# Use of CHO cell synthetic promoters to control transcription of mAb expression vector genes and screening of ER genes for their effect on mAb production

## William Kerry

Submitted for the degree of Doctor of Philosophy


Department of Chemical and Biological Engineering
University of Sheffield

# Acknowledgements

# Abstract

The recombinant manufacture of therapeutic monoclonal antibody (mAb) molecules from Chinese hamster ovary (CHO) cell cultures represents a vital supply of important medicines with global demand. Innovations within this recombinant manufacturing process that can either increase mAb yield or speed up the supply of mAb product are therefore very highly sought after. These innovations could come from genetic engineering of either the mAb expression vector or the CHO host cell. Chapters 3, 4 and 5 of this manuscript focus on plasmid vector engineering, specifically the development and implementation of synthetic promoter sequences to control the transcription of mAb LC/HC and GS selection marker genes, replacing the conventionally-used viral promoters hCMV and SV40 respectively. In chapter 4 different LC/HC synthetic promoter pairs were used, for the first time, in the generation of stably expressing CHO pools. Compared to the use of conventional hCMV promoters to control LC and HC transcription, these synthetic promoter-containing pools consistently recovered $3 - 7$ days faster from MSX selection and in some cases produced higher mean mAb titres (up to 1.3-fold). In chapter 5, synthetic promoters were used for the control of GS transcription in stably expressing CHO pools. The synthetic GS promoter SynSV40_2 was found to generate up to 1.9-fold higher mean cell-specific productivity (qP) when combined with LC/HC synthetic promoters compared to the conventional SV40 promoter. Chapter 6 focused on finding targets for CHO host cell engineering through high-throughput transient transfection screening of genes acting within the endoplasmic reticulum (ER) for their effect on mAb production. The gene pERP1 was shown consistently and robustly to increase qP of mAb-expressing cells by up to 1.6-fold and overexpression of this gene therefore represents a novel CHO host cell engineering target. Overall, the work described within this manuscript led to the discovery of strategies for improving CHO-based recombinant mAb manufacture.

# Contents

## List of tables

## List of figures

## Acronyms and Abbreviations

**AmpR**  Ampicillin resistance

**BiP**  Binding immunoglobulin protein

**CHO**  Chinese hamster ovary

**EEE**  Expression enhancing element

**ELISA**  Enzyme-linked immunosorbent assay

**ER**  Endoplasmic reticulum

**FBOG**  Fed batch overgrow

**GFP**  Green fluorescent protein

**GS**  Glutamine synthetase

**HC**  Heavy chain

**hCMV**  Human Cytomegalovirus

**IgG**  Immunoglobulin G

**IVCD**  Integral of viable cell density

**LC**  Light chain

**mAb**  Monoclonal antibody

**MFI**  Median fluorescence intensity

**MSX**  Methionine sulfoximine

**ORF**  Open reading frame

**Ori**  Bacterial origin of replication

**polyA**  Polyadenylation

**qP**  Cell specific productivity

**qRT-PCR**  Quantitative reverse transcription polymerase chain reaction

**SV40**  Simian virus 40

**TF**  Transcription factor

**TFRE**  Transcription factor regulatory element

**TSS**  Transcription start site

**UPR**  Unfolded protein response

**UTR**  Untranslated region

**VCD**  Viable cell density

# Chapter 1
# Introduction

*This chapter provides a general introduction to the manufacturing of monoclonal antibodies (mAbs) in CHO cells. The worldwide demand for mAb therapeutics is first described along with the structure and function of these molecules. The advantages of CHO cell factories for the recombinant manufacture of mAbs is then explained and alternative expression systems are compared and evaluated. The process of developing a mAb-expressing CHO cell line is explained in detail, with examples of innovations within this process. Finally, the use of synthetic promoters for control of recombinant gene transcription is given particular focus.*

## 1.1 Biopharmaceuticals and monoclonal antibodies

Biopharmaceuticals are biological molecules, usually proteins, that can be used in the treatment of diseases and are manufactured from engineered living cells. The global biopharmaceutical market is ever-growing, with 112 molecules receiving regulatory approval for clinical use in the EU and/or the USA in the period of 2015 - July 2018 (Walsh, 2018). This is in comparison to 60 new approvals across these regions in 2010-2014 and 16 new approvals in 1990-1994, for example. Monoclonal antibodies (mAbs) are the most commonly used biopharmaceuticals. From 2015 to July 2018, 53% of all biopharmaceuticals receiving regulatory approval were mAbs and the total mAb sales value in 2017 was $123 billion (Walsh, 2018). These mAb products are used in the treatment of: various types of cancers; inflammatory conditions such as rheumatoid arthritis, crohn's disease and psoriasis; multiple sclerosis and asthma among many other diseases. The clinical supply of therapeutic mAb molecules is therefore vitally important for patient healthcare across the globe. Consequently, any innovations within the mAb manufacturing process that either speed up the supply of or enable greater yield of therapeutic mAb, helping ultimately to drive down the cost of these drugs, are very highly sought after.

Monoclonal antibodies represent a natural tool of the adaptive immune system that can be recombinantly manufactured on a large scale to enable the treatment of disease. Naturally, B cells of the immune system will have antibodies on their cell

surface that are capable of binding specifically to a single antigen protein that might be displayed by a pathogenic cell (Marshall et al., 2018). This binding event triggers proliferation and differentiation of the B cell into a plasma cell that secretes vast amounts of that same antibody. These plasma cell-secreted antibodies can then bind specifically to their antigen target before recruiting other cells of the immune system to destroy the pathogenic cells bearing that antigen. This specificity of antibody-antigen binding is what makes mAbs such a safe and effective therapeutic molecule, due to the lack of off-target binding that could result in harmful side effects.
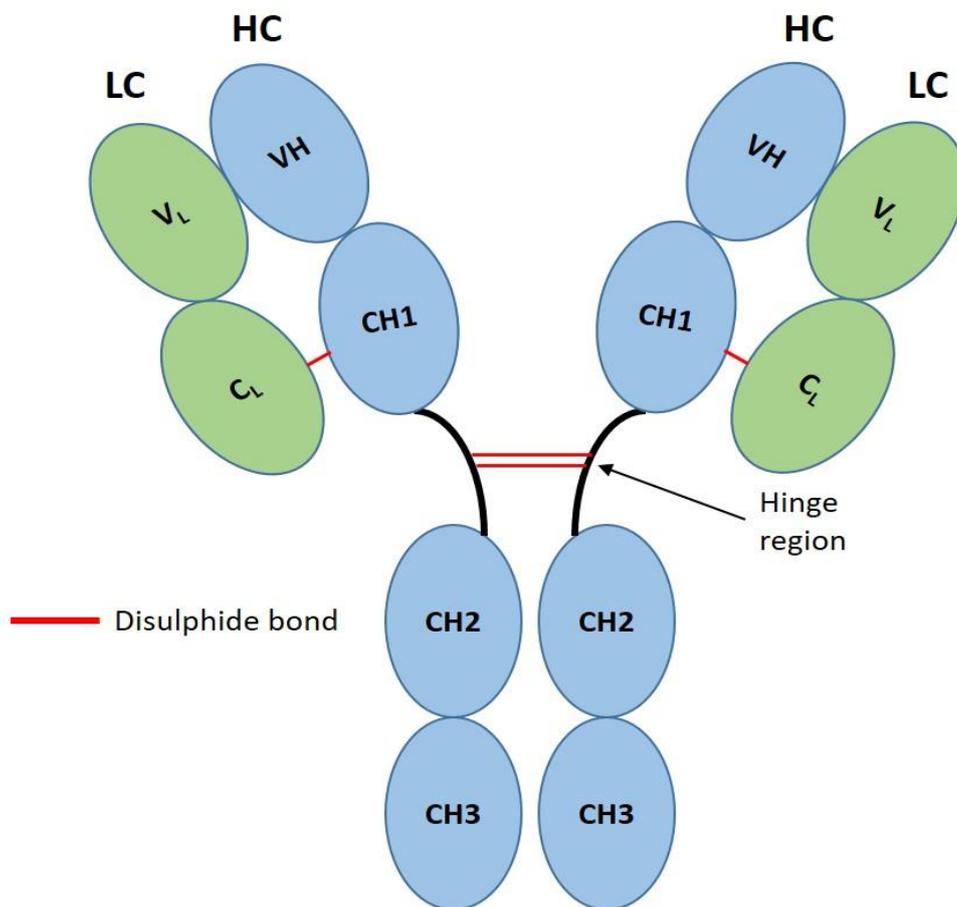


**Figure 1.1 – The structure of an IgG molecule**

An IgG molecule consists of two light chain (LC) polypeptides and two heavy chain (HC) polypeptides joined together by disulphide bonds. The LC contains a variable domain ($V_L$) and a constant domain ($C_L$) while the HC contains one variable domain and three constant domains (VH, CH1, CH2 and CH3 respectively). The two HC polypeptides are joined by disulphide bonds between the hinge regions that connect CH1 and CH2.

In mammals there are five different isotypes of antibody – IgG, IgM, IgA, IgD and IgE, which are distinguished by the sequence and structure of their heavy chain (HC) polypeptides. Of these, IgG has historically been the isotype recombinantly manufactured for therapeutic application. This continues in the modern era, with 27/30 of the novel antibody therapeutics receiving regulatory approval in either the US or the EU between the start of 2018 and June 2020 being IgG or IgG-based products such as IgG-ADC (antibody-drug conjugate) (Kaplon et al., 2020; Kaplon & Reichert, 2018, 2019). IgG molecules are heteromeric structures containing two identical HC polypeptides and two identical light chain (LC) polypeptides. The $\gamma$ HCs of IgG molecules consist of 3 constant domains (CH) and a single variable domain (VH) while the LCs contain a single constant domain ($C_L$) and a single variable domain ($V_L$). IgG molecules contain an internal disulphide bond within each domain of the antibody as well as a set of disulphide bonds connecting the two HCs (known as the hinge region) and a single disulphide bond joins the LC $C_L$ domain with the HC $CH_1$ domain (Liu & May, 2012). There are four subclasses of human IgG that differ mainly in the number of disulphide bonds within the hinge region - IgG1 and IgG4 molecules contain two, IgG2 contains four and IgG3 contains eleven. Figure 1.1 illustrates the structure of an IgG molecule.

