fc, several engineering strategies were tested on Clones 13 and 69 and SEAP production. The C13 and C69 data including the no treatment (NT) controls were normalised to the parental population control culture, whereas SEAP data was normalised to SEAP control culture. Cell culture supernatant was harvested 96 h post-transfection and analysed by ELISA. Error bars represent the standard error of two biological replicates and two technical replicates.

CypB and betaine made less appreciable impact on C69 in which only slight improvements in IVCD were observed (1.2-fold increase; Figure 8.5). We hypothesise that C69 had an innate cellular capability in producing Sp35:Fc (reflected in its higher *q*P) with lower ER stress and therefore was less responsive to CypB and betaine. This is corroborated by the results obtained with easy-to-express (ETE) SEAP where no effects were observed using CypB and betaine as well as the more limited improvement in SEAP specific productivity using PBA. However, treatment with PBA on both C13 and C69

significantly increased their *q*P regardless of their initial level, illustrating that the chemical could affect Sp35:Fc production in a more universal way. Overall, these data illustrated that the effect of molecular and chemical chaperones can be cell line and protein-specific depending on the cell's cellular capability to produce a particular protein.

8.3.3 PDI and ERO1Lβ promote aggregation while chemical chaperones and hypothermic condition suppress aggregation

In an attempt to reduce Sp35:Fc aggregation, the culture supernatants from the individual engineering study were analysed as described in Chapter 4 (Figure 4.5). Western blot analysis of the secreted Sp35:Fc revealed that the molecular chaperones, foldases, chemical chaperones and culture temperature greatly influenced the amount of aggregates produced (Figure 8.6). Whilst chemical chaperones generally suppressed aggregate formation, the increase in folding and assembly rate through overexpression of PDI (particularly in combination with ERO1Lβ) appeared to increase the amount of aggregates. Higher oligomer, likely to be octamer (800 kDa), also became apparent as shown in the Western blot image of Figure 8.6A.

At the extreme, the total amount of Sp35:Fc aggregates increased from 4.8% w/w in the control culture to 10.5% w/w when the recombinant protein was co-expressed with PDI and ERO1Lβ (Figure 8.6B). We hypothesise that whilst ER induction in disulphide bond formation accelerated the Sp35:Fc folding and the overexpression of ERO1Lβ created an altered oxidising environment that was favourable for this reaction, they did not resolve the limiting ER-to-Golgi transport rate in the cells (Chapter 4, Figure 4.6C). Indeed, addressing a limitation in the early secretory pathway without solving the downstream bottleneck may not lead to desired results (Delic et al., 2014) and in the case of Sp35:Fc it would amplify the ER stress and further deteriorate product quality control. In contrast, CypB appeared to have the ability to partially suppress aggregates formation with a 41% reduction from the control (Figure 8.6B). This effect might owe to the dual function of the lumenal protein where it has been implicated both in protein folding (Feige et al., 2009; Kim et al., 2008) and in ERAD of soluble lumenal substrates (Bernasconi et al., 2010). On the other hand, the UPR transactivators and PERK inhibitor did not effectively reduce the amount of aggregates ($P > 0.05$) even though the former is known to regulate the secretory machinery in a global fashion (Schröder, 2008).

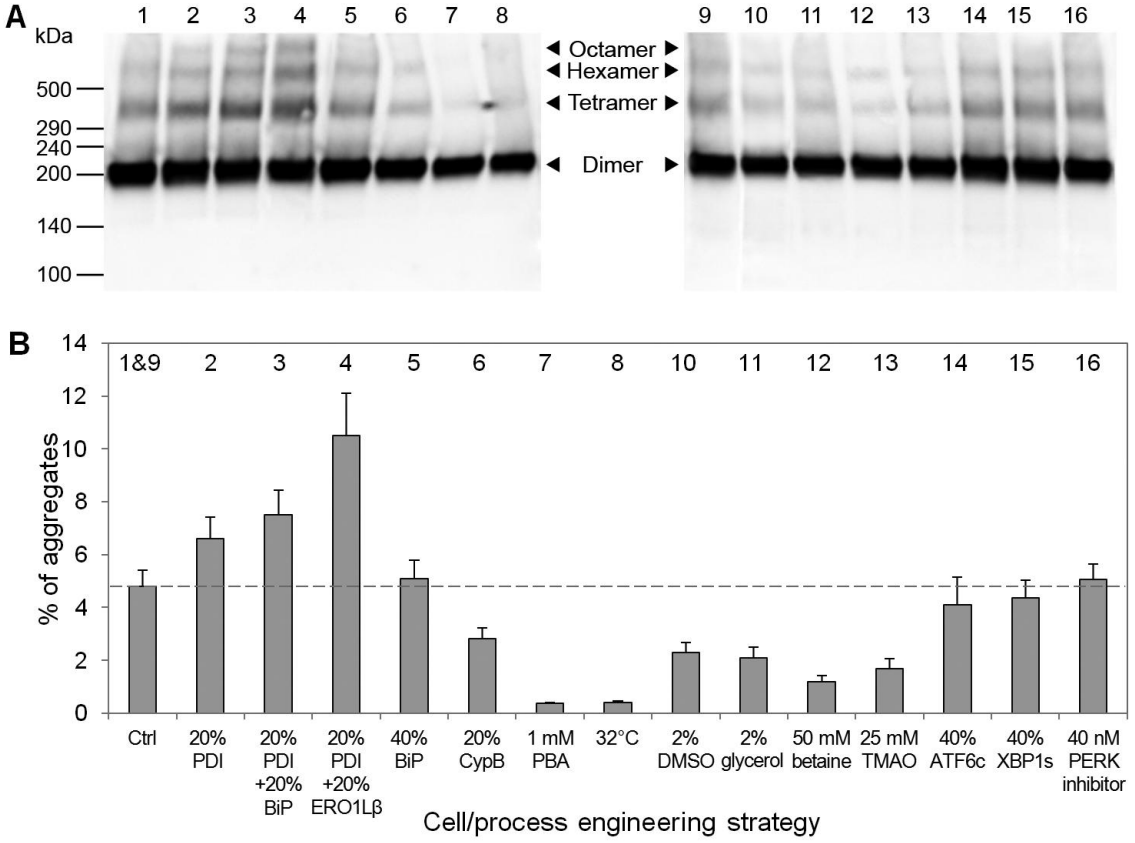

Figure 8.6: Quantitative Western blot of Sp35:Fc from cell culture supernatant revealed the varying amount of aggregates produced under different systems. (A) Representative Western blot image. Culture supernatant was harvested 96 h post-transfection and separated by non-reducing SDS-PAGE in the presence of the alkylating agent NEM to suppress disulphide bond reduction and scrambling. Under reducing condition, all bands were reduced to 100 kDa (data not shown). **(B)** Relative amount of aggregates were quantified using image analysis software. The bar number in the chart corresponds to lane number of the representative Western blot image. Data shown is mean value of two biological and two technical replicates. Error bars represent the standard deviation.

The use of chemical chaperones particularly PBA, as well as hypothermic condition, effectively suppressed aggregates formation (between 52% and 96% reduction; Figure 8.6B). Indeed, several previous studies have shown that chemical chaperones and low culture temperature can reduce the aggregation of recombinant proteins such as β-interferon (Rodriguez et al., 2005; Tharmalingam et al., 2008) and cartilage oligomeric matrix protein angiopoietin-1 (Hwang et al., 2011a,b). As Sp35:Fc aggregates formed intracellularly, we propose that CypB and the chemicals efficiently mediated intracellular folding and target misfolded proteins for further refolding or degradation process, whereas the hypothermic condition shifted the cellular energy used on cell growth/maintenance to the energy extensive recombinant protein folding (discussed in Chapter 4).

8.3.4 Optimal combinations of multiple genes and/or chemical chaperones can be identified using DOE-RSM

We employed a design of experiment–response surface modelling (DOE-RSM) approach to further study the effect of combining two molecules with distinct modes of action, namely BiP and CypB, and a chemical chaperone (PBA) on Sp35:Fc production. Only two functional genes were utilised in order to minimise the transfection cytotoxicity effect. We defined DOE design space boundaries (rDNA amount and chemical concentration ranges) for the three Sp35:Fc production effectors based on the individual study. The Box-Behnken design was chosen due to its minimal number of experiments required while still providing sufficient information to identify key interactions and generate model predictions. Each effector was deployed at a notional low, medium, or high amount/concentration within its design space, yielding 12 discrete coordinates. The mid-point coordinate was replicated five times to determine pure error. Three response variable outputs were measured at each coordinate, specifically the volumetric titre, IVCD and *q*P. All outputs were normalised against control culture that contained neither functional genes nor chemical chaperone.

Quadratic response surface models were used to compare the relationship between CypB, BiP and PBA effectors on titre and IVCD output variables, while two-factor interaction (2FI) response surface model was used on *q*P. The quadratic model derived from the relative titre analysis is as follows:

$$\begin{aligned}\text{Titre} = {} & 1.9 + 0.13(\text{CypB}) + 0.18(\text{BiP}) + 0.22(\text{PBA}) - 0.004(\text{CypB})(\text{BiP}) - \\ & 0.098(\text{PBA})(\text{BiP}) - 0.045(\text{CypB})(\text{PBA}) - 0.12(\text{CypB})^2 - 0.092(\text{BiP})^2 - \\ & 0.27(\text{PBA})^2 \end{aligned} \quad (8.3)$$

Based on the ANOVA, it was found that there were a number of insignificant model terms for IVCD and *q*P outputs, thereby reducing model fit (goodness of fit). CypB for instance was an insignificant effector on *q*P, both individually and in combination with other effectors, where the interaction term was eventually removed from the model equation. The reduced quadratic model for relative IVCD and the linear 2FI model for relative *q*P are described in Equations 8.4 and 8.5, respectively.

$$\begin{aligned}\text{IVCD} = {} & 1.04 + 0.064(\text{CypB}) - 0.069(\text{BiP}) + 0.17(\text{PBA}) - 0.034(\text{PBA})(\text{BiP}) - \\ & 0.036(\text{CypB})(\text{PBA}) - 0.045(\text{BiP})^2 - 0.71(\text{PBA})^2 \end{aligned} \quad (8.4)$$

$$q\text{P} = 1.76 + 0.024(\text{CypB}) + 0.33(\text{BiP}) + 0.55(\text{PBA}) + 0.085(\text{PBA})(\text{BiP}) \quad (8.5)$$

Table 8.1 summarises the response surface models employed. Highly significant models were obtained ($P < 0.05$) for each response and all models displayed an insignificant "lack of fit" indicating comparable variance of modelled and empirical data. Additionally, all models achieved a good agreement between the predicted and adjusted R^2 values (< 0.2 difference) signifying the reliability of the models in predicting the output response to the combination of effectors (Mandenius and Brundin, 2008). The response surface models for relative titre, IVCD and *q*P are graphically represented in Figure 8.7, 8.8 and 8.9, respectively, illustrating the modelled response surfaces for the interaction of two parameters at a constant, mid amount of the third parameter.

Table 8.1: Summary of DOE-RSM analysis of parameters and parameter interactions controlling Sp35:Fc production and CHO cell proliferation.

Response	Factor	Goodness of fit	Probability data > random error	Significant? ($P < 0.05$)	Model predictability (Pred/Adj R^2)
Titre	Model	1.23	0.0001	Yes	0.90/0.94
	CypB	0.13	0.0013	Yes	
	BiP	0.25	0.0002	Yes	
	PBA	0.37	< 0.0001	Yes	
	(CypB)(BiP)	0.00	0.9117	No	
	(PBA)(BiP)	0.04	0.0265	Yes	
	(CypB)(PBA)	0.00	0.2371	No	
	$(CypB)^2$	0.06	0.0097	Yes	
	$(BiP)^2$	0.04	0.0308	Yes	
	$(PBA)^2$	0.31	< 0.0001	Yes	
IVCD	Model	0.34	< 0.0001	Yes	0.72/0.90
	CypB	0.04	0.0041	Yes	
	BiP	0.03	0.0027	Yes	
	PBA	0.23	< 0.0001	Yes	
	(PBA)(BiP)	0.00	0.1893	No	
	(CypB)(PBA)	0.00	0.1642	No	
	$(BiP)^2$	0.00	0.0839	No	
	$(PBA)^2$	0.02	0.0137	Yes	
*q*P	Model	3.35	< 0.0001	Yes	0.88/0.93
	CypB	0.00	0.5922	No	
	BiP	0.89	< 0.0001	Yes	
	PBA	2.43	< 0.0001	Yes	
	(PBA)(BiP)	0.03	0.1967	No	

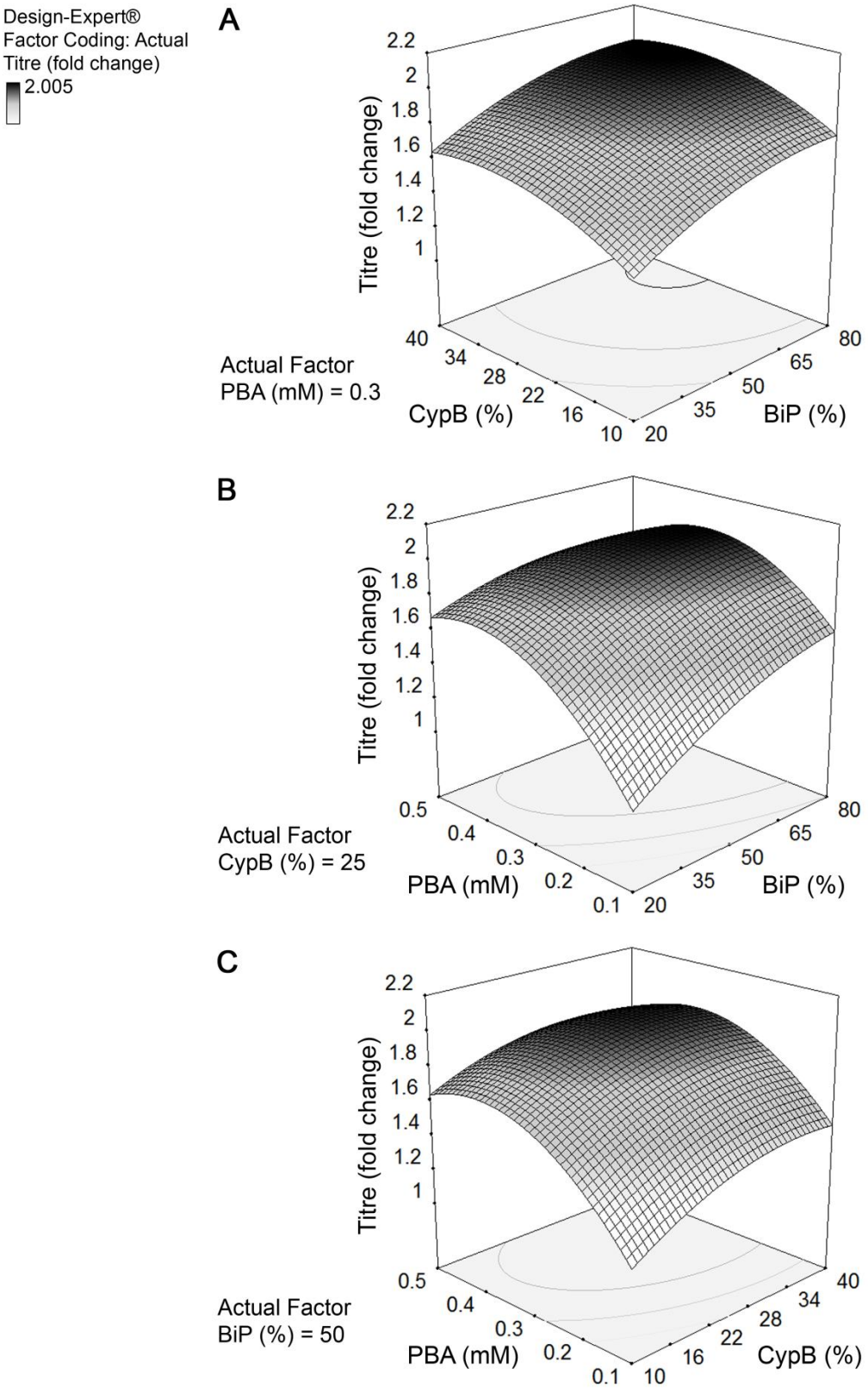

Figure 8.7: DOE response surface models predict Sp35:Fc volumetric titre as a function of CypB and BiP co-expressions and PBA treatment. Modelled response surfaces (96 h culture) are illustrated for the CypB/BiP interaction at fixed 0.3 mM PBA concentration **(A)**, the PBA/BiP interaction at fixed 25% CypB gene load **(B)**, and the PBA/CypB interaction at fixed 50% BiP gene load **(C)**.

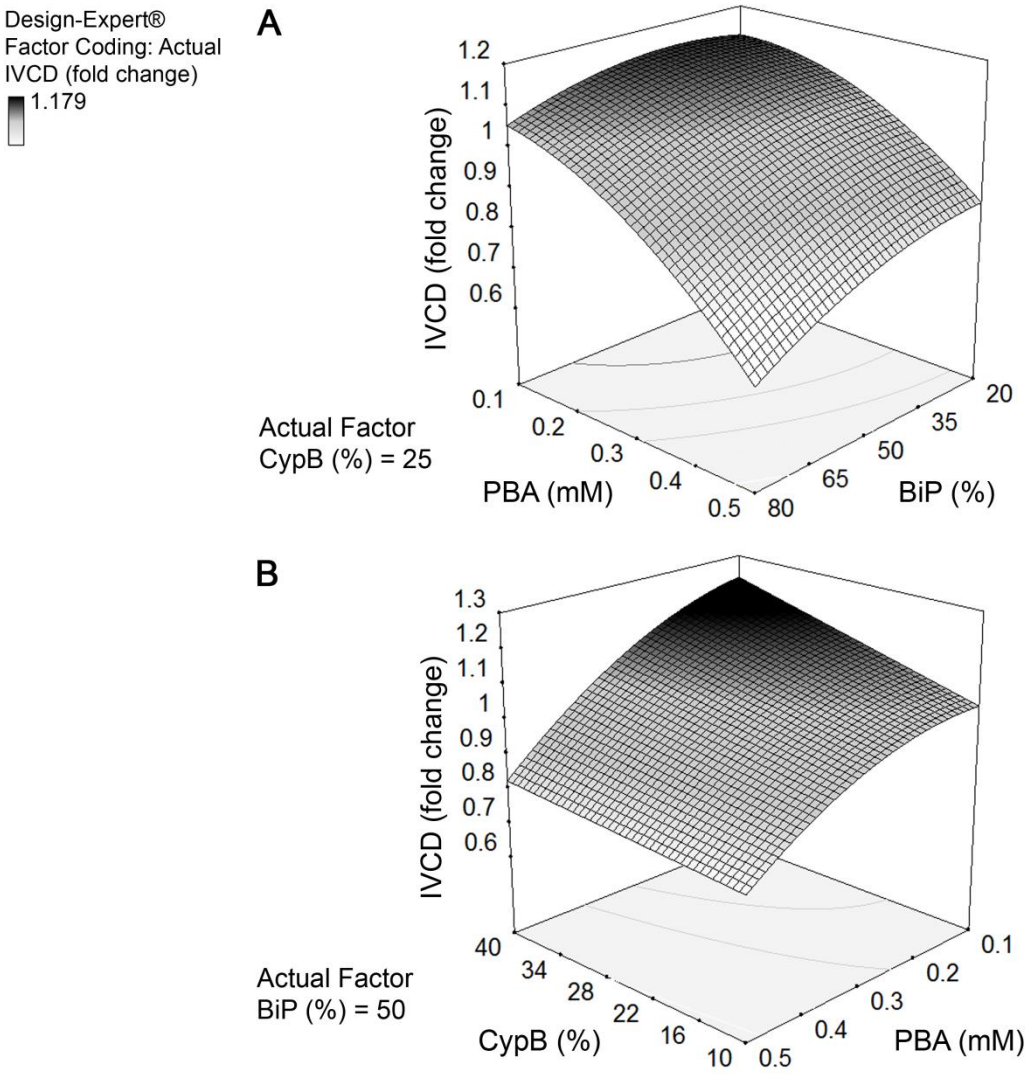

Figure 8.8: DOE response surface models predict CHO cell IVCD as a function of CypB and BiP co-expressions and PBA treatment. Modelled response surfaces (96 h culture) for IVCD are illustrated for the PBA/BiP interaction at fixed 25% w/w CypB gene load **(A)** and the PBA/CypB interaction at fixed 50% w/w BiP gene load **(B)**. The CypB/BiP interaction was insignificant and was removed from the model to improve model fit.