## 1.2 CHO cell factories

When recombinantly manufacturing mAb molecules, there are several criteria that the host cell expression system must meet. Firstly, the host cell expression system must be capable of high levels of mAb productivity. The overall yield of mAb from a recombinant manufacturing process is often referred to as the volumetric titre and is measured as the mass of mAb per volume of cell culture, e.g. grams per litre. The volumetric titre of a recombinant manufacturing process is essentially the product of two other metrics - the density of viable cells secreting mAb (viable cell density - VCD) and the cell-specific productivity level (qP). Both VCD and qP are important metrics when evaluating recombinant production processes or specific innovations within these processes. Due to decades of research and development of manufacturing processes, Chinese Hamster Ovary (CHO) cell recombinant protein volumetric titres have gone from ~100mg/L in the late 1980s to often being on the several g/L scale in the modern era (Jayapal et al., 2007).

Secondly, host cells must be capable of secreting mAbs with humanized N-

linked glycosylation profiles. N-linked glycosylation is a post-translational modification (PTM) involving the attachment of a complex oligosaccharide carbohydrate (AKA a glycan) consisting of several sugar molecules to an asparagine amino acid residue of a protein. In IgG mAb molecules, N-linked glycosylation occurs within the $CH_2$ domain of the HC. Different organisms will naturally produce different N-linked glycans. Human N-linked glycans are branched structures beginning with fucose and N-acetylglucosamine sugars attached to the asparagine residue followed by mannose, more N-acetylglucosamine, galactose and terminating with sialic acid (Boune et al., 2020). CHO cells are capable of secreting proteins bearing these human-like N-glycans. Human-like N-glycosylation of IgG mAbs is crucial for their efficacy, stability and lack of immunogenicity when administered into patients.

Finally, for therapeutic recombinant proteins to receive regulatory approval for clinical use it is hugely beneficial if the host cell expression system has a history of safety and an extensive base of knowledge. CHO was used as the host expression system for manufacture of tissue plasminogen activator in 1986, the first recombinantly-produced biopharmaceutical to receive regulatory approval (Kunert & Reinhart, 2016). CHO cells have continued to be the primary source of recombinant protein manufacture ever since, including mAbs - 84% of mAb products that received regulatory approval between January 2014 and July 2018 were manufactured in CHO cell lines (Walsh, 2018). Decades of knowledge on CHO cell function has therefore been acquired as well as a reputation for production of safe mAb products. Later subsections of this chapter describe the process of generating a mAb-expressing CHO cell line.

## 1.3 Alternative expression systems

### 1.3.1 E. coli cell factories

*E. coli* (*Escherichia coli)* offers some advantages as a potential factory for recombinant protein production. Culture of *E. coli* is relatively easy and inexpensive and, as with other bacterial species, *E. coli* is very fast growing compared with eukaryotic expression systems. The doubling time of *E. coli* in standard laboratory growth conditions is around 20 minutes whereas for CHO cells it is around 20-24 hours (Gibson et al., 2018). Moreover, since *E. coli* has long been a model organism for molecular biology studies, many genetic tools have been developed for insertion, deletion or editing of genes in the *E. coli* genome (Jeong et al., 2013). This extensive knowledge and list of techniques makes generating engineered *E. coli* strains relatively straightforward.

Examples of recently approved therapeutic recombinant proteins manufactured in *E. coli* include Myalepta, a synthetic analogue of the hormone leptin; Admelog, a human insulin analogue and Palynziq, a bacterial enzyme (Walsh, 2018). It has also been shown that *E. coli* are capable of recombinantly producing therapeutic antibody fragments that do not require glycosylation such as single chain variable fragments (ScFvs) and antigen binding fragments (Fabs) (Gupta & Shukla, 2017). However, *E. coli* lacks the eukaryotic protein folding and PTM machinery required to produce complex, humanized proteins like full antibody molecules. Synthetic N-linked glycosylation pathways have been engineered into *E.coli*, which initially produced only relatively simple glycan forms (Valderrama-Rincon et al., 2012). More recently, *E.coli* was engineered to produce recombinant proteins bearing more complex, human-like terminally sialylated N-glycans, albeit with low product yield (Zhu et al., 2020). This represents a promising breakthrough but further engineering is required to increase productivity and the product must be stringently and extensively tested to ensure a high level of product quality including a lack of immunogenicity. Only then could such engineered *E.coli* strains be considered as an alternative host to CHO for production of therapeutic molecules requiring complex, human-like N glycosylation such as mAbs.

Another problem with the manufacture of biopharmaceuticals in *E. coli* is contamination of the product with endotoxin, a bacterially-derived lipid that causes a pro-inflammatory response in humans. Removal of endotoxin during the purification process is both difficult and costly. *E. coli* strains have been specially engineered to not produce endotoxin in an effort to circumvent this problem (Mamat et al., 2015).

Finally, another drawback of the use of *E. coli* for recombinant protein production is that often high level heterologous protein expression leads to the protein forming insoluble aggregates known as inclusion bodies. Proteins in inclusion bodies can be solubilised and re-folded but this is a laborious process with generally low yield (Yamaguchi & Miyazaki, 2014). Alternative approaches are to try and reduce formation of inclusion body aggregates through overexpression of molecular chaperones or addition of chemical chaperones to growth media (Malekian et al., 2019).

### 1.3.2 Yeast cell factories

Yeast cells can also be used for the production of therapeutic recombinant proteins. Yeast cells offer robust growth in harsh fermentation conditions, meaning optimisation of culture parameters is less crucial than with the large-scale culture of mammalian

cells. Genes encoding the recombinant protein product can easily be introduced into the yeast cell genome through transformation of a so-called yeast integrative plasmid (Yip) (Spadiut et al., 2014). Many non-mAb therapeutic proteins produced recombinantly in yeast have received regulatory approval for commercial use. For example Fasturtec(EU)/Elitex(USA), a recombinant urate oxidase enzyme used to treat hyperuricemia is produced in the yeast strain *Saccharomyces cerevisiae (S. cerevisiae)* and Semglee, a recombinant modified form of insulin used to treat diabetes is produced in the yeast strain *Pichia pastoris (P. pastoris)* (Walsh, 2018).

Unlike bacterial cells, yeast cells are capable of post-translational modification of proteins including N-linked glycosylation. However the N-glycans that are naturally attached to proteins by yeast cells are high in mannose residues and are therefore not human-like (Spadiut et al., 2014). Potentially therapeutic antibody fragment molecules that lack the N-glycosylated $CH_2$ domain of a full IgG can easily be expressed in yeast. For example an antigen-binding HC variable domain (a so-called "nanobody") was produced recombinantly in *P. pastoris* and in another study a Fab, a ScFv and another single domain nanobody were produced in *S. cerevisiae* (Rahbarizadeh et al., 2006; Wang et al., 2021).

Efforts have been made to engineer yeast cells to produce human-like N-glycans, which may then allow full-length therapeutic mAbs to be manufactured in these cells. For example in *S. cerevisiae* the enzyme α-1,2-mannosidase was expressed recombinantly, which led to production of glycoproteins bearing more human-like N-glycan structures (Chiba et al., 1998). However these N-glycan structures were not *fully* humanised as they lacked several glycan residues added in human cells (or mammalian systems such as CHO) and the number of mannose residues in these N-glycans was also inconsistent. More recently, researchers have attempted to humanise the N-linked glycosylation pathway in *P. pastoris* through co-expression of multiple transgenes encoding specific glycosylation enzymes, successfully leading to the formation of more human-like N-glycans on *P. pastoris*-derived glycoproteins (Jacobs et al., 2009; Nett et al., 2011). One such glycoengineered *P. pastoris* strain was used to produce 1.6g/L of mAb displaying almost fully humanised N-glycans, lacking only the terminal sialic acid residues (Ye et al., 2011). The results of these studies are promising. If it can be shown consistently that yeast cell systems, particularly *P.* pastoris, are capable of robust, high-level production of mAbs that homogenously display human-like N-glycan structures

then these host cell systems may in the future provide a viable alternative to CHO for recombinant mAb manufacture.

### 1.3.3 Transgenic animals

Another alternative method for manufacture of recombinant therapeutic proteins is the use of transgenic animals as production hosts. Transgenic animals are generated by microinjection of a genetic construct containing a transgene coding for the protein of interest into the genome of a fertilised oocyte. The resulting adult animal then has the transgene contained within every cell of its body including its germ cells, enabling the inheritance of recombinant protein production by the offspring (Shepelev et al., 2018). Traditionally this was done by random genomic integration but the differences in expression level between different genomic loci and the variability in the number of transgenes integrated meant that many transgenic animals needed to be generated before one with acceptably high productivity could be identified. More recently, use of site-specific recombinases or targeted genome editing techniques such as CRISPR-Cas9 have enabled site-specific transgene integration, which helps to negate this variability in expression level. The production costs associated with recombinant protein manufacture from a highly expressing transgenic animal are potentially much lower than the production costs of manufacturing the same recombinant molecule through mammalian cell culture (Maksimenko et al., 2013).

The recombinant protein is harvested from transgenic animals non-invasively, most commonly through milk. Accordingly, mammary gland cell-specific promoters are used to drive transgene transcription (Shepelev et al., 2018). Atryn, a recombinantly-produced human antithrombin III and the C1 esterase inhibitor Ruconest are examples of approved therapeutic proteins produced in the milk of transgenic goats and rabbits respectively (Maksimenko et al., 2013). Transgenic mice and goats have also successfully been used to express mAb molecules (Baruah & Belfort, 2004; Sola et al., 1998; Zhang et al., 2009).

Despite the manufacturing of recombinant therapeutic antibodies through transgenic animals showing promise, there are also significant drawbacks. Doubts remain over the specificity of glycosylation and sialylation in transgenic mammals. For example sialic acid in transgenic goats is in a different form to that used for sialylation of human antibodies and this could possibly lead to immunogenicity of transgenic goat-derived antibodies (Raju et al., 2000). Another problem with the use of transgenic

animals is the complication of the purification process by the presence of endogenous antibody that must be removed from the milk. Goat milk contains around 0.3-0.5 mg/ml of endogenous antibody (Maksimenko et al., 2013). Finally, the risk of zoonotic pathogens contaminating the final protein product is a major challenge within this field (Tripathi & Shrivastava, 2019). Given these doubts and complications, no recombinant therapeutic antibodies produced by transgenic animals have yet been given regulatory approval for clinical use.