Statistical analysis revealed that all parameters significantly influenced titre individually ($P < 0.01$; Table 8.1) which is reflected in Figure 8.7. High amount of BiP and PBA was particularly favourable on product titre, and there was a significant synergistic effect between the two parameters ($P < 0.05$; Table 8.1). However, the titre appeared to decrease when the concentration of PBA was increased above 0.4 mM (Figure 8.7B and C)—the point at which the inhibition of cell growth by PBA outweighed the increase in *q*P. All parameters were also found to have strong individual influence on IVCD ($P < 0.01$) where the response surface models are presented in Figure 8.8. However, CypB did not

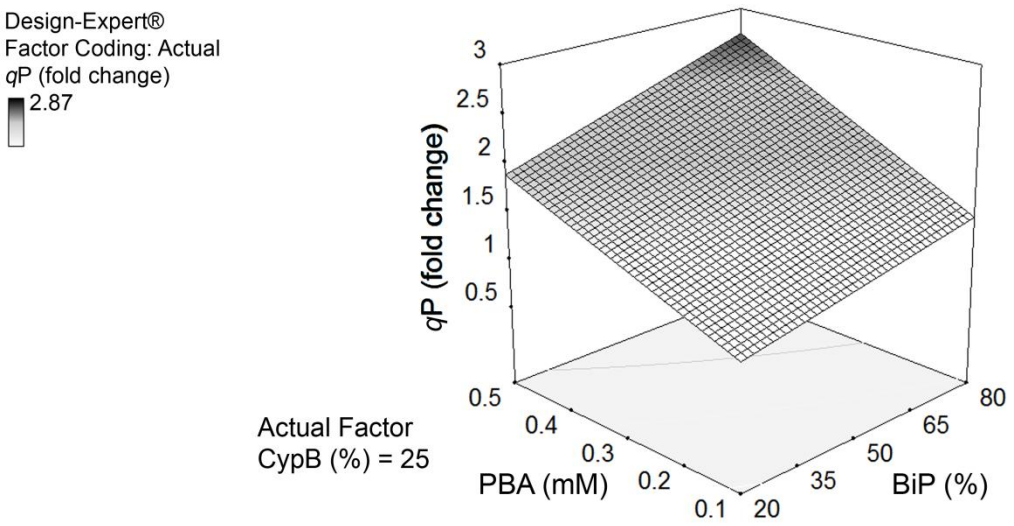

Figure 8.9: DOE response surface models predict CHO cell specific productivity as a function of BiP co-expression and PBA treatment. Modelled response surfaces (96 h culture) for *q*P is illustrated for the PBA/BiP interaction at fixed 25% w/w CypB gene load.

exhibit any interaction with BiP as analysed at fixed 0.3 mM PBA and was removed from the quadratic model in order to improve the model predictability (Pred/Adj R^2; Table 8.1). The deleterious effect of PBA on CypB is also evident in Figure 8.8A and B, where it is clear that CypB's cell growth benefit was completely suppressed at 0.5 mM PBA. Additionally, as the specific productivity is dependent on BiP and PBA (Figure 8.9) and not on CypB ($P > 0.05$; Table 8.1), it can be concluded that the use of CypB alongside such chemical chaperone is incompatible and of limited advantage. The combination of BiP and PBA was able to enhance the *q*P by up to 2.9-fold (Figure 8.9) although there was no interaction between the two strategies ($P > 0.05$; Table 8.1).

Using the response surface models (Equations 8.3–8.5), it is theoretically possible to predict a desired combination of effectors to achieve a specific objective. For example, the models predict that a maximal 2.0-fold increase in volumetric titre with 1.0-fold change in IVCD can be achieved by using the combination of 32% w/w CypB, 70% w/w BiP and 0.32 mM PBA. Moreover, even though our models did not take into account the cell growth/death and protein production during stationary and decline phases, it is still possible to utilise the short-term batch culture data presented here for improved fed-batch culture. Specifically, the combination of 37% w/w CypB, 45% w/w BiP and 0.21 mM PBA is predicted to yield an optimal titre and IVCD (equal weight) of 1.9-fold and 1.1-fold increases, respectively, that could be more advantageous in longer culture duration (i.e., fed-batch culture).

8.3.5 Increased transient production of Sp35:Fc through integrated cell and process engineering

Derived from our cell and process engineering data on short-term Sp35:Fc production, we designed two different biphasic culture strategies for transient fed-batch production. For both biphasic culture strategies, the first phase was centred on cell growth in which we employed co-expression of 20% w/w CypB to enable rapid proliferation of cells. The second phase began once the cells reached the peak VCD (approximately 24×10^6 cells mL^{-1}) on Day 6 where we utilised two different process combinations; (i) 1 mM PBA and temperature shift to 32°C, and (ii) 0.5 mM PBA and 1% v/v glycerol. Additionally, we employed a biphasic culture with only a shift in culture temperature (i.e. without CypB co-expression and chemical addition) to mimic the commonly used biphasic process strategy. This culture was shifted to 32°C on Day 7.

As shown in Figure 8.10A, CypB co-expression facilitated the growth of Sp35:Fc-producing cells, allowing the cultures to reach peak VCD one day earlier compared to the control as well as the "pure" 32°C biphasic culture. With regard to the second phase culture, the combination of 1 mM PBA and 32°C extended the stationary phase substantially, leading to a 24% increase in IVCD compared to the 0.5 mM PBA and 1% v/v glycerol strategy, and a 70% increase over the control (Figure 8.10B). Nevertheless, despite the lower IVCD (Figure 8.10B), the synergistic effect of the latter process strategy outperformed the former in terms of volumetric titre (Figure 8.10C). Specifically, 20% w/w CypB, 0.5 mM PBA and 1% v/v glycerol resulted in a 5.9-fold increase in volumetric titre compared to the control culture of 6.6 µg mL^{-1} while 20% w/w CypB, 1 mM PBA and 32°C gave a 4.4-fold increase. In both cases, the combinations of cell and process engineering strategies are far better than the typical process strategy of shift in culture temperature only (which gave only 1.9-fold increase compared to control).

Furthermore, we tested the optimised strategy (20% w/w CypB, 0.5 mM PBA and 1% v/v glycerol) on the C13 (improved production via rapid cell growth) and C69 (inherently high DTE Sp35:Fc productivity). Results show that C13 displayed significantly improved cell growth and peak VCD during fed-batch culture (Figure 8.11A) which in turn resulted in a 47% increase in IVCD compared to the parental cell line control culture (Figure 8.11B). However, the volumetric titres of both C13 and C69 were limited to about 12 µg mL^{-1} (1.9-fold increase; Figure 8.11C), illustrating the natural "productivity boundary" of the host cell background. Using the optimised, biphasic culture strategy, the volumetric titres were increased to about 41 µg mL^{-1} (Figure 8.11C) hence demonstrating that the use of cell and

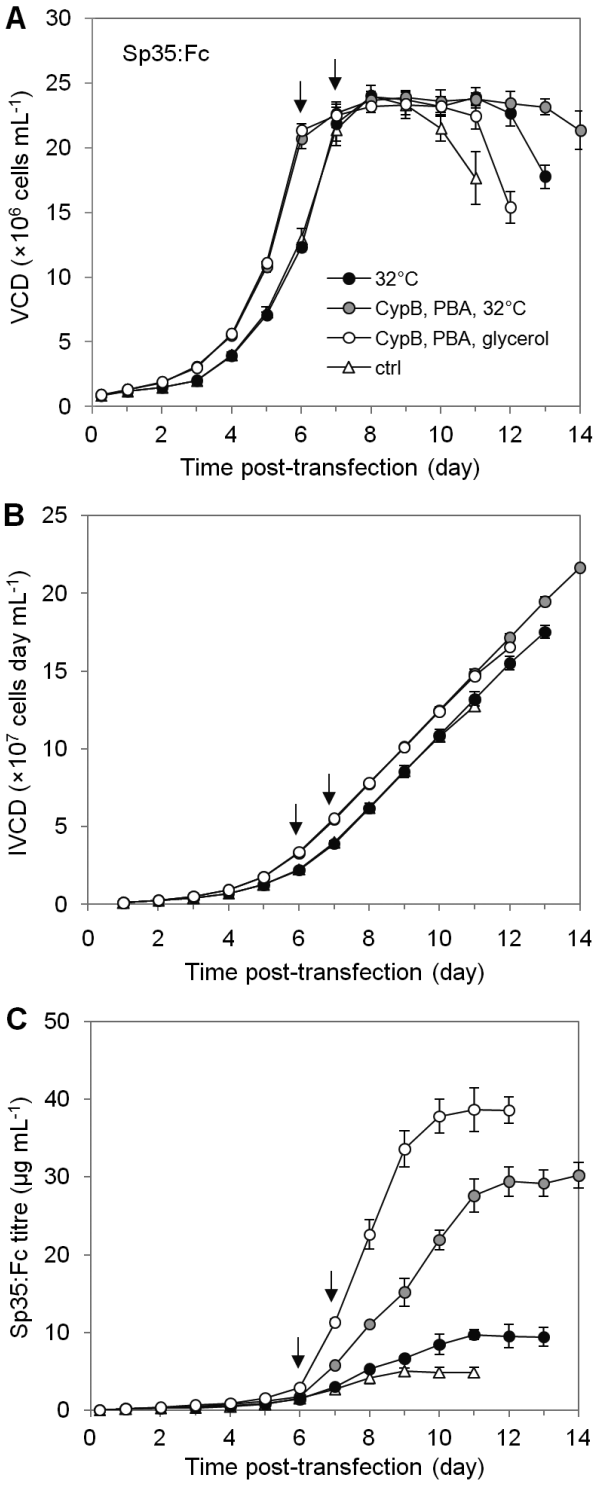

Figure 8.10: Integrated engineering strategies for transient Sp35:Fc productions. 2.5×10^7 cells were transfected with 16.67 µg Sp35:Fc plasmid and subjected to different strategies; a shift to hypothermic condition (●), co-transfection with 20% w/w CypB and addition of 1 mM PBA and shift to hypothermic condition (●), or co-transfection with 20% w/w CypB and treatment with 0.5 mM PBA and 1% v/v glycerol (○). The cells were cultured alongside a control that contained neither CypB gene nor chemical chaperones and maintained at 37°C (△). Arrows indicate the timing of chemical addition and/or temperature shift. **(A)** Viable cell density (VCD) profile. **(B)** Integral of viable cell density (IVCD) profile. **(C)** Sp35:Fc titre profile. Error bars represent the standard deviation of two biological and two technical replicates.

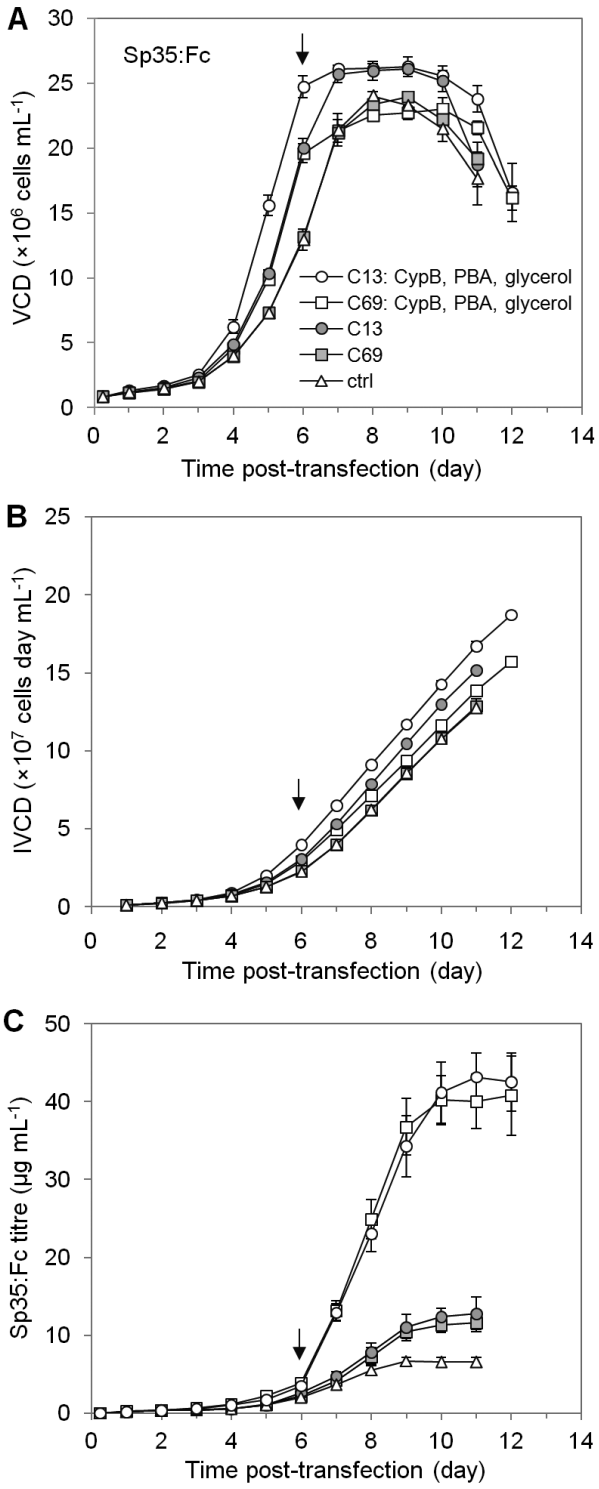

Figure 8.11: Integrated engineering strategies for transient Sp35:Fc productions using two clonally derived cell lines. 2.5×10^7 cells were co-transfected with 16.67 µg Sp35:Fc plasmid and 20% w/w CypB, and treated with 0.5 mM PBA and 1% v/v glycerol using either C13 (○) or C69 (□) cell lines. The cells were cultured alongside a C13 control (●), a C69 control (■) and the parental cell line control (△) that contained neither CypB gene nor chemical chaperones. All cells were maintained at 37°C. Arrow indicates the timing of chemical addition. **(A)** Viable cell density (VCD) profile. **(B)** Integral of viable cell density (IVCD) profile. **(C)** Sp35:Fc titre profile. Error bars represent the standard deviation of two biological and two technical replicates.

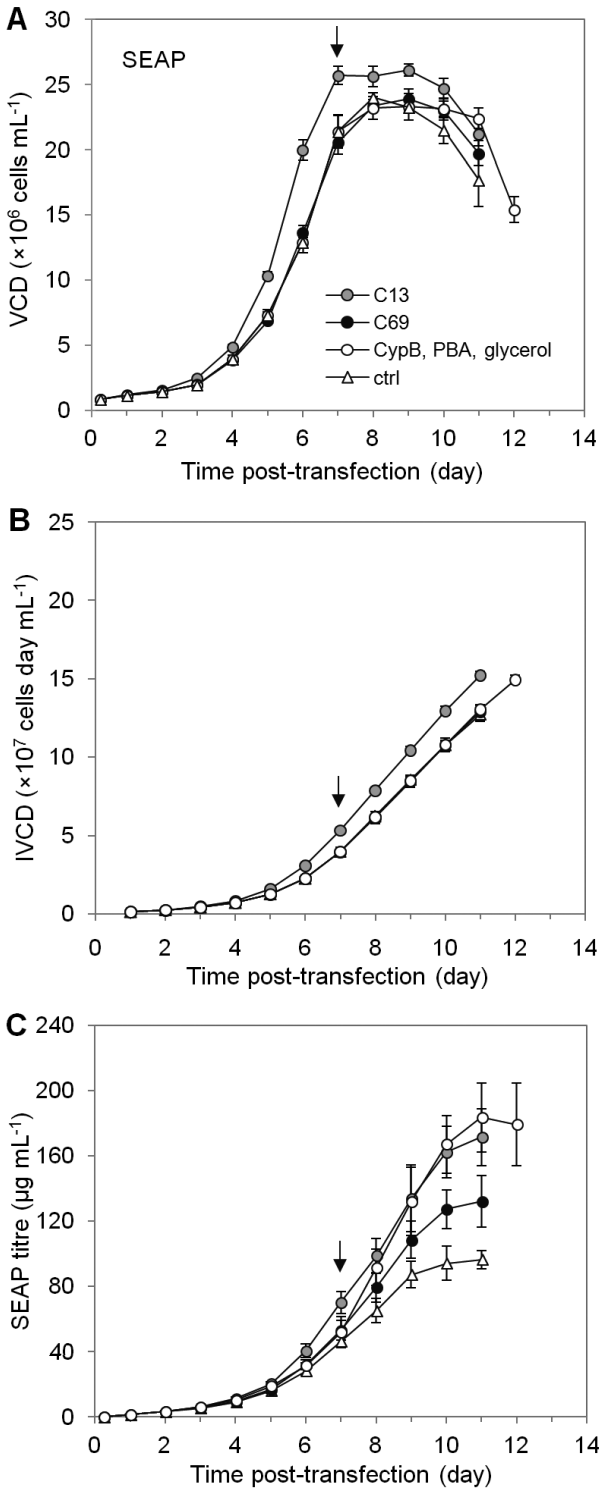

Figure 8.12: Integrated engineering strategies for transient SEAP productions and using two clonal cell lines. 2.5×10^7 cells were transfected with 16.67 µg SEAP plasmid using C13 (●) and C69 (●) cell lines, or co-transfected with 20% w/w CypB and treated with 0.5 mM PBA and 1% v/v glycerol using the parental cell line (○). The cells were cultured alongside a SEAP-producing parental cell line control (△) that contained neither CypB gene nor chemical chaperones. Arrows indicate the timing of chemical addition. **(A)** Viable cell density (VCD) profile. **(B)** Integral of viable cell density (IVCD) profile. **(C)** SEAP titre profile. Error bars represent the standard deviation of two biological and two technical replicates.

process engineering was a much more effectual strategy in controlling the production of DTE protein than exploiting the inherent heterogeneity in the host cell population for improved production.

Conversely, the implementation of the optimised strategy (20% w/w CypB, 0.5 mM PBA and 1% v/v glycerol) on ETE SEAP showed that CypB had no effect on the cell growth (Figure 8.12A) resulting in no improvement in IVCD (Figure 8.12B). Moreover, the additions of PBA and glycerol into the SEAP culture increased the volumetric titre by only 1.9-fold which was comparable to the volumetric titre achieved by C13 without any engineering strategies (Figure 8.12C). Together, these data indicate that the designed engineering strategies employed were more valuable for DTE protein productions. However, as we still observed some improvement on the SEAP titre using PBA and glycerol, this suggests that the treatment of cell culture with chemical chaperones could be an easy-to-implement, generic tool (owing to their versatility of general and specific effects) in improving transient expression compared to the cell engineering strategy.

8.3.6 Chemical chaperones may enhance Sp35:Fc folding and assembly reaction

To assess the aggregate contents and whether the improvements in Sp35:Fc productivity were due to improvement in cellular activities, cell culture supernatant of the final harvest and equal amount of cells on Day 8 of the fed-batch culture were lysed. The extra and intracellular Sp35:Fc polypeptides were immunoblotted as shown in Figure 8.13A. In both cases, no Sp35:Fc aggregates were detected for culture employing chemical(s) treatment, hence proving that the combinations of the molecular/chemical chaperones were effective in inhibiting intracellular aggregate formation during fed-batch culture. Additionally, the amount of aggregates were considerably reduced (67% reduction compared to control) in cells subjected to hypothermic condition alone. A slight decrease (29% reduction compared to control) was also observed with C69 employing no engineering strategies which was consistent with the previous data obtained in Chapter 6 (Figure 6.9).

Further quantitative image analysis of the intracellular samples (Figure 8.13B) revealed that the engineering strategies resulted in increases in the amount of nascent monomer by up to 2.4-fold, likely through an increase in transcriptional rate and/or mRNA stability (Hwang et al., 2011a; Jiang and Sharfstein, 2008). Importantly, we observed that the engineering strategies resulted in increases in dimer to monomer ratio (indicated by

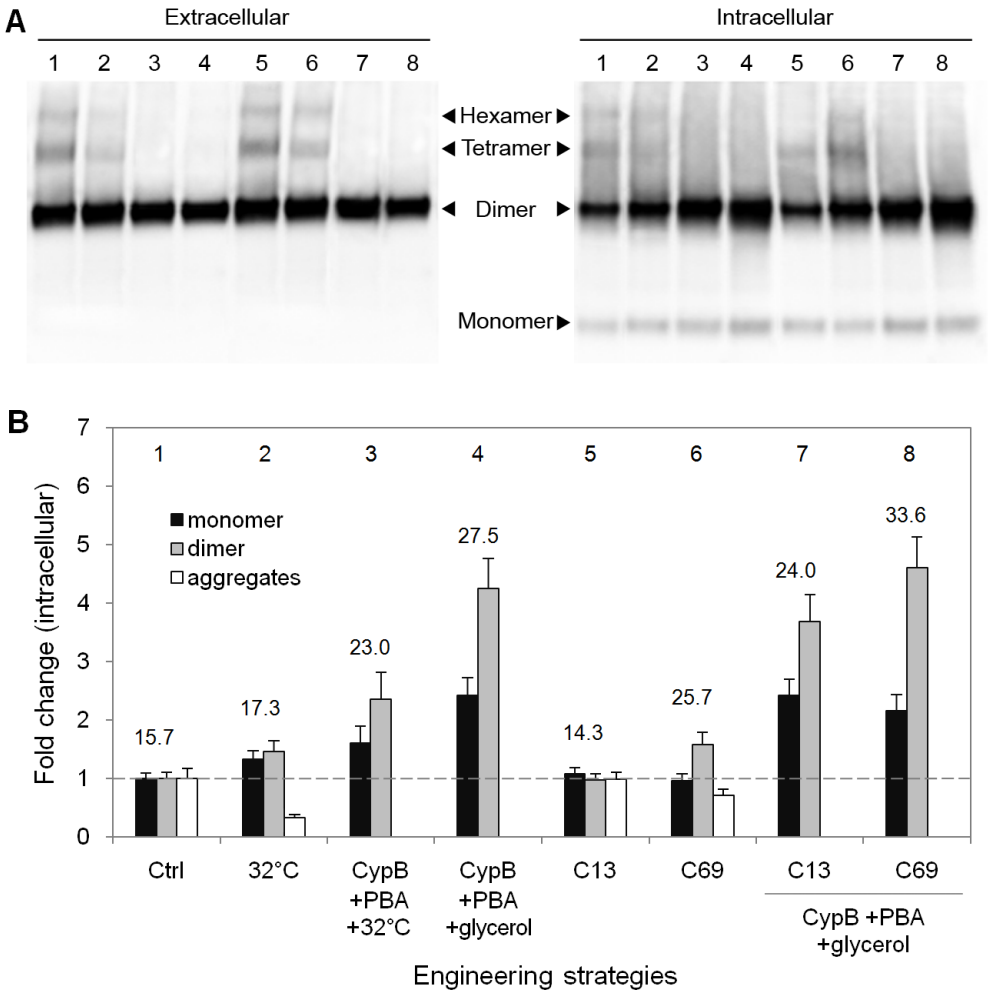

Figure 8.13: Western blot analysis revealed that the improvements in Sp35:Fc production were due to improvement in cellular activities. (A) Representative Western blots of Sp35:Fc polypeptide species in extracellular culture supernatant (final harvest) and intracellular species of equal amount of cells (2×10^4 cells) on Day 8 of fed-batch culture of different culture strategies and clones. The lane number corresponds to the bar chart number for different engineering strategy and/or clone used. **(B)** Intracellular Sp35:Fc unfolded monomer and assembled dimer per cell determined by quantitative Western blot. Error bars represent the standard deviation of two biological and two technical replicates. The dimer to monomer ratio per cell for each engineering strategy and/or clone is indicated above the bars.

the number on respective bars) which imply that the molecular/chemical chaperones, and to a lesser extent the hypothermic condition, restored the cells' cellular function. It is also interesting to point out that C69 which had a relatively high *q*P and low aggregates (Chapter 6, Figures 6.7 and 6.9), had an intrinsically high dimer to monomer ratio (Figure 8.13B) that indicate an inherent efficient folding/assembly processes. We conclude that the integrated cell and process engineering effectively improved the DTE protein production and product quality.

8.4 Discussion

TGE systems are an attractive method for rapid production of research-grade biologics but suffer from relatively low product titres. Whilst improvements through optimisation of large-scale transfection protocols and/or culture media during transient expression are well established (Cain et al., 2013; Daramola et al., 2014; Rajendra et al., 2011; Raymond et al., 2011), for DTE proteins it is imperative to study the effects of engineering the secretory pathway and exocytosis to relieve the posttranslational bottlenecks. Therefore, in this study we investigated the effect of a range of molecular and chemical chaperones and UPR re-programming on CHO cell growth and Sp35:Fc production. Fundamentally, we observed that Sp35:Fc production could be improved via two distinct routes albeit at varying degrees; (i) stimulation of *q*P with general suppression in cell growth (BiP, PDI, ATF6, XBP1, PBA, DMSO, glycerol and hypothermic condition), and (ii) stimulation of cell growth correlated to repression of *q*P (CypB, PERK inhibitor, betaine and TMAO).

The effects of molecular chaperones and foldases are likely to be protein specific and often do not result in the desirable outcomes in the host cells (Borth et al., 2005; Davis et al., 2000; Dorner and Kaufman, 1994; Grubb et al., 2012). In this study, the overexpression protein folding catalyst PDI in CHO cells resulted in increase in aggregates formation, which is in contrast to report by van den Berg et al. (1999) demonstrating that PDI was particularly effective in preventing lysozyme aggregation under crowded conditions *in vitro*. In this context, cell engineering with UPR-related genes is a more useful strategy for enhancing the secretion of recombinant proteins by increasing concentrations of several chaperones in a functionally meaningful ratio (Delic et al., 2014). However, the overexpression of UPR-related genes leads to a general survival disadvantage via increased apoptosis where the effects had been observed in both stable cell lines (Becker et al., 2008) and transient expression of DTE MAbs (Pybus et al., 2014a) which was consistent with our data on Sp35:Fc production. Even though the combination with anti-apoptotic engineering may rescue the cells (e.g., co-expression of x-linked inhibitor of apoptosis; Becker et al., 2010), we also did not observe any benefit of UPR transactivators in preventing Sp35:Fc aggregation, thus rendering them a less attractive strategy.

For DTE Sp35:Fc, our data show that CypB and PERK inhibitor had positive impacts on cell proliferation where the former corroborates recent work in this laboratory on CHO cells expressing DTE MAbs (Pybus et al., 2014). While we may infer that the basic function of the PERK inhibitor is to restrain eIF1α phosphorylation and hence retain global mRNA translation rates, it was interesting to note that CypB was able to partially suppressed

Sp35:Fc aggregation, as well as act in concert with BiP to simultaneously increase both μ and *q*P. These effects might owe to the dual function of the lumenal protein where it has been implicated both in protein folding (Feige et al., 2009; Kim et al., 2008) and in ERAD of soluble lumenal substrates (Bernasconi et al., 2010). However, the impacts of the functional proteins were generally lower compared to the use of chemical additives. As we did not measure the amounts of the functional proteins expressed nor their mRNA levels, further work is required to verify that the limited effects of the cell engineering strategy were not actually due to the inability of the cell hosts to overexpress the functional genes. On a different note, in order to achieve high transfection efficiency with low toxicity, we found that the cell engineering strategy for our fed-batch system was limited to co-expression of a single functional protein, hence necessitating additional process engineering approaches to effectively improve the production titre.

Chemical chaperones, despite have been intensively studied/used for clinical applications (reviewed in Papp and Csermely (2006) and Perlmutter (2002)), are still under-utilised in the therapeutic protein production industry. With the exception of sodium butyrate (and its HDAC inhibitor counterparts), there are limited reports on how these small molecules can benefit the biomanufacturing process, especially for DTE proteins. Another interesting finding of our work is that a moderate concentration of methylamines (betaine and TMAO) in CHO cell culture facilitated cell growth, possibly by alleviating ER stress via improved intracellular Sp35:Fc trafficking (Roth et al., 2012). Even though *q*P was compromised, we speculate that the chemicals could potentially be used in cell line development for DTE protein production. Obviously, sustained UPR activation due to DTE protein overexpression is incompatible with high cell specific growth rate, and stable transfectants whose proliferation is compromised by induction of a UPR do not persist (the main selective pressure within functionally heterogeneous cell populations; Davies et al., 2013). Therefore, by treating a stable clone (or transfectant pool) with a methylamine (or PERK inhibitor) at the beginning of cell line development stage, we may be able to sustain/enrich the cell populations with high producing cells. The UPR-suppressing chemical can then be withdrawn during the manufacturing process and replaced with other chemicals aiming at optimising productivity such as PBA.

Consistent to previous reports, we observed that the addition of PBA, DMSO and glycerol to culture medium significantly increased *q*P and at the same time restrained CHO cell growth in a dose-dependent manner (Hwang et al., 2011a). Importantly, we found that there could be a synergistic effect between two chemicals, where a 7.2-fold increase in *q*P was obtained by combining PBA with glycerol. Different chemical chaperones are known to

act in different ways on recombinant proteins. For example, tauroursodeoxycholic acid has been shown to block the UPR activation in HFE C282Y-producing cells by acting directly on its signal transduction pathway, whereas PBA suppresses the ER stress by enhancing the ER capacity to degrade misfolded HFE C282Y protein therefore preventing the formation of intracellular aggregates (de Almeida et al., 2007). In this study, we hypothesise that PBA and glycerol complemented each other's function in facilitating Sp35:Fc expression, yielding a far higher cell productivity compared to when used alone. However, as the combination of "protective" betaine (Chapter 2, Table 2.1) with PBA did not avert the cytotoxic effect of the latter (Figure 8.3A) we consider that overexpression of antiapoptotic proteins Bcl-2 (Kim and Lee, 2000; Sung and Lee, 2005) and Bcl-x_L (Kim et al., 2011b) or antisense RNA of caspase-3 (Kim and Lee, 2002) may be the effectual way to inhibit chemical chaperone-induced apoptosis. Additionally, our data show that the chemical chaperones were extremely effective in suppressing Sp35:Fc aggregates formation. However, the exact mechanisms underlying the ability of chemical chaperones to inhibit aggregate formation are less clear, and although a number of studies have provided some detailed understanding on this subject (e.g. by shifting the native protein toward more compact confirmation via preferential hydration; Vagenende et al., 2009), the mechanisms are likely to differ from one chemical to another.

Besides chemical additives, it has been reported that the formation of aggregates of recombinant proteins may be reduced by hypothermic culture temperatures (Estes et al., 2015; Hwang et al., 2011b), which was also observed in this study. This is in general agreement with our previous hypothesis (Chapter 4) that Sp35:Fc misfolding/aggregation was due to improper *N*-glycosylation. The distribution of *N*-glycan structures (microheterogeneity) and the levels of site occupancy (macroheterogeneity) may be regulated via culture conditions (including low culture temperature and addition of butyrate; Butler, 2006), which in turn lead to proper functioning of the quality control mechanism to ensure that only correctly folded proteins are transported to the Golgi complex. Additionally, the physical effects of low culture temperature can described by the second law of thermodynamics, where stable systems occupy the states with lowest Gibbs free energy ($\Delta G = \Delta H - T\Delta S$). The rate-limiting step in the folding process is the formation of the folding intermediate, i.e. the conformation that has the highest Gibbs free energy. Lowering the system's temperature (T) would minimise this rate-limiting step, as well as further reduces the enthalpy (ΔH) by lowering the reaction rates of non-bonding interactions, therefore promoting protein folding and stability (Szilágyi et al., 2007).

In addition to their advantages in cell culture processes as presented in this study, chemical chaperones also hold the prospect for utilisation in the downstream formulation process to improve product stability. Likewise, chemicals that have been used in product formulations hold promise to improve the production process. For example, trehalose, which is used as a protectant for several commercially available therapeutic antibodies and proteins, has recently been demonstrated to significantly effective in minimising antibody aggregates as well as in improving the specific and volumetric antibody production titre (Onitsuka et al., 2014). This is highly relevant considering that chemical chaperones are protein and cell line specific. For example, whilst treatment of CHO cell lines with butyrate stimulated up to 15-fold enhancement of yellow fluorescent protein (YFP) production, only up to a 2-fold increase was observed in MAb production (Hunt et al., 2002). Additionally, HDAC inhibitor valproic acid (VPA) had been shown to be more effective in HEK293 cells with 4-fold increase in MAb titres compared to in CHO cells (1.5-fold increase; Backliwal et al., 2008). The data from this study also corroborate those observations in which PBA appeared to have less impact on SEAP production while betaine did not confer any growth benefit onto the SEAP-producing cells.

We have shown that Sp35:Fc production can be controlled through the DOE methodology or biphasic culture strategy. With respect to the former, it was possible to pinpoint the exact functional genes and chemical chaperone concentration required to produce Sp35:Fc with defined levels of cell growth. Desirability analysis identified the combination of CypB, BiP and PBA predicted to yield the highest volumetric titre, without compromising cell growth (i.e., similar IVCD to control). However, the PBA effectively suppressed the CypB's growth benefit. Instead of attempting to reverse the apoptotic effect of the chemical chaperones by transfecting in Bcl-2 or Bcl-x_L (Kim and Lee, 2000; Sung and Lee, 2005), we used a simple biphasic culture that allowed high productivity without compromising cell growth—enabling up to 6-fold increase in volumetric titre. Interestingly, the same strategies were also able to increase SEAP production, although the fact that the impact was less profound (1.9-fold increase) suggested that engineering strategies still need to be protein specific.

Overall, from a functional perspective, for a host cell with a product induced UPR it is feasible to closely control the negative relationship between qP and μ (Figure 8.4). Our data imply that inducing cellular folding/assembly capacity or the UPR components are effective for short-term specific productivity enhancement but incompatible with the creation of productive biomass over the longer term. This is consistent with our preceding observation (Chapter 4, Figure 4,7) where the activation of the UPR to promote protein

folding would prevent cell growth, and it was unlikely that overexpression of ER chaperones actually served to directly "switch-off" the UPR-mediated growth arrest. In contrast, using specific functional genes or chemicals it is possible to either inhibit or offset, via different mechanisms, the effects of the UPR to either permit cell proliferation (e.g. using CypB) or DTE protein expression (e.g. using PBA). Therefore, whilst the optimal combination of functional genes and/or chemicals tends to be protein-specific, from the recognised function(s) of the ER functional effectors we utilised, we can deduce that direct inhibition or alleviation of ER load/stress may be preferable to attempts to indirectly reduce ER load by a general increase in secretory pathway capacity or the rate of forward folding/assembly reaction within the ER. This is also consistent with our *in silico* analysis (Chapter 5, Figures 5.5 and 5.6) which suggests that the translational attenuation mechanism of the UPR (i.e. direct reduction of ER load) is required and sufficient in maintaining cellular homeostasis, whereas engineering the rate of folding/assembly reaction has no apparent benefits, nor it could solve the secondary trafficking bottleneck (rate of ER to *cis*-Golgi) to reduce UPR severity and permit cell biomass accumulation.

Finally, our study illustrates that the optimal engineering strategy for a given recombinant protein requires synergistic combination of genes, chemicals and/or process conditions designed to overcome cellular rate limiting steps simultaneously, enabling the host cell to attain both high *q*P and IVCD. Indeed, the highest TGE titre of 2 g/L reported by Daramola et al. (2014) in CHO cells was obtained by combining powerful expression vector technologies with optimised polyethyleneimine (PEI)-based transfection condition and fed-batch cell culture process. Comparatively, Backliwal et al. (2008) achieved 1.1 g/L titre in HEK293 cells with a combination of optimised expression vector, highly efficient PEI transfection, use of specific growth factor and cell cycle regulators, as well as treatment of cells with VPA. We envisage that the increasingly complex, difficult to produce biologics in the biopharmaceutical industry will necessitate a multi-component cell/process design and engineering platform to create attuned cell phenotype and process condition. Such an approach would require an appropriate multi-gene expression technology (Chatellard et al., 2007; Kriz et al., 2010) and rapid high-throughput statistical assessment of the optimal combination of multiple genetic and chemical components.

Chapter 9

Conclusions and Outlook

This chapter summarises the data presented in this thesis by highlighting the important findings and outcomes, specifically for the production of Sp35:Fc fusion protein and difficult-to-express proteins in general. The prospect of continuation work is also briefly described.

9.1 Conclusions

In biologics development, it can be said that the low-hanging fruits have been harvested. Today, the bioindustry invests far more on R&D and produces fewer new molecules than it did two decades ago. Indeed, while the approval of 112 biopharmaceuticals in the period of 2006–2014 suggests a vibrant sector, only half (51%) were genuinely new biopharmaceutical entities whilst the rest of the products approved were biosimilars, reformulated or me-too versions of previously approved substances (Walsh, 2010, 2014). As the market moves towards treatment for uncommon diseases and more effective drugs, more complex, difficult to express proteins such as bispecific antibodies and fusion proteins are being developed. Consequently, the industry is likely to be impeded by an imposing and growing obstacle resulting from the inability to rapidly and efficiently produce large quantities of recombinant protein at low cost.

The TGE technology platform has the advantage of short development times and low overall cost, and hence is actively pursued to produce a broad range of therapeutic MAbs, proteins and vaccines for preclinical studies. However, transient expression processes employ genetically/functionally diverse parental cell populations, whose intrinsic functional heterogeneity has not been utilised to generate a host cell clone intrinsically suited to the production of the recombinant product. As a result, transient production processes are generally low yielding, and the problems are exacerbated in the case of DTE recombinant proteins. Therefore, several strategies had been developed in this study to improve the transient production of DTE proteins in mammalian systems. The cellular production of model Sp35:Fc fusion protein was first systematically studied through a simple, rapid TGE assay and a mathematical model (Chapters 4 and 5), followed by various mitigation strategies, i.e. clone screening, controllable expression vector, and cell/process engineering (Chapters 6–8).

Data presented in this study support the notion that mammalian cells have limited secretory capacity, and that post-translational mechanisms are often limiting in transient systems (Mason et al., 2012) especially for DTE proteins (Pybus et al., 2014a). The simple and rapid TGE assay system enabled comparable measurement of the basic parameters that underpin product manufacturability, i.e. the relative ability of the host cell to synthesise and secrete a given recombinant product and to proliferate during the production process. In the case of DTE Sp35:Fc, it appears that a sweet-spot exists between the protein folding and aggregates formation to achieve a sub-optimal transgene expression. Even though we tested only two recombinant products in this work (i.e. Sp35:Fc vs. SEAP), the

assay is amenable to medium-high throughput operation and compatible with subsequent analyses of protein molecular variations such as aggregation or glycosylation. Similarly, whilst the intracellular behaviour modelled in this study was specific to Sp35:Fc produced in CHO cells, the model does offer a practical adjunct to other experimental studies. The iterative *in silico* engineering approach can be implemented as a generic framework for Fc-fusion protein productions, with the aim of using it as a platform to better understand the factors underlying cell factory processes, and subsequently propose predictive, rational strategies to overcome the limiting production factors in mammalian production systems.

The data of subclones of the parental CHO-S cell line showed the wide diversity and flexibility of metabolic capacities of CHO cells—a fact known for decades. Indeed, this know-how is used, intuitively, in the bioindustry by evaluating productivities and qualities of product in clonally derived cell populations. Our results provide a fundamental understanding for CHO host cell choice to express difficult to produce recombinant proteins, where the improvement in titre could be achieved either via increase in μ or qP, indicating variations in cellular machinery. However, analysis of the clones revealed that, unlike the production of characteristically easy-to-express GFP, Sp35:Fc productivity was negatively correlated to cell growth, which indicates that it is relatively difficult to obtain clones with improved DTE protein titres. Additionally, the fed-batch data in Chapter 8 demonstrate that the two best-performing clones responded in the same way to the cell/process engineering strategies applied to the parental cell population. Together, these data suggest clone selection is more appropriate for ETE proteins, whereas for DTE proteins cell/process engineering strategy is likely to be far more effective than a reliance on intrinsic clonal heterogeneity. Nevertheless, transcriptomics and proteomics (Charaniya et al., 2009; Dinnis et al., 2006) of the clones that displayed superior DTE protein production functionality could facilitate future employment of direct cell line engineering strategies (reverse engineering).

DTE proteins have mis-matched kinetics of transcription, translation and folding reactions, where overexpression of such proteins using strong promoters generally results in the induction of the UPR (and therefore reduced μ). Cell line engineering for enhanced DTE protein expression to improve cellular machinery is likely to require multi-gene modulation. Therefore, at the initial stage a vector engineering strategy might be more appropriate to achieve both high qP and IVCD, than gene by gene cell line engineering methods. The novel ERSE vectors employed in this study enabled a positive feedback loop within the system, thus enabling a shift from fast-growing cells to high-producing cells during the transient production of DTE Sp35:Fc in fed-batch culture. Although the ERSE-

SV40 vector achieved less than 2-fold increase in Sp35:Fc volumetric titre (and thus required a more effective engineering approach), this study was only the first step towards developing a mammalian controlled/dynamic expression system for DTE proteins and the system has various potential applications. Additionally, as cell/process engineering approach requires identification of a bespoke, protein-product specific strategy, we speculate that this system could be a more generic strategy to mitigate the UPR induction (low μ) associated with DTE recombinant proteins while having control over qP.

The results from cell and process engineering strategies demonstrate that it is possible to closely control the inverse relationship between μ and qP using either functional genes or chemicals. However, given the complexity of mammalian cellular regulation, targeting individual components of the secretory pathway discretely by overexpressing specific functional proteins may not always lead to the desired result (e.g. increase in Sp35:Fc aggregates), and as we have shown for Sp35:Fc, resulted in only modest improvements in μ or qP (<1.5-fold change). These results, however, are not entirely unexpected as our modelling work (Chapter 5, Figure 5.5) suggests that the translational attenuation mechanism played a vital role in maintaining cellular homeostasis, and with this mechanism in place, engineering the folding processes would yield no effects. Moreover, as our early diagnostic assay (Chapter 4, Figure 4.7) detected high levels of endogenous UPR-related proteins, it is also unsurprising that overexpression of these proteins did not lead to significant improvements in qP. This is in contrast to overexpression of ATF6c in SEAP producing cells (Chapter 7, Figure 7.3) which led to a 1.5-fold in SEAP titre.

On the other hand, chemical additives appeared to be very versatile where they were able to both significantly improve cell productivity and growth as well as suppress aggregates formation. Moreover, our work suggests that the use of small molecules should precede other development or engineering strategies for DTE proteins, especially when resources (i.e. time and money) are limited. In addition to having profound effects (Chapter 8, Figure 8.1) and easy implementation (less preparation/validation, easy to titrate, etc), chemical additives appeared to be largely cost-effective. With the exception of TMAO, the costs of chemicals at the highest concentrations used in this study were highly affordable—approximately between 0.2–18% of the cost of CD CHO culture medium ($104/L; Table 9.1). This would enable cost-effective manufacturing should the chemicals be used in large-scale transient/stable productions.

Table 9.1: Cost of different chemical additives per litre culture.