### 1.3.4 Insect cell factories

Insect cells are another host expression system that can be used to produce recombinant proteins. Baculovirus vector systems are routinely used for the introduction of genes encoding recombinant proteins into insect cells. Some Baculovirus/insect cell systems such as MultiBac can efficiently deliver multiple heterologous genes to insect cells for the production of multiprotein complexes or for synthetic cellular engineering (Gupta et al., 2019). A baculovirus/insect cell system has been used to produce disulfide bonded but unglycosylated human anti-malaria vaccine candidate proteins (Lee et al., 2020; Lee et al., 2017). Recombinant protein therapeutics that do not require human-like N-glycosylation have received regulatory approval for manufacture in insect cells, representing around 1% of all USA/EU approvals for human-use biopharmaceuticals (Yee et al., 2018).

N-glycan structures on proteins produced in insect cells are less complex than those on proteins produced in CHO cells, with insect cell-derived N-glycans lacking elongated branches and terminal sialylation (Yee et al., 2018). The less complex N-glycan structures on recombinant proteins produced in insect cells have been shown to reduce circulation time of the protein in the bloodstream of animal models compared to the same proteins produced in mammals (Grossmann et al., 1997). Attempts have been made to engineer insect cells to produce more human-like N-glycan structures. For example, use of the MultiBac baculovirus system to simultaneously express both light and heavy chains of an antibody as well as two different heterologous glycotransferase enzymes in insect cells resulted in successful production of antibodies bearing more human-like N-glycans (Palmberger et al., 2012). However, these glycans were not fully humanised, lacking the terminal sialylation that would require expression of additional glycotransferases. Moreover, only ~50% of the antibodies produced contained these more human-like N-glycans. This incomplete and inconsistent human-like N-

glycosylation is still far behind the acceptable standard for clinical use and more engineering of the baculovirus/insect cell system is required before it could be considered for production of therapeutic recombinant glycoproteins such as mAbs.

Another problem associated with using baculovirus/insect cells for recombinant protein production is that viral replication leads to host cell lysis which in turn causes contamination of the culture with cellular debris, host cell proteins and baculovirus particles, making the product purification process both costly and challenging (Yee et al., 2018). This has necessitated ongoing engineering efforts to either delay insect cell lysis by overexpressing anti-apoptotic proteins or to prevent active viral replication in insect cells by deleting baculovirus genes involved in viral replication. Additionally, it recently came to light that recombinant protein-producing insect cell lines are often infected with adventitious viruses (Geisler & Jarvis, 2018). This created doubt over the safety of therapeutic proteins produced in insect cells and led to the urgent isolation of insect cell lines that are free of adventitious viruses. These problems along with the current inability of insect cells to generate human-like N-glycan structures are the reasons that they are a relatively minor producer of recombinant protein therapeutics.

### 1.3.5 Plant cell factories

Transgenic plant cells, either in the form of plant cell suspension cultures or whole plants, can be used as expression hosts for recombinant protein production. Plants are an attractive proposition due to their relatively cheap, uncomplicated and easy to scale-up cultivation processes (Tripathi & Shrivastava, 2019). The human enzyme glucocerebrosidase, known commercially as ELELYSO®, received regulatory approval for recombinant production in carrot cells in 2012. ELELYSO® continues to be used to treat type 1 Gaucher's disease and represented the first major regulatory approval of a biopharmaceutical manufactured in a transgenic plant (Fox, 2012). Since then, many other plant-produced biopharmaceuticals have entered into clinical trials (Owczarek et al., 2019).

Although plant cells are capable of N-glycosylating recombinant proteins, these plant N-glycan structures differ from human N-glycan structures through the presence of β1,2-xylose and core α1,3-fucose residues. These plant-derived glycosylation residues have immunogenic potential, with specific antibodies against these residues detected in the sera of humans and laboratory animals (Gomord et al., 2005). These differences in plant and human N-glycan profiles have necessitated efforts to engineer

plant cells to produce N-glycan structures lacking these plant-specific residues. For example, RNA interference (RNAi) technology was used to silence expression of glycotransferase enzymes β1,2-xylosyltransferase and α1,3-fucosyltransferase in a *N.benthamiana* cell line (Strasser et al., 2008). This RNAi-engineered cell line was used to produce anti-West Nile virus mAbs that lacked β1,2-xylose and core α1,3-fucose N-glycan residues and instead contained human-like N-glycan structures with high homogeneity (Lai et al., 2014).

The major drawback of using plant expression systems for biopharmaceutical production is their high manufacturing costs due to low product yield and expensive downstream purification. Antibody titre will seldom exceed 100µg/g fresh material in transgenic whole plants or 100mg/L in tobacco cell suspension culture, for example (Schillberg et al., 2019). In contrast, CHO cell suspension cultures are capable of producing antibody at the g/L scale. It is relatively simple to scale-up transgenic plant cell suspension culture or transgenic whole plant cultivation and thereby overcome these low production yields. However, this increase in host biomass places more pressure on downstream protein extraction and purification processes. These downstream protein product recovery steps are difficult and costly compared to the equivalent procedures for CHO production platforms, especially for transgenic whole plants whereby both insoluble plant debris and soluble host cell protein must be removed. As a result of these expensive purification processes, the cost of manufacturing a human antibody in tobacco plants was estimated to be €1,137/g. By contrast, purified antibodies produced in CHO cells generally have a cost of goods of around US$200/g (Schillberg et al., 2019).

## 1.4 Expression, assembly and secretion of an IgG mAb within a CHO cell

Expression of the component LC and HC polypeptides of an IgG, assembly of these polypeptides into the full IgG structure and secretion of this IgG from the CHO cell is a multi-step process. Initially, following entry of the exogenous DNA construct coding for the IgG LC and HC into the nucleus of the cell, these genes will be transcribed through their promoter sequence to generate mRNA. These mRNA molecules will then be exported from the nucleus through nuclear pores into the cytoplasm where they will be attached to ribosomes for translation.

As with all proteins destined for cellular secretion, LC and HC polypeptides are directed into the endoplasmic reticulum (ER) and this occurs in a co-translational
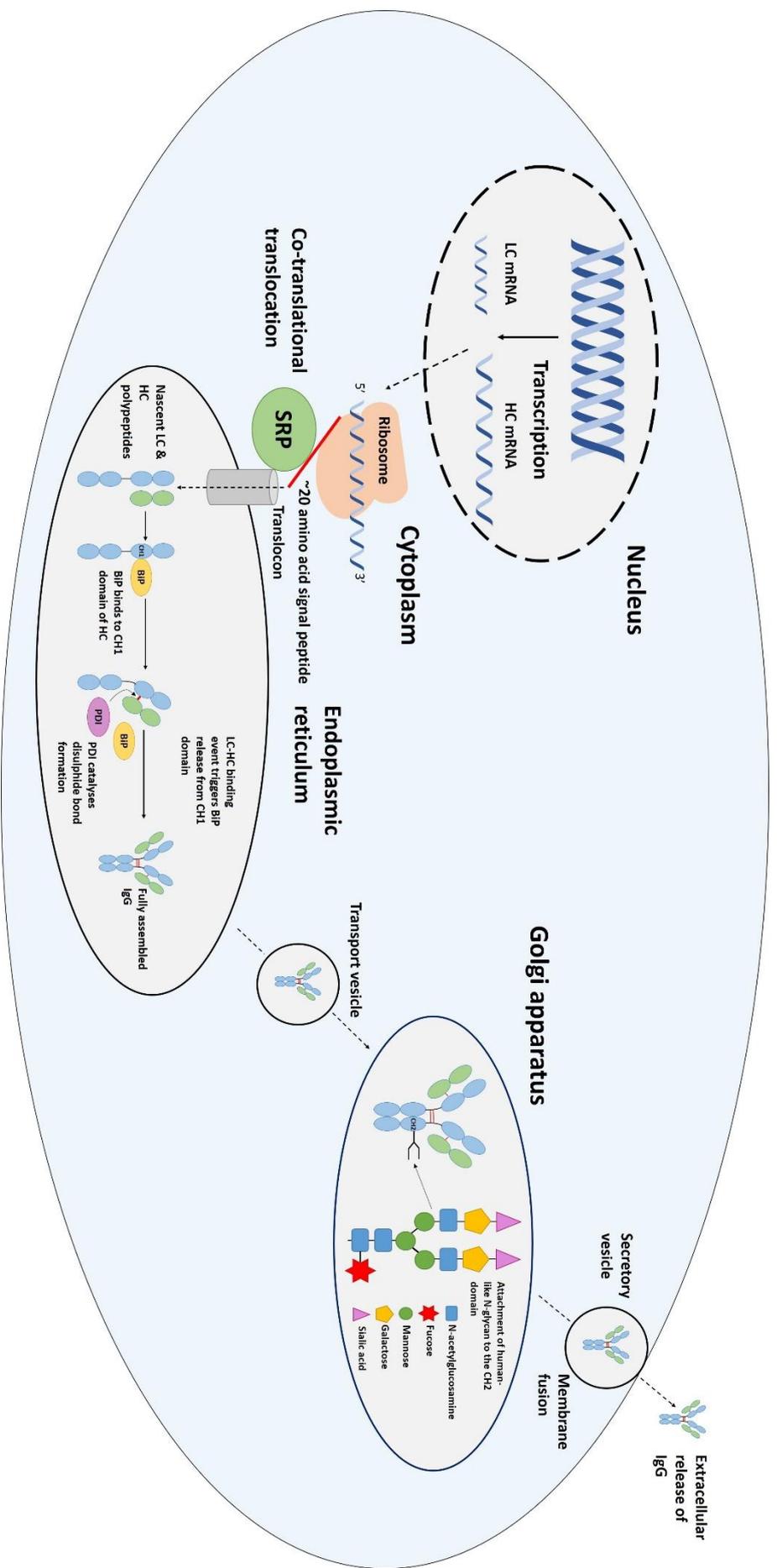
**Figure 1.2 – Production of an IgG mAb in a CHO cell**

A brief overview of the key intracellular events during expression, assembly and secretion of an IgG mAb from a CHO cell is displayed. mRNA transcribed from the LC and HC genes is translated by ribosomes in the cytoplasm and the growing polypeptide chain is fed into the lumen of the ER contranslationally via the translocon channel. In the ER, folding and assembly of the LC and HC polypeptides into a full IgG structure occurs, facilitated by ER resident proteins such as BiP and PDI. N-glycosylation of the IgG molecule begins in the ER before maturation in the Golgi apparatus. The fully assembled, N-glycosylated IgG is then packaged into a secretory vesicle and secreted from the cell.