Chemical	Concentration	Cost/L culture ($)
PBA	1 mM	0.2
PERK inhibitor	40 nM	0.6
Glycerol	2% v/v	2.6
DMSO	2% v/v	7.7
Betaine	100 mM	19
TMAO	50 mM	113

Due to protein-product specificity, there is unlikely to be a generic solution which can be utilised for all recombinant proteins and cell lines. Our data also demonstrated that the optimal engineering intervention for a given DTE recombinant protein requires synergistic interaction between cell engineering and modulation of culture environment to alleviate cellular bottlenecks (and the associated poor expression) without compromising productive cell biomass accumulation. Nevertheless, the number of potentially useful product/process-specific effectors is likely to be very large. For this reason, a high-throughput, standardised platform to determine the optimal balance of functional effectors is highly desirable. Yet, it has to be based on combinatorial, context-dependent empirical modelling approaches to identify the optimal solution. In this regards, our study serves as a paradigm for multivariate optimisation in enabling host cells to achieve both high *q*P and IVCD, and therefore improved DTE protein production titres. We need to replace the current notion that some products are “difficult-to-express” with the concept that a “designer product” requires a “designer cell factory” that does not solely rely on screening natural biological variation for functionality.

9.2 Future Work

Based on work presented in this thesis, the direction of future work could follow several branches. One avenue to explore would be the transient production using a bioreactor under fully optimised conditions. The highest 40 mg/L yield of Sp35:Fc protein in this study was obtained in a small-scale, non-optimised feeding strategy and is still relatively low compared to characteristically easy-to-express protein productions. Large scale transient transfections can be carried out using polyplex-based methods where the detailed underlying mechanisms (Mozley et al., 2014), optimisation technique (Thompson et al., 2012) and large-scale protocols (Daramola et al., 2014) have been described in the literature. Moreover, one major limitation in this study was the progressive loss

recombinant plasmid over long-term culture. Therefore, it would also be useful to include the ERSE-SV40 vectors or other episomal-based expression system in the optimised system. With regard to the former, it could be worth studying whether co-expression of XBP1s (as a replacement or complement to ATF6c) could further improve the yield of the ERSE system. Although our data indicate that the production of Sp35:Fc is limited at post-translational processes, it may also be worthwhile to perform codon optimisation (Chung et al., 2013) on the Sp35:Fc gene to further enhance the expression. Larger volume and/or higher titre would eventually enable comprehensive analysis of product quality especially the *N*- and *O*-glycosylation patterns that presently cannot be tested using our methods due to insufficient quantities.

Building on data presented in Chapter 8 highlighting the various effects of functional gene co-expressions and chemical chaperone treatments, a future line of experimentation could employ a rapid, high-throughput DOE platform to assess various cell lines and recombinant proteins by testing for the optimal combination of multiple genetic/chemical components for a desired synthetic phenotype and appropriate culture processes. While the use of multiple plasmid vectors (i.e. high DNA loads) could result in high transfection toxicity, an alternate strategy for the expression of multiple genes could be through the use of multi-gene engineering technology (Kriz et al., 2010) or a bi-directional promoter system (Chatellard et al., 2007). Factorial designs can be employed to screen the factors affecting the responses and identify the basic interactions between the factors. Alternatively, if some of the factors are known to have little/no interactions or applied during different process

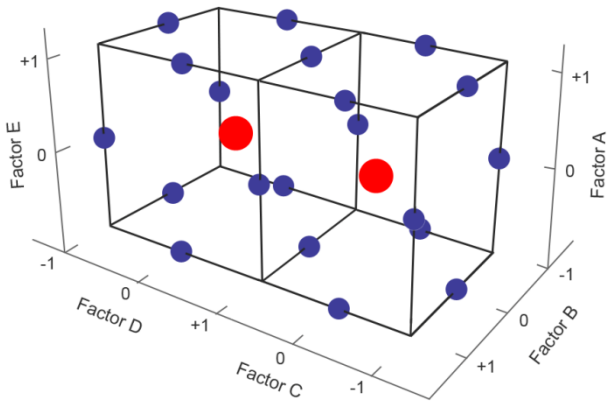

Figure 9.1: Parallel Box-Behnken designs can be used to obtain the optimal responses of 5 factors at a reduced number of experiments (21 design points). Factors A, B and C may represent co-expressions of CypB, XBP1s and BiP, while Factors D and E may represent additions of PBA and glycerol during stationary phase, respectively. The interaction among XBP1s (B), PBA (D) and glycerol (E) is also studied (but not among all 5 factors) as these 3 factors can induce apoptosis. A normal 3-level, 5-factor design would have 46 design points.

condition (e.g. during a biphasic culture), parallel response surface designs (Figure 9.1) can also be utilised. Such parallel technique could significantly reduce the number of experiments required and has been successfully applied to optimise the transient production of MAbs (Daramola et al., 2014) by studying the effects of PEI and DNA concentrations, cell density, culture media, and nutrient supplements. It is also useful to further explore the general and specific effects of different chemical chaperones on the cellular processes such as rDNA/mRNA stability, transcription, secretion and the UPR.

Promising drug candidates in preclinical studies will eventually require stable production of kilogram quantities for the subsequent testing phases and commercialisation. The mathematical model presented in Chapter 5 can potentially be developed as a predictive tool for stable performance from a transient expression platform (i.e., to aid cell line selection), especially for DTE proteins (Figure 9.2). Obviously, the selection of stable cell lines (or the transfectant pool) requires the host cell to combine both reasonable cell growth and recombinant protein expression. For DTE proteins, it seems likely that stable production cell lines would utilise cellular strategies to minimise/avoid UPR induction by decreasing recombinant protein synthesis. As transient production with increasing rDNA dose can determine the specific expression saturation point, further development of the existing model by coupling it to a measurement of cell proliferation and UPR induction may give it the capability to predict the likely specific production range of stable transfectants, and in the case of Sp35:Fc, whether the transfectants will produce aggregates.

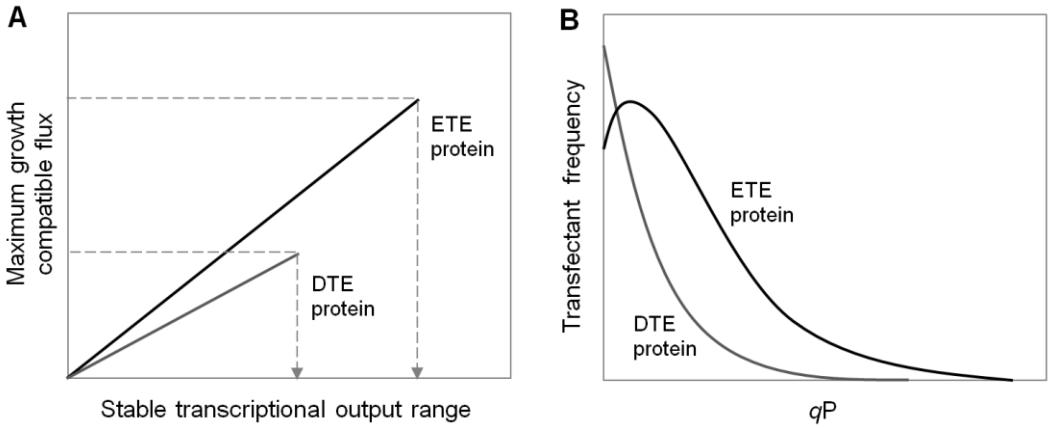

Figure 9.2: Prediction of stable performance from a transient expression platform. UPR induction inhibits cell growth, and stable cell lines effectively avoid UPR induction by decreasing recombinant protein synthesis. This leads to lower stable transcriptional output of DTE proteins compared to ETE proteins **(A)**. A linkage between cell proliferation and UPR induction measurements may grant the existing model a capability to predict the likely specific production range and distribution of stable transfectants **(B)**.

References

Ailor E, Betenbaugh MJ. 1998. Overexpression of a cytosolic chaperone to improve solubility and secretion of a recombinant IgG protein in insect cells. *Biotechnology and Bioengineering*, 58(2-3): 196-203.

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. *Journal of Cell Biology*, 168(3): 389-399.

Allen S, Naim HY, Bulleid NJ. 1995. Intracellular folding of tissue-type plasminogen activator. Effects of disulfide bond formation on N-linked glycosylation and secretion. *Journal of Biological Chemistry*, 270(9): 4797-4804.

Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J, Jirstrand M. 2014. Kinetic models in industrial biotechnology - Improving cell factory performance. *Metabolic Engineering*, 24: 38-60.

Alper H, Fischer C, Nevoigt E, Stephanopoulos G. 2005. Tuning genetic control through promoter engineering. *Proceedings of the National Academy of Sciences of the United States of America*, 102(36): 12678-12683.

Andersen DC, Bridges T, Gawlitzek M, Hoy C. 2000. Multiple cell culture factors can affect the glycosylation of Asn-184 in CHO-produced tissue-type plasminogen activator. *Biotechnology and Bioengineering*, 70(1): 25-31.

Anelli T, Sitia R. 2008. Protein quality control in the early secretory pathway. *EMBO Journal*, 27(2): 315-327.

Antoniou M, Harland L, Mustoe T, Williams S, Holdstock J, Yague E, Mulcahy T, Griffiths M, Edwards S, Ioannou PA, Mountain A, Crombie R. 2003. Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. *Genomics*, 82(3): 269-279.

Arakawa T, Kita Y, Ejima D, Tsumoto K, Fukada H. 2006. Aggregation suppression of proteins by arginine during thermal unfolding. *Protein and Peptide Letters*, 13(9): 921-927.

Arakawa T, Philo JS, Ejima D, Tsumoto K, Arisaka F. 2007. Aggregation analysis of therapeutic proteins, Part 2. Analytical ultracentrifugation and dynamic light scattering. *BioProcess International*, April: 36-47.

Astley K, Al-Rubeai M. 2008. The role of Bcl-2 and its combined effect with p21(CIP1) in adaptation of CHO cells to suspension and protein-free culture. *Applied Microbiology and Biotechnology*, 78(3): 391-399.

Backliwal G, Hildinger M, Chenuet S, Wulhfard S, De Jesus M, Wurm FM. 2008. Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. *Nucleic Acids Research*, 36(15): e96.

Bailey JE. 1998. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress*, 14(1): 8-20.

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. *Biotechnology Letters*, 29(5): 677-684.

Barnes LM, Bentley CM, Dickson AJ. 2004. Molecular definition of predictive indicators of stable protein expression in recombinant NS0 myeloma cells. *Biotechnology and Bioengineering*, 85(2): 115-121.

Barnes LM, Dickson AJ. 2006. Mammalian cell factories for efficient and stable protein expression. *Current Opinion in Biotechnology*, 17(4): 381-386.

Barnes LM, Moy N, Dickson AJ. 2006. Phenotypic variation during cloning procedures: Analysis of the growth behavior of clonal cell lines. *Biotechnology and Bioengineering*, 94(3): 530-537.

Baseman H. 2012. *Commentary on the US Food and Drug Administration's 2011 "Guidance for industry, process validation general principles and practices"*. In: Rathore AS, Sofer G, editors. *Process validation in manufacturing of biopharmaceuticals*. 3rd ed. Florida: Taylor & Francis Group. p 11-50.

Becerra S, Berrios J, Osses N, Altamirano C. 2012. Exploring the effect of mild hypothermia on CHO cell productivity. *Biochemical Engineering Journal*, 60: 1-8.

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. *Journal of Biotechnology*, 135(2): 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. *Journal of Biotechnology*, 146(4): 198-206.

Berg TM, Oyaas K, Levine DW. 1991. Betaine will protect hybridoma cells from hyperosmotic stress. *Biotechnology Techniques*, 5(3): 179-182.

Bernasconi R, Solda T, Galli C, Pertel T, Luban J, Molinari M. 2010. Cyclosporine A-sensitive, Cyclophilin B-dependent endoplasmic reticulum-associated degradation. *PloS One*, 5(9): e13008.

Bertschinger M, Burki C, Backliwal G, Hacker DL, Jordan M, Wurm FM. 2006. Polyethylenimine-based quality control assay for plasmid DNA. *Analytical Biochemistry*, 356(2): 309-311.

Bi JX, Shuttleworth J, Ai-Rubeai M. 2004. Uncoupling of cell growth and proliferation results in enhancement of productivity in p21(C1P1)-arrested CHO cells. *Biotechnology and Bioengineering*, 85(7): 741-749.

Bibila T, Flickinger MC. 1991. A structured model for monoclonal antibody synthesis in exponentially growing and stationary phase hybridoma cells. *Biotechnology and Bioengineering*, 37(3): 210-226.

Bibila TA, Flickinger MC. 1992. Use of a structured kinetic-model of antibody-synthesis and secretion for optimization of antibody-production systems: I. Steady-state analysis. *Biotechnology and Bioengineering*, 39(3): 262-272.

Bibila TA, Robinson DK. 1995. In pursuit of the optimal fed-batch process for monoclonal antibody production. *Biotechnology Progress*, 11(1): 1-13.

Birch JR, Boraston RC, Metcalfe H, Brown ME, Bebbington CR, Field RP. 1994. Selecting and designing cell lines for improved physiological characteristics. *Cytotechnology*, 15(1-3): 11-16.

Blount BA, Weenink T, Vasylechko S, Ellis T. 2012. Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. *PloS One,* 7(3): e33279.

Boland CR, Komarova NL, Goel A. 2009. Chromosomal instability and cancer: not just one CINgle mechanism. *Gut*, 58(2): 163-164.

Bollin F, Dechavanne V, Chevalet L. 2011. Design of Experiment in CHO and HEK transient transfection condition optimization. *Protein Expression and Purification*, 78(1): 61-68.

Bommiasamy H, Back SH, Fagone P, Lee K, Meshinchi S, Vink E, Sriburi R, Frank M, Jackowski S, Kaufman RJ, Brewer JW. 2009. ATF6α induces XBP1-independent expansion of the endoplasmic reticulum. *Journal of Cell Science*, 122(10): 1626-1636.

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. *Biotechnology Progress*, 21(1): 106-111.

Borys MC, Dalal NG, Abu-Absi NR, Khattak SF, Jing Y, Xing Z, Li ZJ. 2010. Effects of culture conditions on N-glycolylneuraminic acid (Neu5Gc) content of a recombinant fusion protein produced in CHO cells. *Biotechnology and Bioengineering*, 105(6): 1048-1057.

Brand HN, Froud SJ, Metcalfe HK, Onadipe AO, Shaw A, Westlake AJ. 1994. *Selection strategies for highly productive recombinant cell lines*. In: Spier RE, Griffiths JB, Berthold W, editors. *Animal cell technology: Products of today, prospects for tomorrow*. 3rd ed. Oxford: Butterworth-Heinemann Ltd. p 55-59.

Bravo R, Miguel Vicencio J, Parra V, Troncoso R, Pablo Munoz J, Bui M, Quiroga C, Rodriguez AE, Verdejo HE, Ferreira J, Iglewski M, Chiong M, Simmen T, Zorzano A, Hill JA, Rothermel BA, Szabadkai G, Lavandero S. 2011. Increased ER-mitochondrial coupling promotes mitochondrial respiration and bioenergetics during early phases of ER stress. *Journal of Cell Science*, 124(13): 2143-2152.

Brock A, Chang H, Huang S. 2009. Non-genetic heterogeneity—a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics*, 10(5): 336-342.

Brown AJ, Sweeney B, Mainwaring DO, James DC. 2014. Synthetic promoters for CHO cell engineering. *Biotechnology and Bioengineering*, 111(8): 1638-1647.

Brown HC, Gangadharan B, Doering CB. 2011. Enhanced biosynthesis of coagulation factor VIII through diminished engagement of the unfolded protein response. *Journal of Biological Chemistry*, 286(27): 24451-24457.

Brown ME, Renner G, Field RP, Hassell T. 1992. Process-development for the production of recombinant antibodies using the glutamine-synthetase (GS) system. *Cytotechnology*, 9(1-3): 231-236.

Browne SM, Al-Rubeai M. 2007. Selection methods for high-producing mammalian cell lines. *Trends in Biotechnology*, 25(9): 425-432.

Browne SM, Al-Rubeai M. 2011. Analysis of an artificially selected GS-NS0 variant with increased resistance to apoptosis. *Biotechnology and Bioengineering*, 108(4): 880-892.

Butler M. 2006. Optimisation of the cellular metabolism of glycosylation for recombinant proteins produced by mammalian cell systems. *Cytotechnology*, 50(1-3): 57-76.

Bulter T, Lee SG, Woirl WWC, Fung E, Connor MR, Liao JC. 2004. Design of artificial cell-cell communication using gene and metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(8): 2299-2304.

Butler M, Meneses-Acosta A. 2012. Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Applied Microbiology and Biotechnology*, 96(4): 885-894.

Cabra V, Vazquez-Contreras E, Moreno A, Arreguin-Espinosa R. 2008. The effect of sulfhydryl groups and disulphide linkage in the thermal aggregation of Z19 α-zein. *Biochimica Et Biophysica Acta-Proteins and Proteomics*, 1784(7-8): 1028-1036.

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-Lα has increased levels of transient protein expression. *Biotechnology Progress*, 29(3): 697-706.

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. *Molecular Biotechnology*, 39(2): 155-158.

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

Carpentier E, Paris S, Kamen AA, Durocher Y. 2007. Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *Journal of Biotechnology*, 128(2): 268-280.

Chaderjian WB, Chin ET, Harris RJ, Etcheverry TM. 2005. Effect of copper sulfate on performance of a serum-free CHO cell culture process and the level of free thiol in the recombinant antibody expressed. *Biotechnology Progress*, 21(2): 550-553.

Chakrabarti A, Chen AW, Varner JD. 2011. A review of the mammalian unfolded protein response. *Biotechnology and Bioengineering*, 108(12): 2777-2793.

Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. 2008. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*, 453(7194): 544-U10.

Charaniya S, Karypis G, Hu W-S. 2009. Mining transcriptome data for function-trait relationship of hyper productivity of recombinant antibody. *Biotechnology and Bioengineering*, 102(6): 1654-1669.

Chatellard P, Pankiewicz R, Meier E, Durrer L, Sauvage C, Imhof MO. 2007. The IE2 promoter/enhancer region from mouse CMV provides high levels of therapeutic protein expression in mammalian cells. *Biotechnology and Bioengineering*, 96(1): 106-117.

Chen F, Kou T, Fan L, Zhou Y, Ye Z, Zhao L, Tan W-S. 2011. The combined effect of sodium butyrate and low culture temperature on the production, sialylation, and biological activity of an antibody produced in CHO cells. *Biotechnology and Bioprocess Engineering*, 16(6): 1157-1165.

Chiang GG, Sisk WP. 2005. Bcl-x_L mediates increased production of humanized monoclonal antibodies in Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 91(7): 779-792.

Chun C, Heineken K, Szeto D, Ryll T, Chamow S, Chung JD. 2003. Application of factorial design to accelerate identification of CHO growth factor requirements. *Biotechnology Progress*, 19(1): 52-57.

Chung BK, Yusufi FN, Mariati, Yang Y, Lee DY. 2013. Enhanced expression of codon optimized interferon gamma in CHO cells. *Journal of Biotechnology*, 167(3): 326-333.

Chung JY, Lim SW, Hong YJ, Hwang SO, Lee GM. 2004. Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 85(5): 539-546.

Clarke J, Porter A, Thorpe R, Davis J. 2002. *Cloning*. In: Davis J, editor. *Basic cell culture: A practical approach*. 2nd ed. Oxford: Oxford University Press.

Codamo J, Hou JJC, Hughes BS, Gray PP, Munro TP. 2011a. Efficient mAb production in CHO cells incorporating PEI-mediated transfection, mild hypothermia and the co-expression of XBP-1. *Journal of Chemical Technology and Biotechnology*, 86(7): 923-934.

Cohen RN, van der Aa MAEM, Macaraeg N, Lee AP, Szoka FC, Jr. 2009. Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *Journal of Controlled Release*, 135(2): 166-174.

Comley, J. 2009. *Design of experiments: Useful statistical tool in assay development or vendor disconnect!* Drug Discovery World, vol. Winter 2009/10, pp. 41-50. www.ddw-online.com/media/32/2273/winter-2009-design-of-experiments.pdf.

Condon RGG, Schaefer EJ, Santoro M, Longley R, Tsao YS, Zurawski SM, Liu Z. 2003. Development of a Chinese hamster ovary cell line for recombinant adenovirus-mediated gene expression. *Biotechnology Progress*, 19(1): 137-143.

Cordoba-Rodriguez RV. 2008. Aggregates in MAbs and recombinant therapeutic proteins: A regulatory perspective. *BioPharm International* 21(11): 44.

Cromwell MEM, Hilario E, Jacobson F. 2006. Protein aggregation and bioprocessing. *The AAPS Journal*, 8(3): E572-E579.

Dahm T, White J, Grill S, Fullekrug J, Stelzer EHK. 2001. Quantitative ER ↔ Golgi transport kinetics and protein separation upon Golgi exit revealed by vesicular integral membrane protein 36 dynamics in live cells. *Molecular Biology of the Cell*, 12(5): 1481-1498.

Dalton AC, Barton WA. 2014. Over-expression of secreted proteins from mammalian cell lines. *Protein Science*, 23(5): 517-525.

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. *Biotechnology Progress*, 30(1): 132-141.

Dasgupta S, Navarrete A-M, Bayry J, Delignat S, Wootla B, Andre S, Christophe O, Nascimbeni M, Jacquemin M, Martinez-Pomares L, Geijtenbeek TB, Moris A, Saint-Remy JM, Kazatchkine MD, Kaveri SV, Lacroix-Desmazes S. 2007. A role for exposed mannosylations in presentation of human therapeutic self-proteins to CD4+T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21): 8965-8970.

Davies SL, Lovelady CS, Grainger RK, Racher AJ, Young RJ, James DC. 2013. Functional heterogeneity and heritability in CHO cell populations. *Biotechnology and Bioengineering*, 110(1): 260-274.

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. *Biotechnology Progress*, 27(6): 1689-1699.

Davis R, Schooley K, Rasmussen B, Thomas J, Reddy P. 2000. Effect of PDI overexpression on recombinant protein secretion in CHO cells. *Biotechnology Progress*, 16(5): 736-743.

de Almeida SF, Picarote G, Fleming JV, Carmo-Fonseca M, Azevedo JE, de Sousa M. 2007. Chemical chaperones reduce endoplasmic reticulum stress and prevent mutant HFE aggregate formation. *Journal of Biological Chemistry*, 282(38): 27905-27912.

De Jesus M, Wurm FM. 2011. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *European Journal of Pharmaceutics and Biopharmaceutics*, 78(2): 184-188.

Delic M, Gongrich R, Mattanovich D, Gasser B. 2014. Engineering of protein folding and secretion-strategies to overcome bottlenecks for efficient production of recombinant proteins. *Antioxidants & Redox Signaling*, 21(3): 414-37.

DePalma A. 2012. Taming difficult to express proteins. *Genetic Engineering & Biotechnology News*, 32(12).

Derouazi M, Martinet D, Schmutz NB, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM. 2006. Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. *Biochemical and Biophysical Research Communications*, 340(4): 1069-1077.

Dey M, Cao C, Dar AC, Tamura T, Ozato K, Sicheri F, Dever TE. 2005. Mechanistic link between PKR dimerization, autophosphorylation, and eIF2α substrate recognition. *Cell*, 122(6): 901-913.

deZengotita VM, Miller WM, Aunins JG, Zhou WC. 2000. Phosphate feeding improves high-cell-concentration NS0 myeloma culture performance for monoclonal antibody production. *Biotechnology and Bioengineering*, 69(5): 566-576.

Dhir S, Morrow KJ, Rhinehart RR, Wiesner T. 2000. Dynamic optimization of hybridoma growth in a fed-batch bioreactor. *Biotechnology and Bioengineering*, 67(2): 197-205.

Dias-Gunasekara S, Gubbens J, van Lith M, Dunne C, Williams JAG, Kataky R, Scoones D, Lapthorn A, Bulleid NJ, Benham AM. 2005. Tissue-specific expression and dimerization of the endoplasmic reticulum oxidoreductase Ero1. *Journal of Biological Chemistry*, 280(38): 33066-33075.

Dietmair S, Nielsen LK, Timmins NE. 2011. Engineering a mammalian super producer. *Journal of Chemical Technology and Biotechnology*, 86(7): 905-914.

Dimelow RJ, Wilkinson SJ. 2009. Control of translation initiation: A model-based analysis from limited experimental data. *Journal of the Royal Society Interface*, 6(30): 51-61.

Dimitrov DS. 2012. Therapeutic proteins. *Methods in Molecular Biology*, 899: 1-26.

Dinnis DM, James DC. 2005. Engineering mammalian cell factories for improved recombinant monoclonal antibody production: Lessons from nature? *Biotechnology and Bioengineering*, 91(2): 180-189.

Dinnis DM, Stansfield SH, Schlatter S, Smales CM, Alete D, Birch JR, Racher AJ, Marshal CT, Nielsen LK, James DC. 2006. Functional proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnology and Bioengineering*, 94(5): 830-841.

Dobson CM. 2003. Protein folding and misfolding. *Nature*, 426(6968): 884-890.

Dorner AJ, Kaufman RJ. 1994. The levels of endoplasmic-reticulum proteins and ATP affect folding and secretion of selective proteins. *Biologicals*, 22(2): 103-112.

Dorner AJ, Krane MG, Kaufman RJ. 1988. Reduction of endogenous GRP78 levels improves secretion of a heterologous protein in CHO cells. *Molecular and Cellular Biology*, 8(10): 4063-4070.

Dorner AJ, Wasley LC, Kaufman RJ. 1989. Increased synthesis of secreted proteins induces expression of glucose-related proteins in butyrate-treated Chinese hamster ovary cells. *Journal of Biological Chemistry*, 264(34): 20602-20607.

Dorner AJ, Wasley LC, Kaufman RJ. 1992. Overexpression of GRP78 mitigates stress induction of glucose regulated proteins and blocks secretion of selective proteins in Chinese hamster ovary cells. *EMBO Journal*, 11(4): 1563-1571.

Dreesen IAJ, Fussenegger M. 2011. Ectopic expression of human mTOR increases viability, robustness, cell size, proliferation, and antibody production of Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 108(4): 853-866.

Du Z, Treiber D, McCoy RE, Miller AK, Han M, He F, Domnitz S, Heath C, Reddy P. 2013. Non-invasive UPR monitoring system and its applications in CHO production cultures. *Biotechnology and Bioengineering*, 110(8): 2184-2194.

Durocher Y, Butler M. 2009. Expression systems for therapeutic glycoprotein production. *Current Opinion in Biotechnology*, 20(6): 700-707.

DuRose JB, Scheuner D, Kaufman RJ, Rothblum LI, Niwa M. 2009. Phosphorylation of eukaryotic translation initiation factor 2 α coordinates rRNA transcription and

translation inhibition during endoplasmic reticulum stress. *Molecular and Cellular Biology*, 29(15): 4295-4307.
Estes B, Hsu Y-R, Tam L-T, 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. *Biotechnology Progress*, 31(1): 258-267.
Estes S, Melville M. 2014. Mammalian cell line developments in speed and efficiency. *Advances in Biochemical Engineering/Biotechnology*, 139: 11-33.
Farzadfard F, Perli SD, Lu TK. 2013. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. *ACS Synthetic Biology*, 2(10): 604-613.
Fast JL, Cordes AA, Carpenter JF, Randolph TW. 2009. Physical instability of a therapeutic Fc fusion protein: Domain contributions to conformational and colloidal stability. *Biochemistry*, 48(49): 11724-11736.
FDA. 2004. *Guidance for industry: PAT - A framework for innovative pharmaceutical development, manufacturing, and quality assurance*. www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf
Feige MJ, Groscurth S, Marcinowski M, Shimizu Y, Kessler H, Hendershot LM, Buchner J. 2009. An unfolded C_H1 domain controls the assembly and secretion of IgG antibodies. *Molecular Cell*, 34(5): 569-579.
Fischer S, Charara N, Gerber A, Woelfel J, Schiedner G, Voedisch B, Geisse S. 2012. Transient recombinant protein expression in a human amniocyte cell line: The CAP-T® cell system. *Biotechnology and Bioengineering*, 109(9): 2250-2261.
Foecking MK, Hofstetter H. 1986. Powerful and versatile enhancer-promoter unit for mammalian expression vectors. *Gene*, 45(1): 101-105.
Follstad BD, Potter AH. 2006. *Cell culture performance with betaine*. United States patent US07067279B1.
Fox SR, Tan HK, Tan MC, Wong S, Yap MGS, Wang DIC. 2005. A detailed understanding of the enhanced hypothermic productivity of interferon-γ by Chinese-hamster ovary cells. *Biotechnology and Applied Biochemistry*, 41: 255-264.
Franceschini G, Macchietto S. 2008. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19): 4846-4872.
Fratantoni JC, Dzekunov S, Singh V, Liu LN. 2003. A non-viral gene delivery system designed for clinical use. *Cytotherapy*, 5(3): 208-210.
Freimark D, Jerome V, Freitag R. 2010. A GFP-based method facilitates clonal selection of transfected CHO cells. *Biotechnology Journal*, 5(1): 24-31.
Fussenegger M, Schlatter S, Datwyler D, Mazur X, Bailey JE. 1998. Controlled proliferation by multigene metabolic engineering enhances the productivity of Chinese hamster ovary cells. *Nature Biotechnology*, 16(5): 468-472.
Gawron K, Jensen DA, Steplewski A, Fertala A. 2010. Reducing the effects of intracellular accumulation of thermolabile collagen II mutants by increasing their thermostability in cell culture conditions. *Biochemical and Biophysical Research Communications*, 396(2): 213-218.
Geisse S. 2009. Reflections on more than 10 years of TGE approaches. *Protein Expression and Purification*, 64(2): 99-107.
Gomez N, Subramanian J, Ouyang J, Nguyen MDH, Hutchinson M, Sharma VK, Lin AA, Yuk IH. 2012. Culture temperature modulates aggregation of recombinant antibody in CHO cells. *Biotechnology and Bioengineering*, 109(1): 125-136.
Goto Y, Hamaguchi K. 1981. Formation of the intrachain disulfide bond in the constant fragment of the immunoglobulin light chain. *Journal of Molecular Biology*, 146(3): 321-340.
Grainger RK, James DC. 2013. CHO cell line specific prediction and control of recombinant monoclonal antibody *N*-glycosylation. *Biotechnology and Bioengineering*, 110(11): 2970-2983.
Grubb S, Guo L, Fisher EA, Brodsky JL. 2012. Protein disulfide isomerases contribute differentially to the endoplasmic reticulum-associated degradation of apolipoprotein B and other substrates. *Molecular Biology of the Cell*, 23(4): 520-532.
Guan B, Krokowski D, Majumder M, Schmotzer CL, Kimball SR, Merrick WC, Koromilas AE, Hatzoglou M. 2014. Translational control during endoplasmic reticulum stress beyond

phosphorylation of the translation initiation factor eIF2α. *Journal of Biological Chemistry*, 289(18): 12593-12611.
Haas AK, Mayer K, Brinkmann U. 2012. Generation of fluorescent IgG fusion proteins in mammalian cells. *Methods in Molecular Biology*, 901: 265-276.
Hacker DL, Nallet S, Wurm FM. 2008. Recombinant protein production yields from mammalian cells: Past, present, and future. *Biopharm International Supplements*, June 2008.
Hall JP. 2007. Applying fusion protein technology to *E. coli*. *BioPharm International*, May 2007.
Han AP, Yu C, Lu L, Fujiwara Y, Browne C, Chin G, FlemingM, Leboulch P, Orkin SH, Chen JJ. 2001. Heme-regulated eIF2α kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. *EMBO Journal*, 20(23): 6909-6918.
Han YK, Ha TK, Kim Y-G, Lee GM. 2011. Bcl-x_L overexpression delays the onset of autophagy and apoptosis in hyperosmotic recombinant Chinese hamster ovary cell cultures. *Journal of Biotechnology*, 156(1): 52-55.
Hannemann A, Christie JK, Flatman PW. 2009. Functional expression of the Na-K-2Cl cotransporter NKCC2 in mammalian cells fails to confirm the dominant-negative effect of the AF splice variant. *Journal of Biological Chemistry*, 284(51): 35348-35358.
Harding HP, Zhang YH, Zeng HQ, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, Stojdl DF, Bell JC, Hettmann T, Leiden JM, Ron D. 2003. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Molecular Cell*, 11(3): 619-633.
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(6): 993-1002.
Hartman TE, Sar N, Genereux K, Barritt DS, He Y, Burky JE, Wesson MC, Tso JY, Tsurushita N, Zhou W, Sauer PW. 2007. Derivation and characterization of cholesterol-independent non-GS NS0 cell lines for production of recombinant antibodies. *Biotechnology and Bioengineering*, 96(2): 294-306.
Hasegawa H, Wendling J, He F, Trilisky E, Stevenson R, Franey H, Kinderman F, Li G, Piedmonte DM, Osslund T, Shen M, Ketchem RR. 2011. *In vivo* crystallization of human IgG in the endoplasmic reticulum of engineered Chinese hamster ovary (CHO) cells. *Journal of Biological Chemistry*, 286(22): 19917-19931.
Hatahet F, Ruddock LW. 2009 Protein disulfide isomerase: A critical evaluation of its function in disulfide bond formation. *Antioxidants & Redox Signaling*, 11(11): 2807-2850.
Hatzimanikatis V, Choe LH, Lee KH. 1999. Proteomics: Theoretical and experimental considerations. *Biotechnology Progress*, 15(3): 312-318.
Hayes NVL, Smales CM, Klappa P. 2010. Protein disulfide isomerase does not control recombinant IgG4 productivity in mammalian cell lines. *Biotechnology and Bioengineering*, 105(4): 770-779.
Heller-Harrison R, Crowe K, Cooley C, Hone M, McCarthy K, Leonard M. 2009. Managing cell line instability and its impact during cell line development. *BioPharm International Supplements*, June 2009: 16-27.
Hernández Bort JA, Stern B, Borth N. 2010. CHO-K1 host cells adapted to growth in glutamine-free medium by FACS-assisted evolution. *Biotechnology Journal*, 5(10): 1090-1097.
Ho Y, Kiparissides A, Pistikopoulos EN, Mantalaris A. 2012. Computational approach for understanding and improving GS-NS0 antibody production under hyperosmotic conditions. *Journal of Bioscience and Bioengineering*, 113(1): 88-98.
Ho Y, Varley J, Mantalaris A. 2006. Development and analysis of a mathematical model for antibody-producing GS-NS0 cells under normal and hyperosmotic culture conditions. *Biotechnology Progress*, 22(6): 560-1569.
Hosokawa N, Tremblay LO, You ZP, Herscovics A, Wada I, Nagata K. 2003. Enhancement of endoplasmic reticulum (ER) degradation of misfolded null Hong Kong α_1-antitrypsin by human ER mannosidase I. *Journal of Biological Chemistry*, 278(28): 26287-26294.
Hossler P, Khattak SF, Li ZJ. 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology*, 19(9): 936-949.

Hsu TA, Betenbaugh MJ. 1997. Coexpression of molecular chaperone BiP improves immunoglobulin solubility and IgG secretion from *Trichoplusia ni* insect cells. *Biotechnology Progress*, 13(1): 96-104.

Hu Z, Guo D, Yip SSM, Zhan D, Misaghi S, Joly JC, Snedecor BR, Shen AY. 2013. Chinese hamster ovary K1 host cell enables stable cell line development for antibody molecules which are difficult to express in DUXB11-derived dihydrofolate reductase deficient host cell. *Biotechnology Progress*, 29(4): 980-985.

Huang S. 2009. Non-genetic heterogeneity of cells in development: more than just noise. *Development*, 136(23): 3853-3862.

Huang Y, 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. *Biotechnology Progress*, 26(5): 1400-1410.

Huggett B. 2013. Public biotech 2012—the numbers. *Nature Biotechnology*, 31(8): 697-703.

Huliak I, Sike A, Zencir S, Boros IM. 2012. The objectivity of reporters: Interference between physically unlinked promoters affects reporter gene expression in transient transfection experiments. *DNA and Cell Biology*, 31(11): 1580-1584.

Hunt L, Batard P, Jordan M, Wurm FM. 2002. Fluorescent proteins in animal cells for process development: Optimization of sodium butyrate treatment as an example. *Biotechnology and Bioengineering*, 77(5): 528-537.

Hurley JM, Chen C-H, Loros JJ, Dunlap JC. 2012. Light-inducible system for tunable protein expression in *Neurospora crassa*. *G3 (Bethesda)*, 2(10): 1207-1212.

Hwang S-J, Jeon C-J, Cho SM, Lee GM, Yoon SK. 2011a. Effect of chemical chaperone addition on production and aggregation of recombinant flag-tagged COMP-Angiopoietin 1 in Chinese hamster ovary cells. *Biotechnology Progress*, 27(2): 587-591.

Hwang S-J, Yoon SK, Koh GY, Lee GM. 2011b. Effects of culture temperature and pH on flag-tagged COMP angiopoietin-1 (FCA1) production from recombinant CHO cells: FCA1 aggregation. *Applied Microbiology and Biotechnology*, 91(2): 305-315.

Igawa T, Tsunoda H, Kuramochi T, Sampei Z, Ishii S, Hattori K. 2011. Engineering the variable region of therapeutic IgG antibodies. *mAbs*, 3(3): 243-252.

Ignatova Z, Gierasch LM. 2006. Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. *Proceedings of the National Academy of Sciences of the United States of America*, 103(36): 13357-13361.

Ioannou YA, Bishop DF, Desnick RJ. 1992. Overexpression of human alpha-galactosidase A results in its intracellular aggregation, crystallization in lysosomes, and selective secretion. *Journal of Cell Biology*, 119(5): 1137-1150.

Jang DJ, Barford JP. 2000. An unstructured kinetic model of macromolecular metabolism in batch and fed-batch cultures of hybridoma cells producing monoclonal antibody. *Biochemical Engineering Journal*, 4(2): 153-168.

Jaqaman K, Danuser G. 2006. Linking data to models: data regression. *Nature Reviews Molecular Cell Biology*, 7(11): 813-819.

Jiang Z, Droms K, Geng Z, Casnocha S, Xiao Z, Gorfrien S, Jacobia S. 2012. Fed-batch cell culture process optimization: A rationally integrated approach. *BioProcess Technical*, 10: 40-45.

Jiang Z, Sharfstein ST. 2008. Sodium butyrate stimulates monoclonal antibody over-expression in CHO cells by improving gene accessibility. *Biotechnology and Bioengineering*, 100(1): 189-194.

Jin F, Kretschmer PJ, Harkins RN, Hermiston TW. 2010. Enhanced protein production using HBV X protein (HBx), and synergy when used in combination with XBP1s in BHK21 cells. *Biotechnology and Bioengineering*, 105(2): 341-349.

Jobbagy Z, Ward JL, Toan SV, Leung GPH, Tse CM. 2002. One-step unidirectional cloning of tandem repeats of DNA fragments: An application for fusion protein production. *Analytical Biochemistry*, 303(1): 104-107.

Johnston SL. 2007. Biologic therapies: what and when? *Journal of Clinical Pathology*, 60(1): 8-17.

Jones J, Nivitchanyong T, Giblin C, Ciccarone V, Judd D, Gorfien S, Krag SS, Betenbaugh MJ. 2005. Optimization of tetracycline-responsive recombinant protein production and

effect on cell growth and ER stress in mammalian cells. *Biotechnology and Bioengineering*, 91(6): 722-732.
Jones PA, Baylin SB. 2007. The epigenomics of cancer. *Cell*, 128(4): 683-692.
Jostock T, Dragic Z, Fang J, Jooss K, Wilms B, Knopf H. 2010. Combination of the 2A/furin technology with an animal component free cell line development platform process. *Applied Microbiology and Biotechnology*, 87(4): 1517-1524.
Kalwy S, Rance J, Young R. 2006. Toward more efficient protein expression. *Molecular Biotechnology*, 34(2): 151-156.
Kaneko Y, Sato R, Aoyagi H. 2010. Evaluation of Chinese hamster ovary cell stability during repeated batch culture for large-scale antibody production. *Journal of Bioscience and Bioengineering*, 109(3): 274-280.
Karra S, Sager B, Karim MN. 2010. Multi-scale modeling of heterogeneities in mammalian cell culture processes. *Industrial & Engineering Chemistry Research*, 49(17): 7990-8006.
Kaufman RJ, Sharp PA, Latt SA. 1983. Evolution of chromosomal regions containing transfected and amplified dihydrofolate-reductase sequences. *Molecular and Cellular Biology*, 3(4): 699-711.
Keightley JA, Shang L, Kinter M. 2004. Proteomic analysis of oxidative stress-resistant cells: A specific role for aldose reductase overexpression in cytoprotection. *Molecular & Cellular Proteomics*, 3(2): 167-175.
Kennard ML, Goosney DL, Monteith D, Roe S, Fischer D, Mott J. 2009. Auditioning of CHO host cell lines using the artificial chromosome expression (ACE) technology. *Biotechnology and Bioengineering*, 104(3): 526-539.
Kim D, Lee YJ. 1993. Effect of glycerol on protein aggregation: Quantitation of thermal aggregation of proteins from CHO cells and analysis of aggregated proteins. *Journal of Thermal Biology*, 18(1): 41-48.
Kim H, Bae S. 2011. Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *American Journal of Translational Research*, 3(2): 166-179.
Kim J, Choi TG, Ding Y, Kim Y, Ha KS, Lee KH, Kang I, Ha J, Kaufman RJ, Lee J, Choe W, Kim SS. 2008. Overexpressed cyclophilin B suppresses apoptosis associated with ROS and Ca^{2+} homeostasis after ER stress. *Journal of Cell Science*, 121(21): 3636-3648.
Kim M, O'Callaghan PM, Droms KA, James DC. 2011a. A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnology and Bioengineering*, 108(10): 2434-2446.
Kim NS, Byun TH, Lee GM. 2001. Key determinants in the occurrence of clonal variation in humanized antibody expression of CHO cells during dihydrofolate reductase mediated gene amplification. *Biotechnology Progress*, 17(1): 69-75.
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. *Biotechnology and Bioengineering*, 71(3): 184-193.
Kim SH, Lee GM. 2009. Development of serum-free medium supplemented with hydrolysates for the production of therapeutic antibodies in CHO cell cultures using design of experiments. *Applied Microbiology and Biotechnology*, 83(4): 639-648.
Kim SJ, Lee GM. 1999. Cytogenetic analysis of chimeric antibody-producing CHO cells in the course of dihydrofolate reductase-mediated gene amplification and their stability in the absence of selective pressure. *Biotechnology and Bioengineering*, 64(6): 41-749.
Kim TK, Eberwine JH. 2010. Mammalian cell transfection: the present and the future. *Analytical and Bioanalytical Chemistry*, 397(8): 3173-3178.
Kim TK, Ryu JS, Chung JY, Kim MS, Lee GM. 2000. Osmoprotective effect of glycine betaine on thrombopoietin production in hyperosmotic Chinese hamster ovary cell culture: Clonal variations. *Biotechnology Progress*, 16(5): 775-781.
Kim Y, Kim JY, Park B, Ahn JO, Jung J, Lee HW, Lee GM, Lee EG. 2011b. Effect of Bcl-xL overexpression on erythropoietin production in recombinant Chinese hamster ovary cells treated with dimethyl sulfoxide. *Process Biochemistry*, 46(11): 2201-2204.
Kim Y, Park B, Ahn JO, Jung J, Lee HW, Lee EG. 2012. New cell line development for antibody-producing Chinese hamster ovary cells using split green fluorescent protein. *BMC Biotechnology*, 12: 24

Kiparissides A, Koutinas M, Kontoravdi C, Mantalaris A, Pistikopoulos EN. 2011. 'Closing the loop' in biological systems modeling - From the *in silico* to the *in vitro*. *Automatica*, 47(6): 1147-1155.