Nucleus

Transcription

LC mRNA

HC mRNA

Cytoplasm

Co-translational translocation

SRP

Ribosome

5'

3'

~20 amino acid signal peptide

Translocon

Endoplasmic reticulum

Nascent LC & HC polypeptides

BiP binds to CH1 domain of HC

CH1

BiP

LC-HC binding event triggers BiP release from CH1 domain

PDI catalyses disulphide bond formation

PDI

BiP

Fully assembled IgG

Transport vesicle

Golgi apparatus

CH2

Attachment of human-like N-glycan to the CH2 domain

N-acetylglucosamine

Galactose

Mannose

Fucose

Sialic acid

Secretory vesicle

Membrane fusion

Extracellular release of IgG

manner, in a process known as co-translational translocation. Briefly, following the emergence from the ribosome of the ~20 amino acid sequence at the very N-terminus of the polypeptide, referred to as a signal peptide, a complex known as the signal recognition particle (SRP) recognises and binds this signal peptide, pausing translation (Zimmermann et al., 2011). The SRP then targets the mRNA-ribosome-signal peptide complex to an ER membrane channel called a translocon. This leads to the resumption of translation and, concurrently, the passing of the nascent polypeptide chain through the translocon and into the lumen of the ER.

The lumen of the ER is where folding and assembly of the LC and HC polypeptides into a full IgG mAb structure occurs. This is a complex process involving many molecular chaperone and foldase components. One of the key components is the molecular chaperone BiP (binding immunoglobulin protein), around which other components form a complex including other chaperones (e.g. GRP94), co-chaperones (e.g. ERdj proteins) and isomerases (e.g. Cyclophilin B) among others (Jansen et al., 2012). BiP, along with associated proteins, binds to both LC and HC polypeptides immediately upon their entry into the ER lumen and stabilises these folding intermediates (Feige et al., 2010). While BiP will form transient interactions with other IgG domains, the interaction with the $CH_1$ domain of the HC is more stable. BiP is only released from the $CH_1$ domain when a LC binds to the HC, which is an important quality control mechanism in the assembly process (Vanhove et al., 2001). Protein disulphide isomerase (PDI) enzymes are another key component of IgG assembly in the ER as these enzymes catalyse the formation of disulphide bonds (Ellgaard & Ruddock, 2005). Fully assembled IgG molecules possess both intrachain and LC-HC interchain disulphide bonds and these are essential to the structure of the molecule (figure 1.1).

Another important CHO cell process in the production of an IgG is N-linked glycosylation - a post-translational modification in which a complex glycan is attached to an asparagine amino acid residue within the $CH_2$ domain of the HC. The human-like N-glycans produced by CHO cells are branched structures beginning with fucose and N-acetylglucosamine residues attached to the asparagine followed by mannose, more N-acetylglucosamine, galactose and terminating with sialic acid residues (figure 1.2) (Boune et al., 2020). This process begins in the ER, with the transfer of a glycan containing two N-acetylglucosamine residues, nine mannose residues and three glucose residues to the asparagine by the oligosaccharyltransferase complex. This glycan is then trimmed by three other ER-located enzymes: glucosidases I and II and endo-

mannosidase, removing three glucose residues and a mannose residue. The fully assembled IgG with this N-glycan attached is then moved in transport vesicles to the Golgi apparatus for further maturation of the N-glycan. In the cis-Golgi, the enzymes mannosidase 1A and 1B remove another three mannose residues. In the medial-Golgi the enzymes GlcNAc transferase I, mannosidase II, GlcNAc transferase II and α-1,6-fucosyltransferase act sequentially to add an N-acetylglucosamine residue, remove two further mannose residues, add another N-acetylglucosamine residue and add a fucose residue respectively. Finally, in the trans-Golgi, galactosyltransferase adds two galactose residues and then sialyltransferase adds the two terminating sialic acid residues.

The fully assembled, N-glycosylated IgG is then packaged into a secretory vesicle and transported along microtubules to the plasma membrane of the CHO cell (Donovan & Bretscher, 2015). This vesicle will then fuse with the plasma membrane, releasing the IgG contents into the extracellular culture media. Figure 1.2 provides an illustrative overview of the different CHO cell processes involved in IgG production described above.

## 1.5 The IgG mAb manufacturing process in CHO: from gene to product

The following subsections describe the different steps involved in stable CHO cell line development, the industrial scale culturing of stable CHO cell lines and the downstream purification of secreted IgG mAb. The flow chart displayed in figure 1.3 summarises these different steps and includes typical timelines. In total, from IgG-encoding DNA to purified product, the process will typically take around 12 months (Agostinetto et al.,2022; Bandaranayake & Almo, 2014; Kelley, 2020).

### 1.5.1 Expression vector engineering

The vector of choice to transfer genes coding for the therapeutic protein product to the genome of CHO cells is plasmid DNA. Typical plasmid vectors used for IgG mAb expression in industry contain 3 genes capable of being expressed in CHO: that of the LC, the HC and a selection marker. The selection marker is used to select for cells that have successfully integrated the vector into their genome and is discussed later in this chapter.

The light and heavy chain genes will typically be housed within separate expression units on the plasmid, with each expression unit containing multiple genetic

components that help to enable strong, constitutive expression. The first of these components is a promoter, which is positioned upstream of the gene and binds transcription factors in the nucleus to enable a high rate of transcription. The human cytomegalovirus major immediate early promoter (hCMV-MIE, referred to hereafter as hCMV) is the most widely used promoter to control both LC and HC transcription. The promoter from the endogenous Chinese Hamster elongation factor 1α (EF1α) gene is also occasionally used (Brown et al., 2014; Deer & Allison, 2004). More recently, synthetic promoters have been constructed to control recombinant gene transcription in CHO cells (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019). Synthetic promoters will be discussed in detail throughout this manuscript.

Different mammalian cell types will have their own unique composition of tRNA molecules and therefore anti-codons. As a result of this, genes used for mAb expression are typically codon optimised for Chinese hamster in order to maximise translation rate of the mRNA (Wang & Guo, 2020). This codon optimisation is a service routinely offered by gene synthesis companies. Translation rate is also increased by the insertion of an intron into the recombinant LC and HC genes, which aids transport of mRNA from the nucleus to the cytoplasm (Tange et al., 2004).

The 3'UTR (untranslated region) sequence immediately downstream of the open reading frame stop codon, also known as the terminator, is required to both stabilise the mRNA molecule and to encourage dissociation of the RNA polymerase II, thus terminating the transcription process. This role is typically taken up by the Simian Virus 40 (SV40) late polyA sequence in industrial mAb expression vectors, although the terminator sequences from the human growth hormone, bovine growth hormone and Herpes Simplex virus thymidine kinase genes have also been used (Wang & Guo, 2020).

In some cases the LC and HC expression units will be flanked by expression enhancing elements (EEEs) that function to promote the formation of open and transcriptionally active chromatin around the transgenes. This is achieved through the prevention of transcriptionally repressive epigenetic marks such as DNA methylation or histone deacetylation, which promote formation of structurally closed heterochromatin that is largely inaccessible to transcription factors. Examples of EEEs are UCOEs (ubiquitous chromatin opening elements), MARs (matrix attachment regions) and insulators (Nematpour et al., 2017).

Finally, plasmid vectors also contain a bacterial origin of replication and an

antibiotic resistance gene driven by a bacterial promoter. During amplification of a plasmid vector in *E.coli* prior to transfection into CHO, these two components enable propagation of the plasmid and selection of *E.coli* containing the plasmid respectively (Wang & Guo, 2020).

### 1.5.2 Transfection of expression vector

After preparing the plasmid vector for expression of the recombinant mAb genes, the next step in the process is transfection of this vector into the CHO host cell line. Several different transfection methods are available to introduce plasmid DNA into the host cell line. Chemical reagents such as the cationic lipid Lipofectamine (ThermoFisher Scientific) or the cationic polymer polyethylenimine (PEI) form complexes with the negatively charged DNA and these chemical-DNA complexes enter the cell via endocytosis (Islam et al., 2014; Karmali & Chaudhuri, 2007). Electroporation is another commonly used method. During electroporation, an electrical pulse is applied to cells which disrupts the cell membrane potential and generates temporary pores in the membrane through which plasmid can enter the cell (Nakamura & Funahashi, 2013). The Nucleofector technology (Lonza) is a specialised form of electroporation that can offer higher transfection efficiencies than traditional electroporation due to the use of pre-optimised electrical parameters and cell type specific buffer solutions (Maasho et al., 2004). Transfection procedures utilising these methods can be classified into two main types: transient or stable.

### 1.5.2.1 Transient transfection of expression vector

In transient transfection, the plasmid vector is generally not linearized (i.e. is left circular) and does not need to code for a selectable marker protein, nor is any selection agent added to the transiently transfected culture to select for cells that have incorporated the expression vector into their genome. As a result, during transient expression procedures the plasmid is mostly maintained extrachromosomally. Transiently transfected cultures will express for several days only, during which plasmid is gradually diluted within the cell population as it is lost during cell division.

Transient transfection enables inexpensive, short-term production of relatively small amounts of developmental material for pre-clinical efficacy and toxicology studies (Daramola et al., 2014). The transient transfection process also enables evaluation of process parameters and potential innovations before committing to an expensive, long-term stable cell line generation process. For example, it may become

clear from transient expression that use of one particular variant of a vector component, such as a promoter or a terminator sequence, leads to superior product titre compared to use of another variant. Similarly, transient expression of a gene may reveal that this gene enhances mAb productivity, making it a potential new target for stable CHO host cell line engineering.

### 1.5.2.2 Stable transfection of expression vector

During stable transfection procedures, on the other hand, plasmid vectors are generally linearized using a restriction enzyme before transfection in order to increase the efficiency of genomic insertion. This stably transfected plasmid will encode a selectable marker protein and transfected cells are cultured with a selection agent added to the culture medium. This way, only those cells that are expressing the selectable marker over a prolonged period and therefore must have incorporated the vector into their genome, are capable of dividing. Selectable markers and selection agents are discussed in detail in the next subsection.

### 1.5.3 Selection of cells that have incorporated the expression vector into their genome

A selection system is required to select for cells within the transfected population that have incorporated the expression vector construct into transcriptionally active sites within their genome. One such system that is commonly used in the biopharmaceutical industry is the glutamine synthetase (GS) selection system (Cockett et al., 1990). Glutamine synthetase is a key enzyme in the synthesis of glutamine, an amino acid that is critical for cell propagation within culture. Therefore, if cells transfected with a plasmid vector containing the GS selection marker gene are cultured in medium lacking glutamine then only those cells in which the plasmid has inserted into a transcriptionally active genomic location will be capable of efficient propagation. In order to increase the selective pressure, i.e. to increase the GS expression level required for a cell to propagate, methionine sulfoximine (MSX) is added to the transfected cell culture (Harfst & Johnstone, 1992). MSX is a specific inhibitor of GS, meaning that following its addition to culture, a cell must produce more GS to overcome this inhibitory effect. The addition of MSX is often described as a gene amplification event, whereby the cell population will become enriched with cells that have duplicated the vector-derived GS gene within their genome. Crucially, because the recombinant mAb genes are linked to the selection marker gene by proximity, any duplication event is extremely likely to also

duplicate these mAb genes. This will ultimately lead to that cell being more productive.

Similar, alternative selection systems exist such as the dihydrofolate reductase (DHFR) system (Kaufman & Sharp, 1982). DHFR catalyses the conversion of dihydrofolate to tetrahydrofolate, an essential co-factor in the synthesis of nucleotides and certain amino acids. In this case the selection agent used to increase selective pressure is methotrexate (MTX), through binding to and inhibiting DHFR. Addition of MTX to cells transfected with a DHFR-containing plasmid vector leads to gene amplification as with MSX in the GS selection system. Many biopharmaceutical companies favour the GS system over DHFR and doubts over the efficacy of the DHFR-MTX system have been discovered. For example it has been shown that transfected cells can adapt to be resistant to MTX in ways other than through amplifying DHFR expression, such as through reducing cell permeability to MTX or through DHFR structure-altering mutations that reduce affinity to MTX (Flintoff & Essani, 1980; Kim et al., 2001).

### 1.5.4 Single cell isolation and screening

Once a pool of stably expressing recombinant CHO cells has been established, single cells must be isolated from this pool before screening of the subsequent monoclonally-derived populations. Production of biopharmaceuticals from monoclonally-derived populations is a requirement for approval from drug regulatory bodies (Priola et al., 2016). This is because the polyclonal pool of transfected cells will be highly heterogeneous both phenotypically and genotypically, due to both the pre-existing heterogeneity within a host CHO cell population before transfection as well as the random and highly variable genomic integration of the mAb expression vector (Lewis et al., 2013; Pilbrough et al., 2009). Use of such a heterogeneous pool in a mAb manufacturing process would lead to changeable and unpredictable results in terms of product titre and quality.

The monoclonally-derived cell line chosen for manufacture must meet several important criteria, such as high productivity, high growth rate, a stable level of productivity over time and high product quality (Le et al., 2015). Clones meeting all of these criteria at once can be rare events within the heterogeneous transfected cell population.

Traditionally, single cells were isolated by multiple rounds of limiting dilution, however this is a laborious and low-throughput process (Priola et al., 2016). Moreover,

this technique does not allow any selective isolation of high producers over low producers, which would help to reduce the number of clones that need to be screened. High-throughput methods capable of more selective isolation were therefore developed such as fluorescence-activated cell sorting (FACS) based protocols whereby affinity matrices are used to capture the secreted protein at the cell surface and this cell surface protein is then detected and approximately quantified through use of fluorescently-labelled antibodies (Black et al., 2011). Similarly, the ClonePix system allows fully automated clone picking through culture of cells on semi-solid medium, which leads to the formation of spatially distinct clonal colonies (Mann, 2007). Secreted protein from these clonal colonies is immobilised on the semi-solid medium and can be quantified with a detection agent. The ClonePix system then selectively picks the more productive clones for further evaluation while discarding the less productive clones.

Once monoclonally-derived populations have been established, thousands of them are usually screened in a low volume, high throughput plate format (Le et al., 2015). A select number of the best performing monoclonally-derived populations will then be screened in larger scale formats such as shake flasks and mini bioreactors. Finally, a select panel of the best performers will be assessed for production stability over many rounds of subculture before a single manufacturing cell line is chosen (Kim et al., 2011). Typically, establishment of a monoclonally-derived population following single cell isolation will take ~ 3 weeks, clone evaluation across different culture formats and scales will take ~ 15 weeks and production stability of cell lines will be assessed over ~12 weeks (Noh et al., 2018; Povey et al., 2014).

### 1.5.5 Bioreactor fed batch culture

Once a single monoclonally-derived cell line suitable for manufacturing of a therapeutic mAb has been identified, the culture of this cell line must be steadily scaled up in volume to eventually provide enough biomass for inoculation of an industrial sized bioreactor (often 20,000 litres). Despite extensive research and innovation within areas of the biomanufacturing process described in above subsections, improvements in the fed batch bioreactor culture process have been the primary source of increases in product titre over the past 30 years. Specifically, improvements have been made in control of bioreactor operating parameters, design of growth media and feeding strategies. Critical bioreactor operating parameters that are controlled throughout the culture include agitation, temperature, dissolved oxygen level and pH (O'Flaherty et al.,

2020).

Despite the common use of fetal bovine serum (FBS) in cell culture media for other applications due to it being rich in components such as vitamins, minerals, hormones and growth factors, serum-free medium is used in recombinant therapeutic protein manufacture. This is due to the batch-to-batch variability of FBS reducing predictability of performance as well as it representing a potential source of contamination (Gstraunthaler, 2003). During the fed batch bioreactor culture, feed solutions are added that may contain amino acids, apoptosis inhibitors, growth factors such as Insulin and IGF-1, histone deacetylase inhibitors such as sodium butyrate and valproic acid, glucose and nucleoside inhibitors amongst other components (Adamson & Walum, 2007; Backliwal et al., 2008; Chen et al., 2011; Gagnon et al., 2011; Mulukutla et al., 2017; Takagi et al., 2017; Zanghi et al., 2000). The concentrations of these components and timing of feed addition is carefully optimised in an effort to provide cells with key nutrients and enable high level recombinant mAb production.

Although fed batch culture set ups remain the most widely used for biomanufacturing, there is substantial interest in the use of perfusion culture to enable continuous manufacturing processes (O'Flaherty et al., 2020). In perfusion culture, fresh growth media is repeatedly pumped into the bioreactor while spent media is removed. Cells within the spent media are separated out through porous membranes and pumped back into the culture, while recombinant protein product is continuously purified. Thus, constant maintenance of key nutrient availability through addition of fresh media and removal of toxic by-products within spent media such as lactate and ammonia allows continuous biomass accumulation with a steady flow of product titre. Cell densities reaching over $6 \times 10^7$ cells/ml have been achieved through perfusion culture (Chotteau et al., 2015; Karst et al., 2017). Moreover, the reduced residence time of the recombinant protein product in culture reduces the chance of product quality variation deriving from cell-secreted enzymes affecting glycan structures on the product. However, the development of downstream purification processes capable of handling the large spent media volumes continually being produced is ongoing.

### 1.5.6 Downstream purification

Once the fed batch culture of recombinant CHO cells is complete, the mAb must then be purified from culture medium. There is a standard purification workflow that manufacturers generally follow, starting with removal of cells and large cell debris from

```
        ┌──────────────────────────────────────────────────────────┐
        │ Construction of plasmid vector encoding IgG LC & HC genes │
        └──────────────────────────────────────────────────────────┘

• Stable transfection
• Selection of transfectants          ~ 2 – 4 weeks
  stably expressing a selection
  marker protein
                                              │
                                              ▼
        ┌──────────────────────────────────────────────────────────┐
        │          Heterogeneous stably expressing cell pool        │
        └──────────────────────────────────────────────────────────┘

• Single cell isolation and
  expansion (~ 3 weeks)
• Evaluation of monoclonally-          ~ 30 weeks
  derived cell lines (~ 15 weeks)
• Evaluation of production
  stability of lead cell lines (~ 12
  weeks)
                                              │
                                              ▼
        ┌──────────────────────────────────────────────────────────┐
        │       Monoclonally-derived, IgG-expressing cell line      │
        └──────────────────────────────────────────────────────────┘

• Development and optimisation
  of tailored bioreactor fed           ~ 6-7 months
  batch culture process and
  downstream IgG purification
  procedure
                                              │
                                              ▼
        ┌──────────────────────────────────────────────────────────┐
        │              Bioreactor scale IgG manufacture             │
        └──────────────────────────────────────────────────────────┘
```

**Figure 1.3 – A typical timeline for generating purified IgG mAb from CHO cells**

The flow chart provides a brief overview of the steps involved in stable CHO cell line development and IgG manufacture. The amount of time typically taken for each step is shown. Development of bioreactor fed batch culture processes and downstream purification procedures may begin before evaluation of monoclonally-derived cell lines is completed. Typically, the entire process takes around 12 months.

the bioreactor culture medium by centrifugation and filtration (Chahar et al., 2020). The mAb is then captured from this clarified medium by protein A affinity chromatography. Protein A, a recombinantly-generated protein of bacterial origin, binds with very high specificity to the HC of monoclonal antibodies. The mAb can be eluted from the protein A column using an acidic buffer of ~ pH 3-4. Storage of the mAb in this low pH buffer before subsequent purification steps also ensures inactivation of any contaminating

viruses. Next, any remaining impurities such as host cell DNA/protein are removed through ion exchange chromatography. Gel filtration chromatography or hydrophobic interaction chromatography may also be used to remove impurities depending on the precise physiochemical properties of the mAb product. Finally, the product is concentrated and transferred to its formulation buffer through diafiltration.