Kitchin K, Flickinger MC. 1995. Alteration of hybridoma viability and antibody secretion in transfectomas with inducible overexpression of protein disulfide-isomerase. *Biotechnology Progress*, 11(5): 565-574.

Kokame K, Kato H, Miyata T. 2001. Identification of ERSE-II, a new cis-actin element responsible for the ATF6-dependent mammalian unfolded protein response. *Journal of Biological Chemistry*, 276(12): 9199-9205.

Kontermann RE. 2011. Strategies for extended serum half-life of protein therapeutics. *Current Opinion in Biotechnology*, 22(6): 868-876.

Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. 2005. Application of global sensitivity analysis to determine goals for design of experiments: An example study on antibody-producing cell cultures. *Biotechnology Progress*, 21(4): 1128-1135.

Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. 2007. Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring. *Computers & Chemical Engineering*, 31(5-6): 392-400.

Kontoravdi C, Pistikopoulos EN, Mantalaris A. 2010. Systematic development of predictive mathematical models for animal cell cultures. *Computers & Chemical Engineering*, 34(8): 1192-1198.

Kopito RR. 2000. Aggresomes, inclusion bodies and protein aggregation. *Trends in Cell Biology*, 10(12): 524-530.

Kou T, Fan L, Zhou Y, Ye Z, Liu X, Zhao L, Tan W. 2011. Detailed understanding of enhanced specific productivity in Chinese hamster ovary cells at low culture temperature. *Journal of Bioscience and Bioengineering*, 111(3): 365-369.

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. *Nature Communications*, 1: 120.

Kromenaker SJ, Srienc F. 1994. Stability of producer hybridoma cell-lines after cell sorting: A case study. *Biotechnology Progress*, 10(3): 299-307.

Ku SCY, Ng DTW, Yap MGS, Chao S-H. 2008. Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. *Biotechnology and Bioengineering*, 99(1): 155-164.

Kubota K, Niinuma Y, Kaneko M, Okuma Y, Sugai M, Omura T, Uesugi M, Uehara T, Hosoi T, Nomura Y. 2006. Suppressive effects of 4-phenylbutyrate on the aggregation of Pael receptors and endoplasmic reticulum stress. *Journal of Neurochemistry*, 97(5): 1259-1268.

Kumar N, Gammell P, Meleady P, Henry, Clynes M. 2008. Differential protein expression following low temperature culture of suspension CHO-K1 cells. *BMC Biotechnology*, 8: 42.

Kunaparaju R, Liao M, Sunstrom NA. 2005. *Epi*-CHO, an episomal expression system for recombinant protein production in CHO cells. *Biotechnology and Bioengineering*, 91(6): 670-677.

Kuryatov A, Mukherjee J, Lindstrom J. 2013. Chemical chaperones exceed the chaperone effects of RIC-3 in promoting assembly of functional α7 AChRs. *PloS One*, 8(4): e62246.

Kuttler F, Mai S. 2006. c-Myc, genomic instability and disease. *Genome Dynamics*, 1: 171-190.

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

Lang S, Benedix J, Fedeles SV, Schorr S, Schirra C, Schäuble 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 Sec61α, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *Journal of Cell Science*, 125(8): 1958-1969.

Le Fourn V, Girod P, Buceta M, Regamey A, Mermod N. 2014. CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. *Metabolic Engineering*, 21: 91-102.

Lecker SH, Goldberg AL, Mitch WE. 2006. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *Journal of the American Society of Nephrology,* 17(7): 1807-1819.

Lee AH, Iwakoshi NN, Glimcher LH. 2003. XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. *Molecular and Cellular Biology*, 23(21): 7448-7459.

Lee GM, Varma A, Palsson BO. 1991. Application of population balance model to the loss of hybridoma antibody productivity. *Biotechnology Progress*, 7(1): 72-75.

Lee GW, Fecko JK, Yen A, Donaldson D, Wood C, Tobler S, Vunuum S, Luo Y, Leonard M. 2007. *Improving the expression of a soluble receptor:Fc fusion protein in CHO cells by coexpression with the receptor ligand*. In: Smith R, editor. *Cell technology for cell products*. Dordrecht: Springer Netherlands. p 29-39.

Lee M, Diamond ME, Yates JL. 1999. Genetic evidence that EBNA-1 is needed for efficient, stable latent infection by Epstein-Barr virus. *Journal of Virology*, 73(4): 2974-2982.

Legmann R, Schreyer HB, Combs RG, McCormick EL, Russo AP, Rodgers ST. 2009. A predictive high-throughput scale-down model of monoclonal antibody production in CHO cells. *Biotechnology and Bioengineering*, 104(6): 1107-1120.

Levine B, Sinha S, Kroemer G. 2008. Bcl-2 family members: Dual regulators of apoptosis and autophagy. *Autophagy*, 4(5): 600-606.

Lilie H, McLaughlin S, Freedman R, Buchner J. 1994. Influence of protein disulfide isomerase (PDI) on antibody folding *in vitro*. *Journal of Biological Chemistry*, 269(19): 14290-14296.

Lim Y, Wong NSC, Lee YY, Ku SCY, Wong DCF, Yap MGS. 2010. Engineering mammalian cells in bioprocessing - current achievements and future perspectives. *Biotechnology and Applied Biochemistry*, 55: 175-189.

Little RD, Schildkraut CL. 1995. Initiation of latent DNA replication in the Epstein-Barr virus genome can occur at sites other than the genetically defined origin. *Molecular and Cellular Biology*, 15(5): 2893-2903.

Liu C, Chen L. 2007. Enhanced recombinant M-CSF production in CHO cells by glycerol addition: model and validation. *Cytotechnology*, 54(2): 89-96.

Liu CH, Chu IM, Hwang SM. 2001. Enhanced expression of various exogenous genes in recombinant Chinese hamster ovary cells in presence of dimethyl sulfoxide. *Biotechnology Letters*, 23(20): 1641-1645.

Liu J, Nguyen MDH, Andya JD, Shire SJ. 2005. Reversible self-association increases the viscosity of a concentrated monoclonal antibody in aqueous solution. *Journal of Pharmaceutical Sciences*, 94(9): 1928-1940.

Liu X, Liu J, Wright TW, Lee J, Lio P, Donahue-Hjelle L, Ravnikar P, Wu F. 2010. Isolation of novel high-osmolarity resistant CHO DG44 cells after suspension of DNA mismatch repair. *BioProcess International*, 8(4): 68-76.

Liu Z, Cashion LM, Twu JJ. 1997. A systematic comparison of relative promoter/enhancer activities in mammalian cell lines. *Analytical Biochemistry*, 246(1): 150-152.

Lobito AA, Kimberley FC, Muppidi JR, Komarow H, Jackson AJ, Hull KM, Kastner DL, Screaton GR, Siegel RM. 2006. Abnormal disulfide-linked oligomerization results in ER retention and altered signaling by TNFR1 mutants in TNFR1-associated periodic fever syndrome (TRAPS). *Blood*, 108(4): 1320-1327.

Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J. 2000. Molecular Cell Biology, 4th edition. New York: W. H. Freeman.

Lu Y, Harding SE, Rowe AI, Davis KG, Fish B, Varley P, Gee C, Mulot S. 2008. The effect of a point mutation on the stability of IgG4 as monitored by analytical ultracentrifugation. *Journal of Pharmaceutical Sciences*, 97(2): 960-969.

Ma C, Martin S, Trask B, Hamlin JL. 1993. Sister chromatid fusion initiates amplification of the dihydrofolate reductase gene in Chinese hamster cells. *Genes & Development*, 7(4): 605-620.

Ma Z, Yi X, Zhang Y. 2008. Enhanced intracellular accumulation of recombinant HBsAg in CHO cells by dimethyl sulfoxide. *Process Biochemistry*, 43(6): 690-695.

Mahler H-C, Friess W, Grauschopf U, Kiese S. 2009. Protein aggregation: Pathways, induction factors and analysis. *Journal of Pharmaceutical Sciences*, 98(9): 2909-2934.

Majors BS, Betenbaugh MJ, Pederson NF, Chiang GG. 2008. Enhancement of transient gene expression and culture viability using Chinese hamster ovary cells overexpressing Bcl-x_L. *Biotechnology and Bioengineering*, 101(3): 567-578.
Makrides SC. 1999. Components of vectors for gene transfer and expression in mammalian cells. *Protein Expression and Purification*, 17(2): 183-202.
Mandenius C, Brundin A. 2008. Bioprocess Optimization using design-of-experiments methodology. *Biotechnology Progress*, 24(6): 1191-1203.
Mariati, Ho SCL, Yap MGS, Yang Y. 2012. Post-transcriptional regulatory elements for enhancing transient gene expression levels in mammalian cells. *Methods and Protocols*, 801: 125-135.
Marquardt T, Helenius A. 1992. Misfolding and aggregation of newly synthesized proteins in the endoplasmic reticulum. *Journal of Cell Biology*, 117(3): 505-513.
Mason M, Sweeney B, Cain K, Stephens P, Sharfstein ST. 2012. Identifying bottlenecks in transient and stable production of recombinant monoclonal-antibody sequence variants in Chinese hamster ovary cells. *Biotechnology Progress*, 28(3): 846-855.
Mayer M, Kies U, Kammermeier R, Buchner J. 2000. BiP and PDI cooperate in the oxidative folding of antibodies *in vitro*. *Journal of Biological Chemistry*, 275(38): 29421-29425.
Mazzarella RA, Marcus N, Haugejorden SM, Balcarek JM, Baldassare JJ, Roy B, Li LJ, Lee AS, Green M. 1994. ERp61 is GRP58, a stress-inducible luminal endoplasmic reticulum protein, but is devoid of phosphatidylinositide-specific phospholipase C activity. *Archives of Biochemistry and Biophysics*, 308(2): 454-460.
McCullough KD, Martindale JL, Klotz LO, Aw TY, Holbrook NJ. 2001. Gadd153 sensitizes cells to endoplasmic reticulum stress by down-regulating Bc12 and perturbing the cellular redox state. *Molecular and Cellular Biology*, 21(4): 1249-1259.
McKinney KL, Dilwith R, Belfort G. 1995. Optimizing antibody production in batch hybridoma cell culture. *Journal of Biotechnology*, 40(1): 31-48.
McLeod J, O'Callaghan PM, Pybus LP, Wilkinson SJ, Root T, Racher AJ, James DC. 2011. An empirical modeling platform to evaluate the relative control discrete CHO cell synthetic processes exert over recombinant monoclonal antibody production process titer. *Biotechnology and Bioengineering*, 108(9): 2193-2204.
Mead EJ, Chiverton LM, Smales CM, von der Haar T. 2009. Identification of the Limitations on recombinant gene expression in CHO cell lines with varying luciferase production rates. *Biotechnology and Bioengineering*, 102(6): 1593-1602.
Mehier-Humbert S, Guy RH. 2005. Physical methods for gene transfer: Improving the kinetics of gene delivery into cells. *Advanced Drug Delivery Reviews*, 57(5): 733-753.
Mei B, Low SC, Krassova S, Peters RT, Pierce GF, Dumont JA. 2013. *Monomeric Fc-fusion proteins*. In: Schmidt SR, editor. *Fusion protein technologies for biopharmaceuticals: Application and challenges*. 1st ed. New Jersey: John Wiley & Sons. p 107-118.
Meng YG, Liang J, Wong WL, Chisholm V. 2000. Green fluorescent protein as a second selectable marker for selection of high producing clones from transfected CHO cells. *Gene*, 242(1-2): 201-207.
Merlo LMF, Pepper JW, Reid BJ, Maley CC. 2006. Cancer as an evolutionary and ecological process. *Nature Reviews Cancer*, 6(12): 924-935.
Merritt SE, Palsson BO. 1993. Loss of Antibody Productivity is highly reproducible in multiple hybridoma subclones. *Biotechnology and Bioengineering*, 42(2): 247-250.
Mezghrani A, Fassio A, Benham A, Simmen T, Braakman I, Sitia R. 2001. Manipulation of oxidative protein folding and PDI redox state in mammalian cells. *EMBO Journal*, 20(22): 6288-6296.
Mi S, Lee X, Shao ZH, Thill G, Ji BX, Relton J, Levesque M, Allaire N, Perrin S, Sands B, Crowell T, Cate RL, McCoy JM, Pepinsky RB. 2004. LINGO-1 is a component of the Nogo-66 receptor/p75 signaling complex. *Nature Neuroscience*, 7(3): 221-228.
Mi S, Pepinsky RB, McCoy J. 2013. *Treatment of conditions involving demyelination.* European patent EP2543384A2.
Miller WM, Blanch HW, Wilke CR. 1988. A kinetic analysis of hybridoma growth and metabolism in batch and continuous suspension culture: Effect of nutrient concentration, dilution rate, and pH. *Biotechnology and Bioengineering*, 32(8): 947-965.

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. *Biotechnology and Bioengineering*, 107(2): 337-346.

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. *Biotechnology and Bioengineering*, 98(3): 611-615.

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. *Biotechnology Progress*, 30(5): 1161-1170.

Naderi S, Meshram M, Wei C, McConkey B, Ingalls B, Budman H, Scharer J. 2011. Development of a mathematical model for evaluating the dynamics of normal and apoptotic Chinese hamster ovary cells. *Biotechnology Progress*, 27(5):1197-1205.

Nehlsen K, Schucht R, da Gama-Norton L, Kroemer W, Baer A, Cayli A, Hauser H, Wirth D. 2009. Recombinant protein expression by targeting pre-selected chromosomal loci. *BMC Biotechnology*, 9:100.

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. *Applied Microbiology and Biotechnology*, 97(6): 2531-2539.

Novoa I, Zeng HQ, Harding HP, Ron D. 2001. Feedback inhibition of the unfolded protein response by *GADD34*-mediated dephosphorylation of eIF2α. *Journal of Cell Biology*, 153(5): 1011-1021.

Nyberg GB, Balcarcel RR, Follstad BD, Stephanopoulos G, Wang DI. 1999. Metabolic effects on recombinant interferon-gamma glycosylation in continuous culture of Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 62(3): 336-347.

O'Callaghan PM, James DC. 2008. Systems biotechnology of mammalian cell factories. *Briefings in Functional Genomics & Proteomics*, 7(2): 95-110.

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. *Biotechnology and Bioengineering*, 106(6): 938-951.

O'Connor S, Li E, Majors BS, He L, Placone J, Baycin D, Betenbaugh MJ, Hristova K. 2009. Increased expression of the integral membrane protein ErbB2 in Chinese hamster ovary cells expressing the anti-apoptotic gene Bcl-x_L. *Protein Expression and Purification*, 67(1): 41-47.

Oh MK, Scoles DR, Haipek C, Strand AD, Gutmann DH, Olson JM, Pulst SM. 2003. Genetic heterogeneity of stably transfected cell lines revealed by expression profiling with oligonucleotide microarrays. *Journal of Cellular Biochemistry*, 90(5): 1068-1078.

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. *Biotechnology and Bioengineering*, 100(2): 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. *Journal of Bioscience and Bioengineering*, 106(6): 568-573.

Onitsuka M, Tatsuzawa M, Asano R, Kumagai I, Shirai A, Maseda H, Omasa T. 2014. Trehalose suppresses antibody aggregation during the culture of Chinese hamster ovary cells. *Journal of Bioscience and Bioengineering*, 117(5): 632-638.

Pagani M, Fabbri M, Benedetti C, Fassio A, Pilati S, Bulleid NJ, Cabibbo A, Sitia R. 2000. Endoplasmic reticulum oxidoreductin 1-Lβ (*ERO1-Lβ*), a human gene induced in the course of the unfolded protein response. *The Journal of Biological Chemistry*, 275 (31): 23685-23692.

Palermo DP, Degraaf ME, Marotti KR, Rehberg E, Post LE. 1991. Production of analytical quantities of recombinant proteins in Chinese hamster ovary cells using sodium butyrate to elevate gene expression. *Journal of Biotechnology*, 19(1): 35-47.

Pandhal J, Wright PC. 2010. *N*-Linked glycoengineering for human therapeutic proteins in bacteria. *Biotechnology Letters*, 32(9): 1189-1198.

Papp E, Csermely P. 2006. *Chemical chaperones: Mechanisms of action and potential use*. In: Starke BR, Gaestel M, editors. *Molecular chaperones in health and disease.* Berlin: Springer Berlin Heidelberg. p 405-416.
Pepinsky RB, Arndt JW, Quan C, Gao Y, Quintero-Monzon O, Lee X, Mi S. 2014. Structure of the LINGO-1-anti-LINGO-1 Li81 antibody complex provides insights into the biology of LINGO-1 and the mechanism of action of the antibody therapy. *Journal of Pharmacology and Experimental Therapeutics*, 350(1): 110-123.
Perlmutter DH. 2002. Chemical chaperones: A pharmacological strategy for disorders of protein folding and trafficking. *Pediatric Research*, 52(6): 832-836.
Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. 2004. Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology*, 14(2): 103-114.
Pfaffl MW, Georgieva TM, Georgiev IP, Ontsouka E, Hageleit M, Blum JW. 2002. Real-time RT-PCR quantification of insulin-like growth factor (IGF)-1, IGF-1 receptor, IGF-2, IGF-2 receptor, insulin receptor, growth hormone receptor, IGF-binding proteins 1, 2 and 3 in the bovine species. *Domestic Animal Endocrinology*, 22(2): 91-102.
Philo JS. 2006. Is any measurement method optimal for all aggregate sizes and types? *The AAPS Journal*, 8(3): E564-E571.
PhRMA. 2012. *The R&D process: The road to new medicines. 2012 Pharmaceutical industry profile*. www.phrma.org/sites/default/files/159/phrma_industry_profile.pdf.
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. *Biotechnology and Bioengineering*, 108(2): 386-394.
Piechaczek C, Fetzer C, Baiker A, Bode J, Lipps HJ. 1999. A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. *Nucleic Acids Research*, 27(2): 426-428.
Pilbrough W, Munro TP, Gray P. 2009. Intraclonal protein expression heterogeneity in recombinant CHO cells. *PloS One*, 4(12): e8432.
Prasad S, Khadatare PB, Roy I. 2011. Effect of chemical chaperones in improving the solubility of recombinant proteins in *Escherichia coli*. *Applied and Environmental Microbiology*, 77(13): 4603-4609.
Prentice HL, Ehrenfels BN, Sisk WP. 2007. Improving performance of mammalian cells in fed-batch processes through "bioreactor evolution". *Biotechnology Progress*, 23(2): 458-464.
Pybus LP, Dean G, West NR, Smith A, Daramola O, Field R, Wilkinson SJ, James DC. 2014a. Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 111(2): 372-385.
Pybus LP, James DC, Dean G, Slidel T, Hardman C, Smith A, Daramola O, Field R. 2014b. Predicting the expression of recombinant monoclonal antibodies in Chinese hamster ovary cells based on sequence features of the CDR3 domain. *Biotechnology Progress*, 30(1): 188-197.
Pörtner R, Schilling A, Lüdemann I, Märkl H. 1996. High density fed-batch cultures for hybridoma cells performed with the aid of a kinetic model. *Bioprocess Engineering*, 15(3): 117-124.
Qin JY, Zhang L, Clift KL, Hulur I, Xiang AP, Ren B-Z, Lahn BT. 2010. Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PloS One*, 5(5): e10611.
Raj A, van Oudenaarden A. 2008. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2): 216-226.
Rajendra Y, Kiseljak D, Baldi L, Hacker DL, Wurm FM. 2011. A simple high-yielding process for transient gene expression in CHO cells. *Journal of Biotechnology*, 153(1-2): 22-26.
Raymond C, Tom R, Perret S, Moussouami P, L'Abbe D, St-Laurent G, Durocher Y. 2011. A simplified polyethylenimine-mediated transfection process for large-scale and high-throughput applications. *Methods*, 55(1): 44-51.
Robinson D, Chu L. 2007. *Process development and design*. In: Stacey G, Davis J, editors. *Medicines from animal cell culture*. Chichester: John Wiley & Sons. p 173-185.

Rodriguez J, Spearman M, Huzel N, Butler M. 2005. Enhanced production of monomeric interferon-beta by CHO cells through the control of culture conditions. *Biotechnology Progress*, 21(1): 22-30.