### 1.6 Design of synthetic mammalian promoters

### 1.6.1 Mammalian promoter function

In mammalian genomes, transcriptional regulation is mediated by two adjacent segments of DNA (Haberle & Stark, 2018). The first is the core or minimal promoter region (referred to in this manuscript as the core promoter). The core promoter is a short DNA sequence located close to the transcription start site and is bound by the transcription initiation complex, which consists of six so-called general transcription factors and RNA polymerase II, the enzyme responsible for generating the mRNA strand. For example, TATA box and BRE are two common core promoter elements and are bound by the general transcription factors TBP and TFIIB respectively.

Upstream of the core promoter is a proximal promoter region or an enhancer (referred to in this manuscript as a proximal promoter region). The proximal promoter contains binding sites for specific transcription factors within the cell. After binding to the proximal promoter, these specific transcription factors will then interact with specific co-factors - transactivators will interact with co-activators while transsilencers will interact with co-repressors. These co-factors can then regulate the frequency of binding of the transcription initiation complex to the core promoter and thus regulate the level of gene transcription. Synthetic promoter engineering efforts have focused on the creation of new proximal promoter regions. Attempts to engineer core promoters have been shown to generate minimal improvement compared to use of natural core promoters commonly used in mammalian transgene expression (Johari et al., 2019).

### 1.6.2 Design strategy 1: Synthetic transcription factor – promoter pairs

The use of synthetic transcription factors that bind to user-defined DNA sequences is one strategy that has been used to control transgene transcription in mammalian cells. For example the Cas9 endonuclease, which can be targeted to specific genomic sequences via a co-expressed guide RNA, has been utilised in creation of a synthetic transcription factor through fusion of a nuclease-inactive Cas9 with a transactivator

(Chavez et al., 2015). Similarly, transcription activator-like effectors (TALEs) are proteins found naturally in plant pathogenic bacteria where they activate transcription of specific plant genes (Moscou & Bogdanove, 2009). These TALEs can be engineered to bind almost any target DNA sequence and they have been used to activate transcription of specific endogenous genes in human cells *in vitro* (Perez-Pinera et al., 2013).

Enabling transgene transcription through these fully synthetic transcription factor – promoter pairs could offer some advantages in recombinant biopharmaceutical production. Firstly, there would be no burden placed on the host cell through the use of endogenous transcription factors, which could otherwise create unpredictable changes in the natural transcriptome of the host cell. Second, these synthetic transcription factor – promoter pairs should function similarly across different mammalian cell types as their output is not dependent on the particular complement of host transcription factors, making them a transcriptional control system with flexible usage. However, high level co-expression of another recombinant protein (i.e. the synthetic transcription factor) alongside the recombinant therapeutic protein would require exceptional protein production capacity and host cells possessing this capacity would likely be extremely rare or non-existent.

The use of synthetic transcription factors together with insertion of their specific binding sites upstream of a transgene can also facilitate inducible transgene expression. For example, Rössger et al constructed a chimeric transcription factor consisting of a transsilencing domain fused to a bile acid-responsive repressive element (Rössger et al., 2014). This transcription factor successfully silenced activity of a promoter containing its binding site until the addition of bile acid, which then repressed the silencing and switched on target gene expression. Such inducible expression systems could in theory be useful for recombinant biopharmaceutical manufacture. For example, transgene expression could be switched off during the biomass accumulation phase of culture and then switched on once biomass has sufficiently accumulated. However, the compound required to induce transgene expression would need to be compatible with efficient cell growth in culture and would need to meet all regulatory safety standards. Finding such a compound may be challenging.

### 1.6.3 Design strategy 2: Mutation of natural promoter sequences

Another method of creating synthetic promoters is to introduce mutations into natural promoters that are known already to function in mammalian cells. For example, Ferreira

and colleagues introduced mutations directly into the hCMV and EF1α promoters, two promoters commonly used to drive recombinant gene transcription in mammalian cells (Ferreira et al., 2011). Mutations were introduced by both site-directed mutagenesis within the TATA and CAAT box core promoter elements of the promoters as well as by random mutagenesis. This led to generation of synthetic promoters displaying a 40-fold range in expression level. Similarly the SV40 promoter, another promoter often used to drive recombinant gene transcription in mammalian cells, has also been mutated by deleting 5' portions of the sequence (Fan et al., 2013). As intended, these deletions created new synthetic promoters with weaker transcriptional activity compared to the SV40 promoter.

While mutating existing, naturally evolved promoters can create new synthetic promoters of varying transcriptional strengths, this strategy does not enable the precise design of synthetic promoters with pre-determined and predictable transcriptional strength. In addition, mutating a promoter sequence will most likely reduce transcriptional output of that promoter, meaning this strategy is unlikely to enable the creation of synthetic promoters that are more transcriptionally active than commonly used natural promoters.

### 1.6.4 Design strategy 3: Joining of DNA fragments

Synthetic mammalian promoters have also been constructed through the rational combination of DNA fragments. For example, Ogawa et al constructed synthetic promoters by randomly ligating synthetic oligonucleotides containing transcription factor binding sites (Ogawa et al., 2007). Two out of the eleven synthetic promoters displayed greater transcriptional activity than a CMV promoter in human PC-3 cells. When these two synthetic promoters were introduced into various other human cell types they displayed highly variable activity levels, likely due to the variation in transcription factor expression among the different cell types. In another study, short sequence motifs shown to be enriched in highly active human promoters compared to in the genome were overlapped to create synthetic promoters that displayed transcriptional activity in different mammalian cell lines (Grabherr et al., 2011). A cruder approach was taken by Schlabach et al, who took every possible 10-mer DNA sequence, assembled each one in to 100bp 10-fold tandem repeats and then screened these 100bp synthetic elements for transcriptional activity across multiple mammalian cell lines (Schlabach et al., 2010). Similar to the results of Ogawa et al, screening revealed that

some of these synthetic elements were capable of producing transcriptional activity comparable to that of a CMV promoter but the performance of these elements varied hugely between different cell lines. Finally, in another study a chimeric promoter was created for use in mammalian cells by fusing fragments from the SV40 promoter, the human β actin promoter and the human ubiquitin C promoter (Tornøe et al., 2002). Randomisation of the sequences thought to be separating transcription factor binding sites within this chimeric promoter resulted in a synthetic promoter library spanning a 10-fold range in transcriptional activity. Although these studies are examples of successful creation of transcriptionally active synthetic promoters, these promoters were broadly designed to function across different mammalian cells. They were not specifically designed to utilise the unique repertoire of transcription factors within a particular cell line e.g. a CHO cell.

### 1.6.5 Design strategy 4: TFRE 'building block' design approach

Recently, a new strategy has been used for the construction of CHO cell synthetic promoters (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019). This strategy involves the screening of individual transcription factor regulatory element (TFRE) sequence blocks (binding sites for specific transcription factors) for transcriptional activity in CHO cells. Information about the relative transcriptional activity given by individual TFRE blocks can then be used to design synthetic proximal promoters in which different TFRE blocks are combined to achieve a specific overall level of promoter activity. This strategy therefore exploits the unique complement of transcription factors expressed in CHO cells and has been shown to enable the precise design of synthetic promoters with pre-determined levels of transcriptional output over a wide range (over 10-fold). This TFRE 'building block' design approach is discussed in more detail in chapter 3 of this manuscript.

### 1.7 Project aims

This chapter has provided an introduction to the process of manufacturing mAb molecules in CHO cells. The work described within this manuscript aimed to generate solutions that improve this process by either increasing volumetric mAb titres or by reducing the development time of stably expressing CHO cell pools. Chapter 3 describes the design of CHO cell synthetic promoters using the TFRE 'building block' design approach while chapters 4 and 5 describe the use of these synthetic promoter

sequences to control the transcription of mAb LC/HC genes (chapter 4) and the GS selection marker gene (chapter 5) in a stable expression mode. Chapter 6 describes high-throughput transient transfection screening of genes acting within the ER for their effect on mAb production, with the aim of discovering novel targets for CHO host cell engineering.

# Chapter 2

# Materials and Methods

*This chapter details the materials and methods that were used in the experiments described throughout this manuscript.*

### 2.1 Expression vector construction

Restriction endonuclease cut fragments were purified through separation on an agarose gel followed by extraction from the agarose gel using the Qiagen QIAquick gel extraction kit. All restriction endonucleases used were supplied by New England Biolabs. To ligate fragments with complementary sticky ends, either the rapid DNA Ligation kit (Merck) or the NEB Quick Ligase (New England Biolabs) were used. Before transfection of plasmids, correct assembly was confirmed by Sanger sequencing.

### 2.1.1 Single gene LC and HC expression vectors

Both the LC and HC-expressing plasmids for mAb 2 had an expression unit with the layout displayed in Figure 2.1. hCMV proximal promoters were contained between EcoRI and NheI restriction sites in both plasmids. The LC expression unit had a lambda LC poly A between SbfI and NotI sites while the HC expression unit had an SV40 poly A in the same position. The mAb 1 LC-expressing plasmid was made by ligating an NheI-SbfI cut backbone, containing a lambda LC poly A between SbfI and NotI sites of the expression unit, with a NheI-SbfI cut fragment containing the CMV core promoter along with the mAb 1 LC ORF. The mAb 1 HC-expressing plasmid was made by ligating a BamHI-NotI cut backbone with a BamHI-NotI cut fragment encompassing the entire mAb 1 HC expression unit.

### 2.1.2 Stable expression vectors containing LC and HC synthetic promoters

LC/HC dual gene stable expression vectors were created by three-way Golden Gate assembly of two entry plasmids with a GS-containing destination plasmid. Both LC and HC entry plasmids contained expression units with the layout displayed in figure 2.1. Plasmids containing the synthetic proximal promoter sequences shown in appendix table A2 flanked by EcoRI and NheI recognition sites were synthesised (GeneArt, ThermoFisher Scientific) and digested with these restriction enzymes. These EcoRI-NheI cut fragments were then purified and ligated between EcoRI and NheI sites of mAb 1 LC/HC entry plasmids, replacing the existing hCMV proximal promoters.
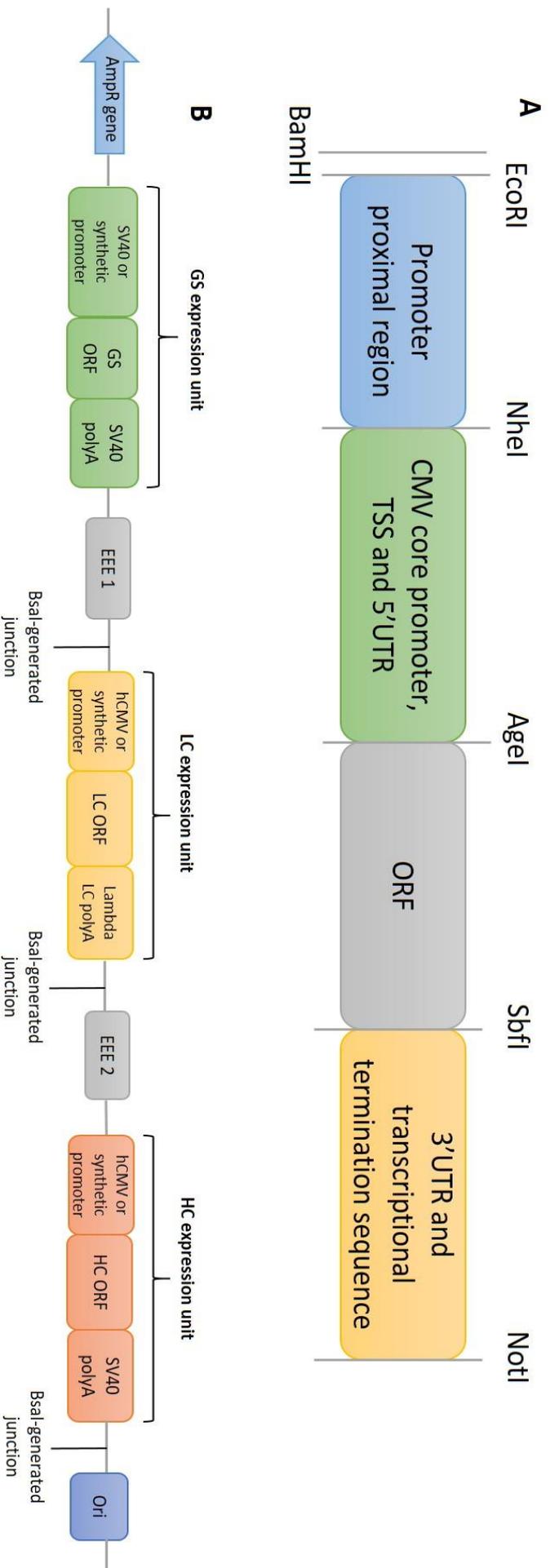
**Figure 2.1 - Expression unit layout and linearised view of GS-containing dual gene vector**

**A)** displays the standardised layout of the gene expression unit used in this work. Gene expression components listed can be changed for alternative components by generation of sticky ends using the restriction endonucleases shown. **B)** shows a linearised view of the stable expression vector layout used in this work. Abbreviations: CMV – cytomegalovirus; hCMV – human cytomegalovirus; TSS – transcription start site; UTR – untranslated region; ORF – open reading frame; GS – glutamine synthetase; SV40 – simian virus 40; EEE – expression enhancing element; polyA – polyadenylation signal; Ori – bacterial origin of replication.

Next, a fragment containing an expression enhancing element (EEE) flanked by ASCI-generated overhangs was ligated with ASCI-linearised and dephosphorylated (Antarctic Phosphatase, New England Biolabs) mAb 1 HC entry plasmids containing either synthetic proximal promoters or the hCMV proximal promoter between EcoRI and NheI sites of the expression unit. The mAb 2 LC and HC entry plasmids were created by using AgeI and SbfI to replace the LC/HC ORFs in mAb 1 entry plasmids with the mAb 2 LC/HC ORFs.

Three-way Golden Gate assembly reactions (Golden Gate Assembly Kit (BsaI-HFv2), New England Biolabs) were performed with LC and HC entry plasmids as well as a destination plasmid containing the GS selection marker expression unit and a second EEE sequence. BsaI-generated overhangs were designed to enable ordered assembly of the LC expression unit to the 5' of the HC expression unit. Figure 2.1 shows a linearised view of the golden gate assembly product.

### 2.1.3 Stable expression vectors containing GS synthetic promoters

The EEE and GS expression unit-containing destination plasmid described above was mutated using the Q5 Site-directed mutagenesis kit (New England Biolabs) to insert BclI, AsiSI and PacI restriction endonuclease recognition sites into the SV40 promoter sequence. The PacI site was inserted to the 3' of the SV40 promoter transcription start site (TSS). The BclI site was inserted to the 5' of the start of the SV40 promoter sequence while the AsiSI site was inserted between BclI and PacI.

A sequence block containing the CMV core promoter and TSS was PCR amplified with primers containing AsiSI (forward primer) and PacI (reverse primer) recognition sites. The destination plasmid was cut with AsiSI and PacI and treated with Antarctic phosphatase (New England Biolabs) to prevent re-circularisation via complementary AsiSI- and PacI-generated overhangs. The CMV core PCR fragment was then cut with AsiSI and PacI and ligated into this phosphatase-treated backbone.

The synthetic proximal promoters GS1 - GS6, listed in appendix table A3, were synthesised as single-stranded oligonucleotides (Integrated DNA Technologies) flanked by BclI and AsiSI recognition sites. These sequences were PCR amplified to make them double-stranded, cut with BclI and AsiSI and ligated into CMV core-containing, BclI and AsiSI cut destination plasmid backbones.

Sequence blocks containing the synthetic proximal promoters SynSV40_1 and

SynSV40_2, listed in appendix table A3, upstream of a CMV core sequence with an AsiSI site inbetween, a BclI site to the 5' and a PacI site to the 3' were synthesised (GeneArt, ThermoFisher Scientific). These sequence blocks were cut with BclI and PacI and ligated into BclI and PacI cut destination plasmid backbones.

These new destination plasmids, containing different synthetic promoters within the GS expression unit, were then used in golden gate assembly reactions to insert LC and HC expression units.

### 2.1.4 Endoplasmic reticulum effector gene expression vectors

Human protein coding ORFs of genes listed in table 6.1 with a common Kozak sequence (AGGCCACC) inserted immediately to the 5' of the ATG start codon were synthesised flanked by AgeI and SbfI recognition sites (GeneArt, ThermoFisher Scientific). These two restriction enzymes were then used to clone each of these ORFs into a simple vector backbone containing an hCMV promoter to drive transgene transcription.

### 2.2 Transformation and plasmid preparation

Ampicillin was used throughout this work at a concentration of 100µg/ml for selective amplification of plasmid DNA in *E.coli*. Plasmid was transformed in to one of Invitrogen Subcloning Efficiency DH5α (ThermoFisher Scientific), NEB 5-alpha (New England Biolabs), NEB 10-beta (New England Biolabs), Dam⁻/Dcm⁻ competent *E.coli* (New England Biolabs) or Zymo Mix and Go DH5 Alpha (Zymo Research) competent cells according to manufacturer's protocols and cells were plated onto LB agar with ampicillin for overnight incubation at 37°C. Overnight LB broth with ampicillin cultures of transformed *E.coli* colonies were prepared and one of the Qiagen QIAprep Spin Miniprep kit; Plasmid Plus Midi Kit or Plasmid Plus Maxi Kit (Qiagen) was then used to purify plasmid DNA according to manufacturer's instructions. Plasmids were eluted in nuclease-free water. Nanodrop One or Nanodrop 2000 (ThermoFisher Scientific) spectrophotometers were used to determine concentration and purity of plasmid preparations.

### 2.3 Transient 96 well Nucleocuvette plate transfections

The Lonza 96 well Nucleocuvette plate and 96 well Shuttle device were used for high-throughput transient transfection of CHO cells. $2 \times 10^6$ viable cells and varying amounts of plasmid DNA were combined in individual wells of the 96 well

nucleocuvette plate. The total volume of DNA-cell mixture in each well of the 96 well nucleocuvette plate was 20µl with 2µl of this being the DNA solution. To enable efficient mixing and transfer of this DNA-cell mixture to wells of the nucleocuvette plate, the mixture was first set up at an excess in a separate 96 well plate. To prepare the DNA-cell solution, viable cell density of exponential phase cultures was first measured on a Vi-cell XR Cell Viability Analyzer (Beckman Coulter) before a cell pellet was prepared by centrifugation, resuspended in the transfection solution and combined with the plasmid DNA to be transfected. The pulse settings applied to the 96 well Nucleocuvette plate were: Part 1 – FF, Part 2 – 158. After nucleofection, cells were seeded into Nunc 96 deep well plates (ThermoFisher) and incubated for between 24 – 120 hours. To measure mAb titre from these high-throughput transient transfections, either the RD-Biotech FastELISA Human Immunoglobulin Quantification Kit or an HPLC protein A method was used.

## 2.4 Generation of stably expressing CHO pools

### 2.4.1 Preparation of DNA for stable transfection

100µg of plasmid was linearised overnight with either PvuI (Roche) or SwaI (New England Biolabs). 5PRIME 2ml Phase Lock Gel tubes (Heavy) (Quantabio) were used for phase lock gel extraction of linearised plasmid. Two successive rounds of organic-aqueous phase separation were performed, with phenol:chloroform and chloroform used as the organic extraction reagents. DNA was then ethanol precipitated, air dried and resuspended in nuclease-free water.