Ron D, Walter P. 2007. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature Reviews Molecular Cell Biology*, 8(7): 519-529.

Roth SD, Schuettrumpf 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(9): e44505.

Ruiz JC, Wahl GM. 1990. Chromosomal destabilization during gene amplification. *Molecular and Cellular Biology*, 10(6): 3056-3066.

Running Deer J, Allison DS. 2004. High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1α gene. *Biotechnology Progress*, 20(3): 880-889.

Ryu JS, Kim TK, Chung JY, Lee GM. 2000. Osmoprotective effect of glycine betaine on foreign protein production in hyperosmotic recombinant Chinese hamster ovary cell cultures differs among cell lines. *Biotechnology and Bioengineering*, 70(2): 167-175.

Sandadi S, Ensari S, Kearns B. 2005. Heuristic optimization of antibody production by Chinese hamster ovary cells. *Biotechnology Progress*, 21(5): 1537-1542.

Sareneva T, Pirhonen J, Cantell K, Kalkkinen N, Julkunen I. 1994. Role of *N*-glycosylation in the synthesis, dimerization and secretion of human interferon-γ. *Biochemical Journal*, 303(3): 831-840.

Schäuble N, Lang S, Jung M, Cappel S, Schorr S, Ulucan Ö, Linxweiler J, Dudek J, Blum R, Helms V, Paton AW, Paton JC, Cavalié A, Zimmermann R. 2012. BiP-mediated closing of the Sec61 channel limits Ca^{2+} leakage from the ER. *EMBO Journal*, 31(15): 3282-3296.

Schiedner G, Hertel S, Bialek C, Kewes H, Waschuetza G, Volpers C. 2008. Efficient and reproducible generation of high-expressing, stable human cell lines without need for antibiotic selection. *BMC Biotechnology*, 8: 13.

Schilling CH, Edwards JS, Palsson BO. 1999. Toward metabolic phenomics: Analysis of genomic data using flux balances. *Biotechnology Progress*, 15(3): 288-295.

Schlabach MR, Hu JK, Li M, Elledge SJ. 2010. Synthetic design of strong promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 107(6): 2538-2543.

Schmid G, Schlaeger EJ, Wipf B. 2001. Non-GMP plasmid production for transient transfection in bioreactors. *Cytotechnology*, 35(3): 157-164.

Schmidt H, Jirstrand M. 2006. Systems Biology Toolbox for MATLAB: A computational platform for research in systems biology. *Bioinformatics*, 22(4): 514-515.

Schröder M. 2008. Engineering eukaryotic protein factories. *Biotechnology Letters*, 30(2): 187-196.

Schröder M, Kaufman RJ. 2005. The mammalian unfolded protein response. *Annual Review of Biochemistry*, 74: 739-789.

Schröder M, Korner C, Friedl P. 1999. Quantitative analysis of transcription and translation in gene amplified Chinese hamster ovary cells on the basis of a kinetic model. *Cytotechnology*, 29(2): 93-102.

Schröder M, Schafer R, Friedl P. 2002. Induction of protein aggregation in an early secretory compartment by elevation of expression level. *Biotechnology and Bioengineering*, 78(2): 131-140.

Seewöster T, Lehmann J. 1997. Cell size distribution as a parameter for the predetermination of exponential growth during repeated batch cultivation of CHO cells. *Biotechnology and Bioengineering*, 55(5): 793-797.

Sevier CS, Kaiser CA. 2006. Disulfide transfer between two conserved cysteine pairs imparts selectivity to protein oxidation by Ero1. *Molecular Biology of the Cell*, 17(5): 2256-2266.

Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, Yu X, Yang LM, Tan BK, Rosenwald A, Hurt EM, Petroulakis E, Sonenberg N, Yewdell JW, Calame K, Glimcher LH, Staudt LM. 2004. XBP1, downstream of Blimp-1, expands the secretory

apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity*, 21(1): 81-93.
Shelikoff M, Sinskey AJ, Stephanopoulos G. 1994. The effect of protein synthesis inhibitors on the glycosylation site occupancy of recombinant human prolactin. *Cytotechnology*, 15(1-3): 195-208.
Shukla AA, Gupta P, Han X. 2007. Protein aggregation kinetics during Protein A chromatography: Case study for an Fc fusion protein. *Journal of Chromatography A*, 1171(1-2): 22-28.
Shusta EV, Raines RT, Pluckthun A, Wittrup KD. 1998. Increasing the secretory capacity of Saccharomyces cerevisiae for production of single-chain antibody fragments. *Nature Biotechnology*, 16(8): 773-777.
Sidoli FR, Mantalaris A, Asprey SP. 2004. Modelling of mammalian cells and cell culture processes. *Cytotechnology*, 44(1-2): 27-46.
Simmen T, Lynes EM, Gesson K, Thomas G. 2010. Oxidative protein folding in the endoplasmic reticulum: Tight links to the mitochondria-associated membrane (MAM). *Biochimica Et Biophysica Acta-Biomembranes*, 1798(8): 1465-1473.
Sinacore MS, Drapeau D, Adamson SR. 2000. Adaptation of mammalian cells to growth in serum-free media. *Molecular Biotechnology,* 15(3): 249-257.
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. *Biotechnology and Bioengineering*, 88(4): 474-488.
Smith CL, Bolton A, Nguyen G. 2010. Genomic and epigenomic instability, fragile sites, schizophrenia and autism. *Current Genomics*, 11(6): 447-469.
Smith JD, Tang BC, Robinson AS. 2004. Protein disulfide isomerase, but not binding protein, overexpression enhances secretion of a non-disulfide-bonded protein in yeast. *Biotechnology and Bioengineering*, 85(3): 340-350.
Snijder B, Pelkmans L. 2011. Origins of regulated cell-to-cell variability. *Nature Reviews Molecular Cell Biology*, 12(2): 119-125.
Sommerfeld S, Strube J. 2005. Challenges in biotechnology production—generic processes and process optimization for monoclonal antibodies. *Chemical Engineering and Processing*, 44(10): 1123-1137.
Sriburi R, Bommiasamy H, Buldak GL, Robbins GR, Frank M, Jackowski S, Brewer JW. 2007. Coordinate regulation of phospholipid biosynthesis and secretory pathway gene expression in XBP-1(S)-induced endoplasmic reticulum biogenesis. *Journal of Biological Chemistry*, 282(10): 7024-7034.
Stansfield SH, Allen EE, Dinnis DM, Racher AJ, Birch JR, James DC. 2007. Dynamic analysis of GS-NS0 cells producing a recombinant monoclonal antibody during fed-batch culture. *Biotechnology and Bioengineering*, 97(2): 410-424.
Stapleton JA, Endo K, Fujita Y, Hayashi K, Takinoue M, Saito H, Inoue T. 2012. Feedback control of protein expression in mammalian cells by tunable synthetic translational inhibition. *ACS Synthetic Biology*, 1(3): 83-88.
Stinski MF, Isomura H. 2008. Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. *Medical Microbiology and Immunology*, 197(2): 223-231.
Stockholm D, Benchaouir R, Picot J, Rameau P, Neildez TMA, Landini G, Laplace-Builhe C, Paldi A. 2007. The origin of phenotypic heterogeneity in a clonal cell population *in vitro*. *PloS One*, 2(4): e394.
Strand J, Huang C-T, Xu J. 2013. Characterization of Fc-fusion protein aggregates derived from extracellular domain disulfide bond rearrangements. *Journal of Pharmaceutical Sciences*, 102(2): 441-453.
Strnad J, Brinc M, Spudic V, Jelnikar N, Mirnik L, Carman B, Kravanja Z. 2010. Optimization of cultivation conditions in spin tubes for Chinese hamster ovary cells producing erythropoietin and the comparison of glycosylation patterns in different cultivation vessels. *Biotechnology Progress*, 26(3): 653-663.
Suen KF, Turner MS, Gao F, Liu B, Althage A, Slavin A, Ou W, Zuo E, Eckart M, Ogawa T, Yamada M, Tuntland T, Harris JL, Trauger JW. 2010. Transient expression of an IL-

23R extracellular domain Fc fusion protein in CHO vs. HEK cells results in improved plasma exposure. *Protein Expression and Purification*, 71(1): 96-102.
Sung YH, Lee GM. 2005. Enhanced human thrombopoietin production by sodium butyrate addition to serum-free suspension culture of Bcl-2-overexpressing CHO cells. *Biotechnology Progress*, 21(1): 50-57.
Suzuki E, Ollis DF. 1990. Enhanced antibody production at slowed growth rates: Experimental demonstration and a simple structured model. *Biotechnology Progress*, 6(3): 231-236.
Swiech K, Picanco-Castro V, Covas DT. 2012. Human cells: New platform for recombinant therapeutic protein production. *Protein Expression and Purification*, 84(1): 147-153.
Szilágyi A, Kardos J, Osváth S, Barna L, Závodszky P. 2007. *Protein Folding*. In: Lajtha A, Banik N, editors. *Handbook of Neurochemistry and Molecular Neurobiology*. New York: Springer US. p 303-343.
Tabuchi H, Sugiyama T, Tanaka S, Tainaka S. 2010. Overexpression of taurine transporter in Chinese hamster ovary cells can enhance cell viability and product yield, while promoting glutamine consumption. *Biotechnology and Bioengineering*, 107(6): 998-1003.
Tachibana H, Cheng XJ, Watanabe K, Takekoshi M, Maeda F, Aotsuka S, Kaneda Y, Takeuchi T, Ihara S. 1999. Preparation of recombinant human monoclonal antibody Fab fragments specific for *Entamoeba histolytica*. *Clinical and Diagnostic Laboratory Immunology*, 6(3): 383-387.
Takahashi Y, Nishikawa M, Takiguchi N, Suehara T, Takakura Y. 2011. Saturation of transgene protein synthesis from mRNA in cells producing a large number of transgene mRNA. *Biotechnology and Bioengineering*, 108(10): 2380-2389.
Taylor FR, Prentice HL, Garber EA, Fajardo HA, Vasilyeva E, Pepinsky RB. 2006. Suppression of sodium dodecyl sulfate-polyacrylamide gel electrophoresis sample preparation artifacts for analysis of IgG4 half-antibody. *Analytical Biochemistry*, 353(2): 204-208.
Tharmalingam T, Sunley K, Butler M. 2008. High yields of monomeric recombinant beta-interferon from macroporous microcarrier cultures under hypothermic conditions. *Biotechnology Progress*, 24(4): 832-838.
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. *Biotechnology Progress*, 28(1): 179-187.
Thorburn A. 2008. Apoptosis and autophagy: regulatory connections between two supposedly different processes. *Apoptosis*, 13(1): 1-9.
Tigges M, Fussenegger M. 2006. Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. *Metabolic Engineering*, 8(3): 264-272.
Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M. 2009. A tunable synthetic mammalian oscillator. *Nature*, 457(7227): 309-312.
Tornøe J, Kusk P, Johansen TE, Jensen PR. 2002. Generation of a synthetic mammalian promoter library by modification of sequences spacing transcription factor binding sites. *Gene*, 297(1-2): 21-32.
Trusina A, Papa FR, Tang C. 2008. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *Proceedings of the National Academy of Sciences of the United States of America*, 105(51): 20280-20285.
Tveten K, Holla OL, Ranheim T, Berge KE, Leren TP, Kulseth MA. 2007. 4-Phenylbutyrate restores the functionality of a misfolded mutant low-density lipoprotein receptor. *FEBS Journal*, 274(8): 1881-1893.
Tyedmers J, Mogk A, Bukau B. 2010. Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology*, 11(11): 777-788.
Tzur A, Kafri R, LeBleu VS, Lahav G, Kirschner MW. 2009. Cell growth and size homeostasis in proliferating animal cells. *Science*, 325(5937): 167-171.
Vagenende V, Yap MGS, Trout BL. 2009. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry*, 48(46): 11084-11096.
van Berkel PHC, Gerritsen J, Perdok G, Valbjorn J, Vink T, van de Winkel JGJ, Parren PWHI. 2009. N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG. *Biotechnology Progress*, 25(1): 244-251.

Van Craenenbroeck K, Vanhoenacker P, Haegeman G. 2000. Episomal vectors for gene expression in mammalian cells. *European Journal of Biochemistry*, 267(18): 5665-5578.

van den Berg B, Ellis RJ, Dobson CM. 1999. Effects of macromolecular crowding on protein folding and aggregation. *EMBO Journal*, 18(24): 6927-6933.

Vaz AIF, Vicente LN. 2007. A particle swarm pattern search method for bound constrained global optimization. *Journal of Global Optimization*, 39(2): 197-219.

Wahl GM, Vitto L, Padgett RA, Stark GR. 1982. Single-copy and amplified CAD genes in Syrian Hamster chromosomes localized by a highly sensitive method for *in situ* hybridization. *Molecular and Cellular Biology*, 2(3): 308-319.

Wahl LM, Gerrish PJ, Saika-Voivod I. 2002. Evaluating the impact of population bottlenecks in experimental evolution. *Genetics*, 162(2): 961-971.

Walsh G. 2010. Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, 28(9): 917-924.

Walsh G. 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32(10): 992-1000.

Walsh G, Jefferis R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nature Biotechnology*, 20(10): 1241-1252.

Wang T, Fodor S, Hapuarachchi S, Jiang XG, Chen K, Apostol I, Huang G. 2013. Analysis and characterization of aggregation of a therapeutic Fc-fusion protein. *Journal of Pharmaceutical and Biomedical Analysis*, 72: 59-64.

Wang W. 2005. Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics*, 289(1-2):1 -30.

Wang W, Nema S, Teagarden D. 2010. Protein aggregation—Pathways and influencing factors. *International Journal of Pharmaceutics*, 390(2): 89-99.

Wang W, Yi X, Zhang Y. 2007. Gene transcription acceleration: Main cause of hepatitis B surface antigen production improvement by dimethyl sulfoxide in the culture of Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 97(3): 526-535.

Wang Z, Ma X, Zhao L, Fan L, Tan W-S. 2012. Expression of anti-apoptotic 30Kc6 gene inhibiting hyperosmotic pressure-induced apoptosis in antibody-producing Chinese hamster ovary cells. *Process Biochemistry*, 47(5): 735-741.

West N. 2014. PhD thesis: Development of a tunable mammalian protein expression system and an investigation of promoter interference in three promoters often utilized in the production of biopharmaceuticals. University of Sheffield, UK.

Wilson AS, Power BE, Molloy PL. 2007. DNA hypomethylation and human diseases. *Biochimica Et Biophysica Acta-Reviews on Cancer*, 1775(1): 138-162.

Wulhfard S, Baldi L, Hacker DL, Wurm F. 2010. Valproic acid enhances recombinant mRNA and protein levels in transiently transfected Chinese hamster ovary cells. *Journal of Biotechnology*, 148(2-3): 128-132.

Wulhfard S, Kiseljak D, Baldi L, Hacker DL, Wurm FM. 2012. *Transgene mRNA levels and stability are key factors to enhance transient gene expression in CHO DG44 cells*. In: Jenkins N, Barron N, Alves P, editors. *Proceedings of the 21st annual meeting of the European Society for Animal Cell Technology (ESACT), Dublin, Ireland, June 7-10, 2009*. Dordrecht: Springer Netherlands. p 121-124.

Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology*, 22(11): 1393-1398.

Xia W, Bringmann P, McClary J, Jones PP, Manzana W, Zhu Y, Wang SJ, Liu Y, Harvey S, Madlansacay MR, McLean K, Rosser MP, MacRobbie J, Olsen CL, Cobb RR. 2006. High levels of protein expression using different mammalian CMV promoters in several cell lines. *Protein Expression and Purification*, 45(1): 115-124.

Xie LZ, Wang DIC. 1994. Applications of improved stoichiometric model in medium design and fed-batch cultivation of animal-cells in bioreactor. *Cytotechnology*, 15(1-3): 17-29.

Yam GH, Gaplovska-Kysela K, Zuber C, Roth J. 2007. Sodium 4-phenylbutyrate acts as a chemical chaperone on misfolded myocilin to rescue cells from endoplasmic reticulum stress and apoptosis. *Investigative Ophthalmology & Visual Science*, 48(4): 1683-1690.

Yamamoto K, Yoshida H, Kokame K, Kaufman RJ, Mori K. 2004. Differential contributions of ATF6 and XBP1 to the activation of endoplasmic reticulum stress-responsive cis-acting elements ERSE, UPRE and ERSE-II. *Journal of Biochemistry*, 136(3): 343-350.

Yates JL, Warren N, Sugden B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature*, 313(6005): 812-815.

Ye J, Alvin K, Latif H, Hsu A, Parikh V, Whitmer T, Tellers M, Edmonds MCdlC, Ly J, Salmon P, Markusen JF. 2010. Rapid protein production using CHO stable transfection pools. *Biotechnology Progress*, 26(5): 1431-1437.

Yee JC, Gatti MdL, Philp RJ, Yap M, Hu W. 2008. Genomic and proteomic exploration of CHO and hybridoma cells under sodium butyrate treatment. *Biotechnology and Bioengineering*, 99(5): 1186-1204.

Yim SS, An SJ, Kang M, Lee J, Jeong KJ. 2013. Isolation of fully synthetic promoters for high-level gene expression in *Corynebacterium glutamicum*. *Biotechnology and Bioengineering*, 110(11): 2959-2969.

Yoshida H, Haze K, Yanagi H, Yura T, Mori K. 1998. Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins: Involvement of basic leucine zipper transcription factors. *Journal of Biological Chemistry*, 273(50): 33741-33749.

Yoshida H, Matsui T, Hosokawa N, Kaufman RJ, Nagata K, Mori K. 2003. A time-dependent phase shift in the mammalian unfolded protein response. *Developmental Cell*, 4(2): 265-271.

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(7): 881-891.

Yoshida H, Okada T, Haze K, Yanagi H, Yura T, Negishi M, Mori K. 2000. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Molecular and Cellular Biology*, 20(18): 6755-6767.

Yoshida H, Yoshizawa T, Shibasaki F, Shoji S, Kanazawa I. 2002. Chemical chaperones reduce aggregate formation and cell death caused by the truncated Machado-Joseph disease gene product with an expanded polyglutamine stretch. *Neurobiology of Disease*, 10(2): 88-99.

Yoshikawa Y, Nakanishi F, Ogura Y, Oi D, Omasa T, Katakura Y, Kishimoto M, Suga K. 2000. Amplified gene location in chromosomal DNA affected recombinant protein production and stability of amplified genes. *Biotechnology Progress*, 16(5): 710-715.

Zabner J, Fasbender AJ, Moninger T, Poellinger KA, Welsh MJ. 1995. Cellular and molecular barriers to gene-transfer by a cationic lipid. *Journal of Biological Chemistry*, 270(32): 18997-19007.

Zarrin AA, Malkin L, Fong I, Luk KD, Ghose A, Berinstein NL. 1999. Comparison of CMV, RSV, SV40 viral and V lambda 1 cellular promoters in B and T lymphoid and non-lymphoid cell lines. *Biochimica Et Biophysica Acta-Gene Structure and Expression*, 1446(1-2): 135-139.

Zeitelhofer M, Vessey JP, Xie Y, Tuebing F, Thomas S, Kiebler M, Dahm R. 2007. High-efficiency transfection of mammalian neurons via nucleofection. *Nature Protocols*, 2(7): 1692-1704.

Zhang L, Chou CP, Moo-Young M. 2011. Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system. *Biotechnology Advances*, 29(6): 923-929.

Zhang PC, McGrath BC, Reinert J, Olsen DS, Lei L, Gill S, Wek SA, Vattem KM, Wek RC, Kimball SR, Jefferson LS, Cavener DR. 2002. The GCN2 eIF2α kinase is required for adaptation to amino acid deprivation in mice. *Molecular and Cellular Biology*, 22(19): 6681-6688.

Zhu A, Hurst R. (2002). Anti-N-glycolylneuraminic acid antibodies identified in healthy human serum. *Xenotransplantation*, 9(6): 376-381.

Appendix A

Plasmid maps

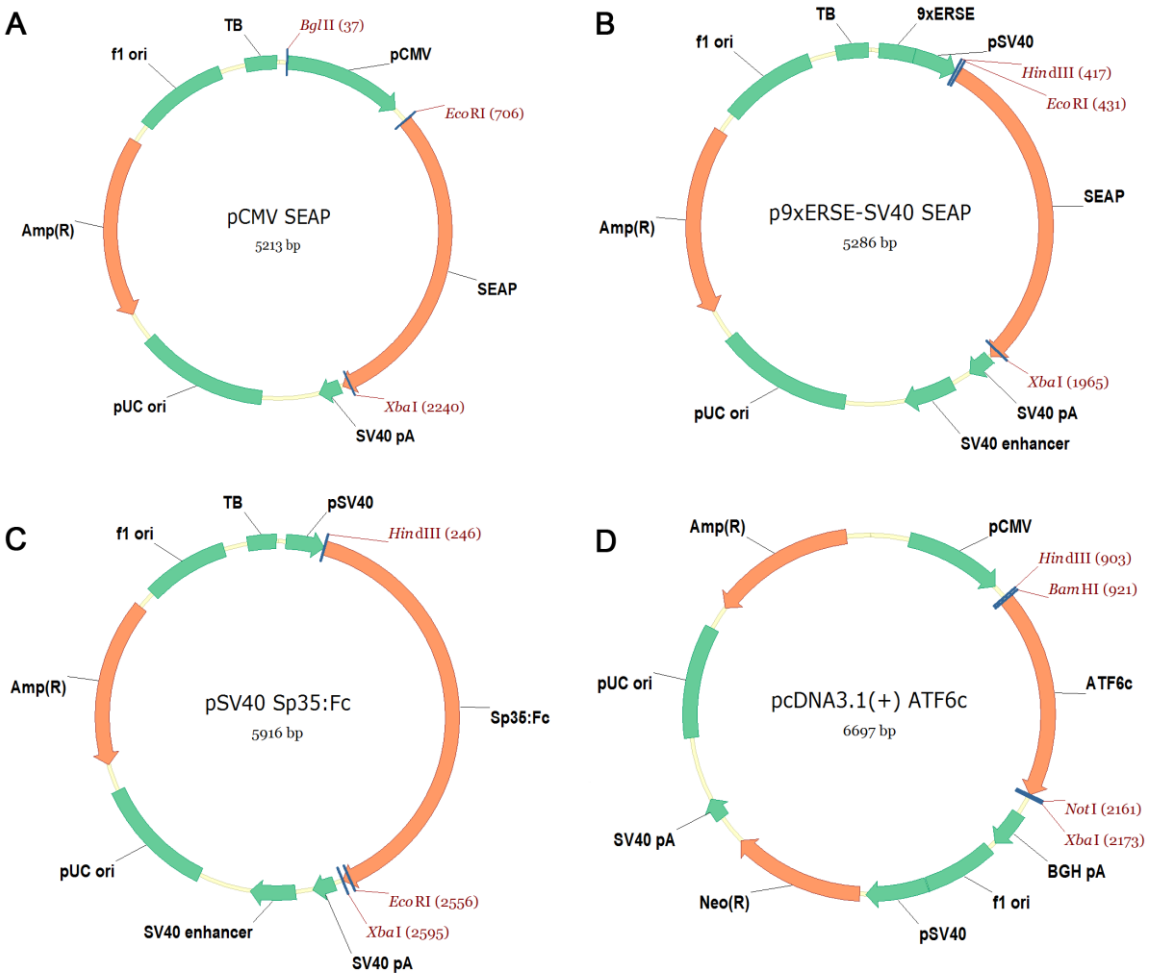

Figure A1: Plasmids constructed/provided by other people for use in this study. (A) The CMV SEAP vector encoding SEAP. The plasmid was modified from pSEAP2-Control vector (Clontech) by replacing the original SV40 promoter with a CMV promoter and removing the SV40 enhancer. **(B)** The 9xERSE-SV40 vector encoding SEAP. 9 units of ERSE sequence were inserted upstream of the SV40 promoter of pSEAP2-Control vector. Similar 9xERSE-SV40 vectors encoding Sp35:Fc or ATF6c were also constructed (not shown). **(C)** The SV40 Sp35:Fc vector encoding Sp35:Fc was constructed by replacing the SEAP gene in pSEAP2-Control with an Sp35:Fc gene. **(D)** The pcDNA3.1(+) vector (Life Technologies) driven by a CMV promoter with ATF6c gene insert. A pcDNA3.1(+) vector encoding XBP1s was also constructed (not shown).

Appendix B

SBToolBox2 codes

```
********** MODEL NAME
Sp35:Fc fusion protein production in CHO cells

********** MODEL NOTES
Aggregation is assumed to be irreversible and occur in the ER
Parameter values shown are for 3 ug rDNA load

********** MODEL STATES
%Define mass balance equations
d/dt(mRNA) = Ngene*Sm-kdeg*mRNA-u*mRNA %mRNA
d/dt(P) = TP*mRNA-2*RP2-u*P %Monomer
d/dt(P2ER) = RP2-2*RP4-RP6-kERP2*P2ER-kuP2*P2ER-u*P2ER %Dimer in ER
d/dt(P4ER) = RP4-RP6-kERP4*P4ER-kuP2*P4ER-u*P4ER %Tetramer in ER
d/dt(P6ER) = RP6-kERP6*P6ER-kuP2*P6ER-u*P6ER %Hexamer in ER
d/dt(Pu) = kuP*P-u*Pu %Ubiquitinated monomer
d/dt(P2u) = kuP2*P2ER-u*P2u %Ubiquitinated dimer
d/dt(P2G) = kERP2*P2ER-kGP2*P2G-u*P2G %Dimer in Golgi complex
d/dt(P4G) = kERP4*P4ER-kGP4*P4G-u*P4G %Tetramer in Golgi complex
d/dt(P6G) = kERP6*P6ER-kGP6*P6G-u*P6G %Hexamer in Golgi complex

%Define initial conditions
mRNA(0) = 1e3
P(0) = 7e3
P2ER(0) = 9e4
P4ER(0) = 400
P6ER(0) = 50
Pu(0) = 300
P2u(0) = 300
P2G(0) = 300
P4G(0) = 1
P6G(0) = 0.1

********** MODEL PARAMETERS
%Define parameters (fitted values)
u = 0.0225714
Ngene = 72000
Sm = 0.001333
tdeg = 11.9463
TP = 42.0174
kfa = 0.000406347
kaggP4 = 6.01546e-8
kaggP6 = 8.04124e-8
kuP = 0.000917286
kuP2 = 7.91967e-5
thERP2 = 3.1003
thERP4 = 5.04359
thERP6 = 5.41146
thGP2 = 0.0104184
thGP4 = 0.0155464
thGP6 = 0.0173787

********** MODEL VARIABLES
%Define mRNA degradation and Sp35:Fc transport rates
kdeg = 0.6931/tdeg
kERP2 = 0.6931/thERP2
kERP4 = 0.6931/thERP4
```

```
kERP6 = 0.6931/thERP6
kGP2 = 0.6931/thGP2
kGP4 = 0.6931/thGP4
kGP6 = 0.6931/thGP6

%Define folding/assembly and aggregation reactions
RP2 = kfa*P*P
RP4 = kaggP4*P2ER*P2ER
RP6 = kaggP6*P2ER*P4ER

%Define specific productivity rate of Sp35:Fc species
qP = 3.32e-7*(kGP2*P2G)*24 %qP for dimer
qPaggP4 = 6.64e-7*(kGP4*P4G)*24 %qP for tetramer
qPaggP6 = 9.96e-7*(kGP6*P6G)*24 %qP for hexamer

********** MODEL REACTIONS


********** MODEL FUNCTIONS


********** MODEL EVENTS


********** MODEL MATLAB FUNCTIONS
```

Appendix C

MATLAB codes

```
%%This m file calculates local sensitivity analysis for Sp35:Fc model for
3 ug rDNA load

%Initialise global parameters
global u Ngene Sm kdeg TP kfa kaggP4 kaggP6 kuP kuP2 kERP2 kERP4 kERP6
kGP2 kGP4 kGP6

%Define fixed parameters
u = 0.0226;
Ngene = 72000;
lamda = 3.32e-7;

%Define perturbable parameters to their initial value
Smi = 0.001333;
tdegi = 11.95;
TPi = 42.02;
kfai = 0.0004063;
kaggP4i = 6.015e-008;
kaggP6i = 8.041e-008;
kuPi = 0.0009173;
kuP2i = 7.920e-005;
thERP2i = 3.100;
thERP4i = 5.044;
thERP6i = 5.411;
thGP2i = 0.01042;
thGP4i = 0.01555;
thGP6i = 0.01738;

%Define variables
kdegi = log(2)/tdegi;
kERP2i = log(2)/thERP2i;
kERP4i = log(2)/thERP4i;
kERP6i = log(2)/thERP6i;
kGP2i = log(2)/thGP2i;
kGP4i = log(2)/thGP4i;
kGP6i = log(2)/thGP6i;

%Perform local sensitivity analysis on the parameters
for parameterchange = 1:8

    %Define/reset parameters to their initial value
    Sm = Smi;
    kdeg = kdegi;
    TP = TPi;
    kfa = kfai;
    kaggP4 = kaggP4i;
    kaggP6 = kaggP6i;
    kuP = kuPi;
    kuP2 = kuP2i;
    kERP2 = kERP2i;
    kGP2 = kGP2i;

    x0 = [1e3 7e3 9e4 400 50 300 1 0.1 300 300]; %Initial conditions
    tspan = [0 700];
```

```
if parameterchange == 1 %Perturb Sm
    %Run mass balance equations with +1% perturbation
    Sm = Smi*1.01;
    [t,x] = ode45(@Sp35Fc,tspan,x0); %Call ODE function
    qPupper = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPupperf = qPupper(matsize(1));
    %Run mass balance equations with -1% perturbation
    Sm = Smi*0.99;
    [t,x] = ode45(@Sp35Fc,tspan,x0); %Call ODE function
    qPlower = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPlowerf = qPlower(matsize(1));
    %Calculate response coefficient on qP
    RCSm = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

elseif parameterchange == 2 %Perturb kdeg
    kdeg = kdegi*0.99;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPupper = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPupperf = qPupper(matsize(1));
    kdeg = kdegi*1.01;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPlower = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPlowerf = qPlower(matsize(1));
    RCkdeg = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

elseif parameterchange == 3 %Perturb TP
    TP = TPi*1.01;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPupper = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPupperf = qPupper(matsize(1));
    TP = TPi*0.99;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPlower = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPlowerf = qPlower(matsize(1));
    RCTP = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

elseif parameterchange == 4 %Perturb kfa
    kfa = kfai*1.01;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPupper = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPupperf = qPupper(matsize(1));
    kfa = kfai*0.99;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPlower = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPlowerf = qPlower(matsize(1));
    RCkfa = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

elseif parameterchange == 5 %Perturb kagg
    kaggP4 = kaggP4i*0.99;
    kaggP6 = kaggP6i*0.99;
    [t,x] = ode45(@Sp35Fc,tspan,x0);
    qPupper = lamda*kGP2*x(:,8)*24;
    matsize = size(x);
    qPupperf = qPupper(matsize(1));
```

```
        kaggP4 = kaggP4i*1.01;
        kaggP6 = kaggP6i*1.01;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPlower = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPlowerf = qPlower(matsize(1));
        RCkagg = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

     elseif parameterchange == 6 %Perturb kuP
        kuP = kuPi*0.99;
        kuP2 = kuP2i*0.99;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPupper = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPupperf = qPupper(matsize(1));
        kuP = kuPi*1.01;
        kuP2 = kuP2i*1.01;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPlower = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPlowerf = qPlower(matsize(1));
        RCku = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

    elseif parameterchange == 7 %Perturb kERP2
        kERP2 = kERP2i*1.01;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPupper = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPupperf = qPupper(matsize(1));
        kERP2 = kERP2i*0.99;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPlower = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPlowerf = qPlower(matsize(1));
        RCkERP2 = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01

    else %Perturb kGP2
        kGP2 = kGP2i*1.01;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPupper = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPupperf = qPupper(matsize(1));
        kGP2 = kGP2i*0.99;
        [t,x] = ode45(@Sp35Fc,tspan,x0);
        qPlower = lamda*kGP2*x(:,8)*24;
        matsize = size(x);
        qPlowerf = qPlower(matsize(1));
        RCkGP2 = (qPupperf - qPlowerf)/(qPupperf + qPlowerf)/0.01
    end
end
```

```
%%This ODE function is called for to simulate mass balance equations
function [massbalance] = Sp35Fc(t,x)

  %Initialise global parameters
  global Ngene u Sm kdeg TP kfa kaggP4 kaggP6 kuP kuP2 kERP2 kERP4
kERP6
```

```
    kGP2 kGP4 kGP6

    %Define folding/assembly and aggregation reactions
    RP2 = kfa*x(2)^2;
    RP4 = kaggP4*x(3)^2;
    RP6 = kaggP6*x(3)*x(4);

    % x(1): mRNA = Ngene*Sm-kdeg*mRNA-u*mRNA
    % x(2): [P] = TP*mRNA-2*RP2-kuP*[P]-u*[P]
    % x(3): [P2]ER = RP2-2*RP4-RP6-kERP2*[P2]ER-kuP2*[P2]ER-u*[P2]ER
    % x(4): [P4]ER = RP4-RP6-kERP4*[P4]ER-kuP2*[P4]ER-u*[P4]ER
    % x(5): [P6]ER = RP6-kERP6*[P6]ER-kuP2*[P6]ER-u*[P6]ER
    % x(6): [Pu] = kuP*[P]-u*[Pu]
    % x(7): [P2u] = kuP2*[P2]ER-u*[P2u]
    % x(8): [P2]G = kERP2*[P2]ER-kGP2*[P2]G-u*[P2]G
    % x(9): [P4]G = kERP4*[P4]ER-kGP4*[P4]G-u*[P4]G
    % x(10): [P6]G = kERP6*[P6]ER-kGP6*[P6]G-u*[P6]G

    %Run mass balance equations
    massbalance =
        [Ngene*Sm-kdeg*x(1)-u*x(1);
        TP*x(1)-2*RP2-kuP*x(2)-u*x(2);
        RP2-2*RP4-RP6-kERP2*x(3)-kuP2*x(3)-u*x(3);
        RP4-RP6-kERP4*x(4)-kuP2*x(4)-u*x(4);
        RP6-kERP6*x(5)-kuP2*x(5)-u*x(5);
        kuP*x(2)-u*x(6);
        kuP2*x(3)-u*x(7);
        kERP2*x(3)-kGP2*x(8)-u*x(8);
        kERP4*x(4)-kGP4*x(9)-u*x(9);
        kERP6*x(5)-kGP6*x(10)-u*x(10)];
end
```

Appendix D

SV40-ERSE systems

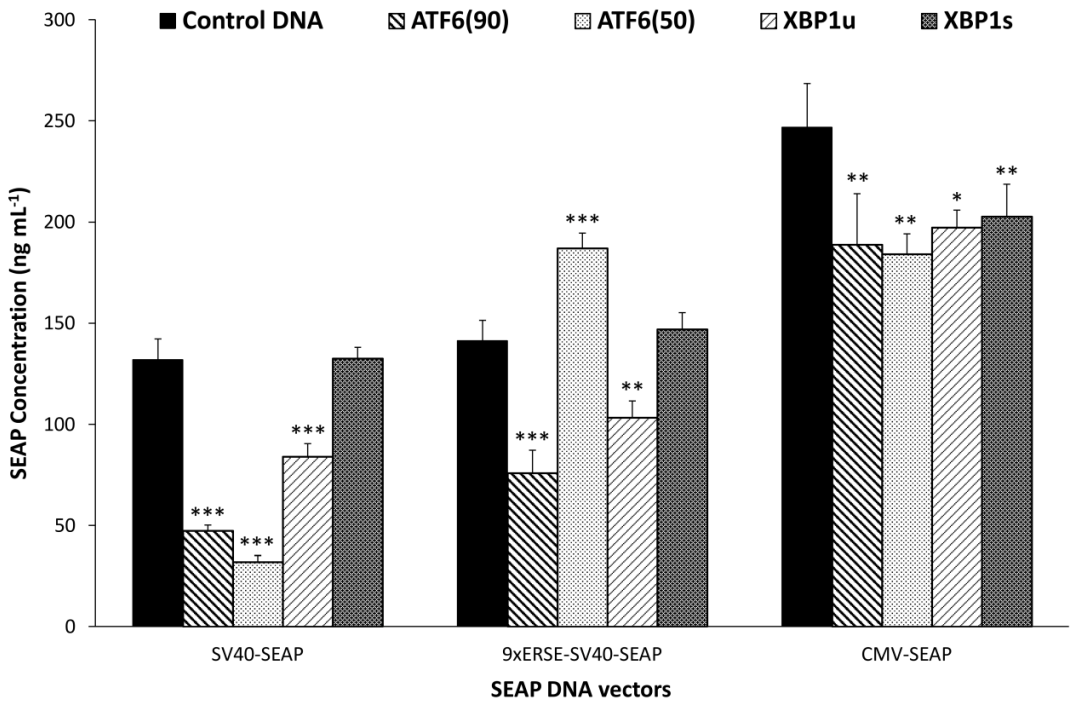

Figure D1: The effect of ATF6(90), ATF6(50), XBP1µ and XBP1 on SEAP protein expression from three different SEAP DNA vectors. SEAP protein expression for 0.5 µg of SV40-SEAP, 9xERSESV40-SEAP and CMV-SEAP vector DNA co-transfected with the 0.5 µg of UPR activator vectors (ATF6(90), ATF6(50), XBP1u and XBP1s) in CHOK1SV cells. The total amount of transfected DNA was kept constant using the –ve control DNA vector. Cell media was collected 48 hours post-transfection and analysed for SEAP protein expression. For each SEAP vector, mean values significantly different (Dunnett's test) from their control are indicated by asterisks (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). N = 3, error bars represent + 1 S.D. Figure and text are reproduced with permission form West (2014).

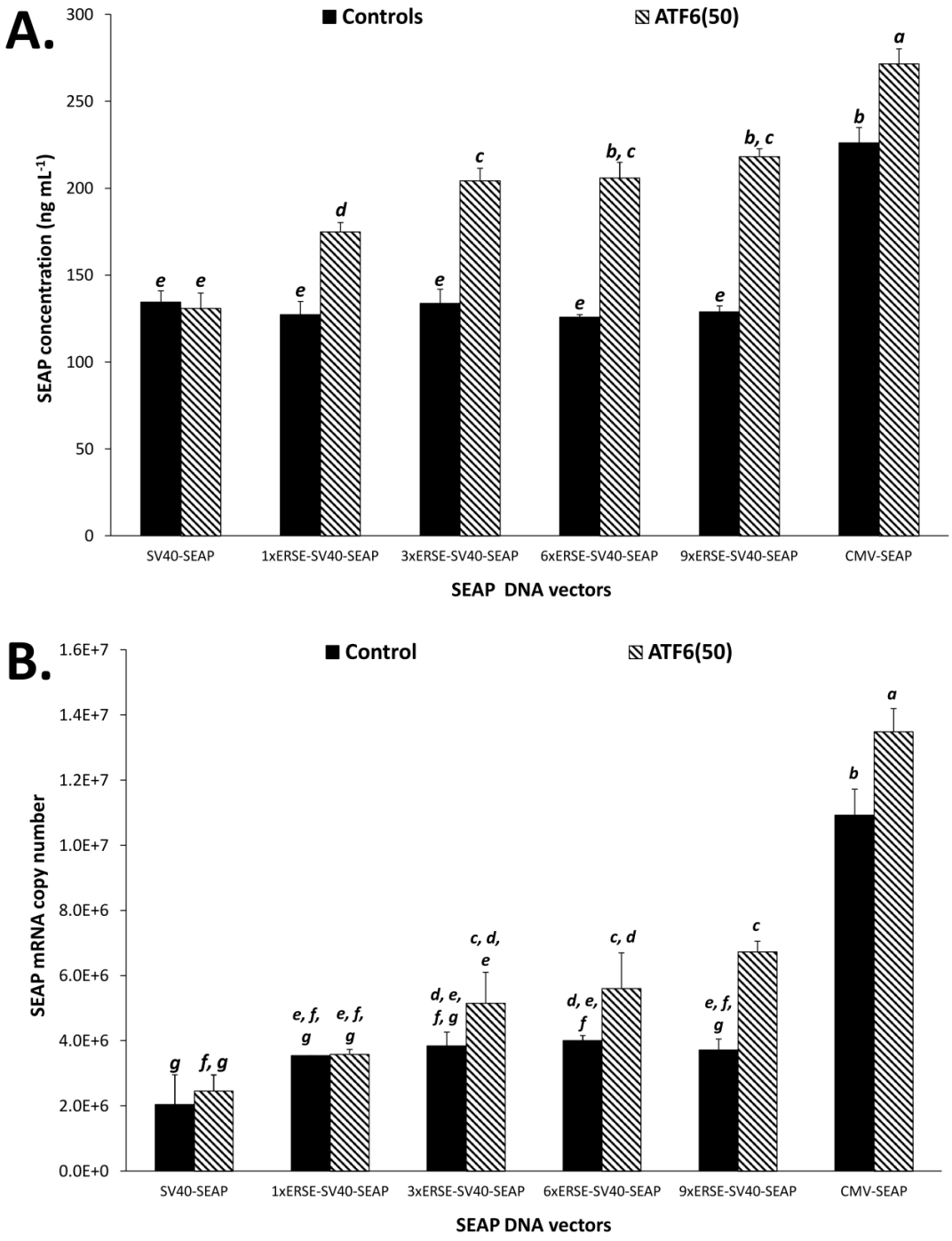

Figure D2: The effect of different numbers of the ERSE sequence on SEAP protein and mRNA copy number when co-transfected with the ATF6(50) DNA vector. (A) SEAP protein expression for 0.5 µg of SV40-SEAP, 1xERSE-SV40-SEAP, 3xERSE-SV40-SEAP, 6xERSE-SV40-SEAP, 9xERSE-SV40-SEAP and CMV-SEAP vectors when co-transfected with 0.125 µg ATF6(50) vector DNA in CHOK1SV cells. **(B)** SEAP mRNA copy numbers for the different SEAP vectors taken form samples as used in **A**. Cell media was collected 48 hours post-transfection and analysed for SEAP protein expression. For each SEAP vector, values with different letters differ significantly from each other (Tukey's test, $p < 0.05$). N = 3, Error bars represent + 1 S.D. Figure and text are reproduced with permission form West (2014).