### 2.4.2 Transfection of stable expression vectors and MSX selection of pools

Linearised plasmid DNA was transfected into CHO host cells using the Nucleofector 2b device (Lonza). Cells were resuspended in Nucleofector solution V (Lonza) and combined with linearised plasmid DNA in the cuvettes provided before the electrical pulse was applied to this cell/DNA solution via the 2b device. Cuvettes of transfected cells were combined into CD-CHO (ThermoFisher Scientific) in a T175cm$^2$ flask (Corning), which represented one pool. 24 hours post-transfection MSX was added to each pool before static incubation to allow recovery. The % viability and VCD of MSX-selected pools was monitored using the Vi-cell XR Cell Viability Analyzer (Beckman Coulter). Once pools had recovered they were then transferred to a shaking culture set-up in 125ml Erlenmeyer flasks (Corning).

## 2.5 Fed batch overgrow

45ml FBOG cultures were set up in 250ml vented Erlenmeyer flasks (Corning). Cultures were maintained for either 11 or 12 days through regular addition of two proprietary feed solutions. Glucose concentration in cultures was maintained >0 and kept consistent between cultures by measuring regularly using a YSI 2900 Biochemistry Analyzer (Xylem) and adding glucose solution to the cultures as appropriate. Supernatant samples were taken for antibody titre analysis by an HPLC protein A method. Cell density measurements were taken throughout culture using the Vi-cell XR Cell Viability Analyzer (Beckman Coulter).

## 2.6 qRT-PCR

### 2.6.1 RNA extraction and reverse transcription

Cell pellets containing $5 \times 10^6$ cells taken from high viability cultures were prepared by centrifugation. RNA was extracted from cell pellets using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. For reverse transcription (RT) of RNA into cDNA, the SuperScript IV First-Strand Synthesis System (ThermoFisher Scientific) was used with oligo d(T) primers according to the manufacturer's instructions. Negative control RT reactions in which the reverse transcriptase enzyme was replaced by water were also run.

### 2.6.2 Measurement of mRNA levels from stable pools

qRT-PCR was performed using TaqMan probes, TaqMan Universal Master Mix II and a QuantStudio 12K Flex Real-Time PCR System (ThermoFisher Scientific). cDNA samples prepared from stable pools were each run in duplicate with TaqMan probes against both the target sequence and the Mmadhc housekeeping gene. Negative control RT reactions were also run with TaqMan probes against both the target sequence and the Mmadhc housekeeping gene to ensure that Ct numbers were substantially higher than when cDNA samples were run with these probes, thereby inferring that RNA preparations did not contain target DNA at significant levels.

To perform absolute quantification of LC and HC mRNA from LC/HC synthetic promoter-containing stable pools, standard curves were set up with 10-fold serial dilutions of linearised and purified plasmid DNA containing the target LC or HC sequence. These standard curves were run in duplicate with LC or HC TaqMan probes. Dilution curves were created by 2-fold serial dilution of a single cDNA sample and

these were run in duplicate with Mmadhc Taqman probes. These dilution curves were used to inform on how differences in starting cDNA amount impacted Mmadhc Ct number. LC and HC standard curves and Mmadhc dilution curves are shown in the appendix. Ct numbers given by stable pool cDNA samples with LC and HC probes were converted to target copy numbers using the LC and HC standard curves. Ct numbers given by stable pool cDNA samples with the Mmadhc probe were used to normalise these LC and HC copy numbers for variation in starting cDNA amount/RT reaction efficiency based on the Mmadhc dilution curves.

To perform relative fold change quantification of endogenous GS mRNA from GS synthetic promoter-containing stable pools, the ΔΔCT method was used based on sample Ct numbers acquired with endogenous GS and Mmadhc Taqman probes. In order to check the specificity of the endogenous GS Taqman probe for its target sequence, this probe was also run with a serial dilution of cDNA prepared from untransfected CHO host RNA and a serial dilution of linearised plasmid containing the expression vector GS gene.

## 2.7 Flow cytometry

### 2.7.1 Measurement of transient GFP expression

CHO host cells transfected with the pMAX-GFP plasmid (Lonza) and cultured for 24 hours in Nunc 96 well plates (ThermoFisher Scientific) within static conditions (37˚C, 5% (V/V) $CO_2$) were analysed for GFP expression on an Attune Autosampler flow cytometer (ThermoFisher Scientific). TO-PRO-3 cell viability dye (ThermoFisher Scientific) was also added to transfected cells before flow cytometry analysis to help identify the proportion of each transfected cell population that was viable. GFP transfected samples and a mock transfected negative control were analysed as described in figure 4.1.

### 2.7.2 Measurement of LC and HC expression from stable pools

Cell pellets were taken from high viability stable pool cultures and washed with D-PBS before being fixed by resuspension in pre-chilled 70% methanol + 30% D-PBS solution. Fixed cells were then washed in 1% BSA + D-PBS solution before being stained by resuspension in solution containing an APC-conjugated goat anti-human IgG Fcγ (Jackson Immunoresearch) and a FITC-conjugated goat anti-human λLC (Southern Biotechnology) and incubation in the dark for ~1 hour. Stained cells were then washed

again in 1% BSA + D-PBS and resuspended in 1% BSA + D-PBS for flow cytometry analysis. Flow cytometry was performed on an LSRFortessa Cell Analyzer (BD Biosciences) with controls set up and samples analysed as described figure 4.10.

## 2.8 CHO cell culture maintenance

CHO cells were maintained in vented Erlenmeyer flasks (Corning) within shaking incubators (Infors HT) set to 37˚C, 140rpm and 5% (V/V) $CO_2$. CD-CHO growth medium (ThermoFisher Scientific) was used and this was supplemented with 6mM L-Glutamine for the routine passaging of CHO host cells pre-transfection. Passages were performed every 3-4 days and new subcultures were inoculated at a VCD of 2-3 x $10^5$/ml. The Vi-cell XR Cell Viability Analyzer (Beckman Coulter) was used to measure total and viable cell density as well as % viability of cultures. Host cells were passaged at least 3 times following revival from liquid nitrogen storage before being transfected in order to allow the cell population to adapt to the shaking culture conditions. A maximum of 20 passages was performed before transfection in order to minimise the effect of genetic drift on cellular phenotypes.

# Chapter 3

# CHO cell synthetic promoter design

*This chapter explains the design strategy that was used to create CHO cell synthetic promoters and the rationale behind this strategy.*

## 3.1 Introduction

### 3.1.1 Limitations of currently used promoters

The current convention when manufacturing mAb molecules in CHO cells is to use an expression vector containing both LC and HC genes under the control of hCMV promoters. The use of two hCMV promoters is a sub-optimal solution to achieving high stable production titres for several reasons. Firstly, hCMV promoters can lead to production instability of IgG-producing CHO cells, i.e. a reduction in cell-specific productivity (qP) over time. This production instability can come in two ways: hCMV has been shown to be prone to transcriptional silencing via DNA methylation and, in addition, two copies of an identical hCMV sequence within a single expression construct can allow homologous recombination events between these two hCMV sequences that leads to recombinant gene deletions (Kim et al., 2011).

Secondly, hCMV is an unnecessarily large vector component and therefore contributes to the creation of large expression vectors, thus reducing transfection efficiency compared to if the expression vector was smaller. Synthetic promoters specifically designed for CHO cell activity have been shown to enable similar levels of transcriptional output to hCMV with much shorter sequences (Brown et al., 2014). Finally, the use of two hCMV promoters may not enable LC and HC to be transcribed at ratios other than 1:1. The expression of LC in excess of HC has been shown to be important for maximising production titres of some mAbs (Pybus et al., 2014). This idea is explored in detail in chapter 4 of this manuscript.

Conventionally, the SV40 promoter is used to control GS selection marker gene transcription in mAb expression vectors used for stable CHO pool creation (Fan et al., 2013). While this promoter displays transcriptional activity in mammalian cells and is used successfully as part of the GS selection system, it may not produce the optimal level of GS gene transcription for optimised mAb production titres. By using a promoter with less transcriptional activity than the SV40 promoter to control GS transcription, the level of selective pressure on stably transfected CHO pools will be greater and this

could lead to greater recombinant mAb productivity. This idea is explored in detail in chapter 5 of this manuscript.

### 3.1.2 TFRE 'building block' design strategy

As described in more detail in subsections 1.6.2 – 1.6.4, mammalian synthetic proximal promoters have been designed previously using a variety of strategies. Broadly, these strategies include the creation of purely synthetic transcription factor – promoter pairs, mutation of natural promoter sequences and the joining of certain DNA fragments (Chavez et al., 2015; Ferreira et al., 2011; Grabherr et al., 2011; Ogawa et al., 2007; Perez-Pinera et al., 2013; Rössger et al., 2014; Schlabach et al., 2010; Tornøe et al., 2002). These methods have all successfully enabled the creation of synthetic promoter sequences showing transcriptional activity in mammalian cells. However, none of these approaches enabled the precise design of synthetic promoters displaying pre-determined transcriptional outputs in CHO cells.

As an alternative to these other synthetic promoter design strategies, synthetic promoter libraries displaying broad ranges of transcriptional output in CHO cells have been created previously using a transcription factor regulatory element (TFRE) 'building block' design approach (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019). Unlike other synthetic promoter construction methods, this method has been shown to reliably enable assembly of promoters displaying pre-determined transcription output levels in CHO cells. A TFRE is a short nucleotide sequence that binds to a discrete transcription factor within the nucleus. This design strategy is predicated on the billboard model of promoter activity, which explains that overall transcriptional output from a promoter is simply the sum of transcriptional outputs given by each individual TFRE block, regardless of the order in which these TFRE blocks are arranged (Arnosti & Kulkarni, 2005). This is akin to the way that information on a billboard can be arranged in any order without altering the total sum of information given.

There was a common workflow for the TFRE 'building block' strategy used previously in these studies (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019) (figure 3.1). Firstly, the relative level of transcriptional output in CHO cells given by many different individual TFRE blocks was determined. This was done by measurement of SEAP (secreted alkaline phosphatase) reporter gene expression following transient transfection of so called "homotypic" promoter constructs containing several copies of the individual TFRE block of interest upstream of a CMV

core promoter sequence. This determination of individual TFRE block transcriptional output then enabled the combination of different TFRE blocks into "heterotypic" promoters with different levels of overall transcriptional output (figure 3.1).
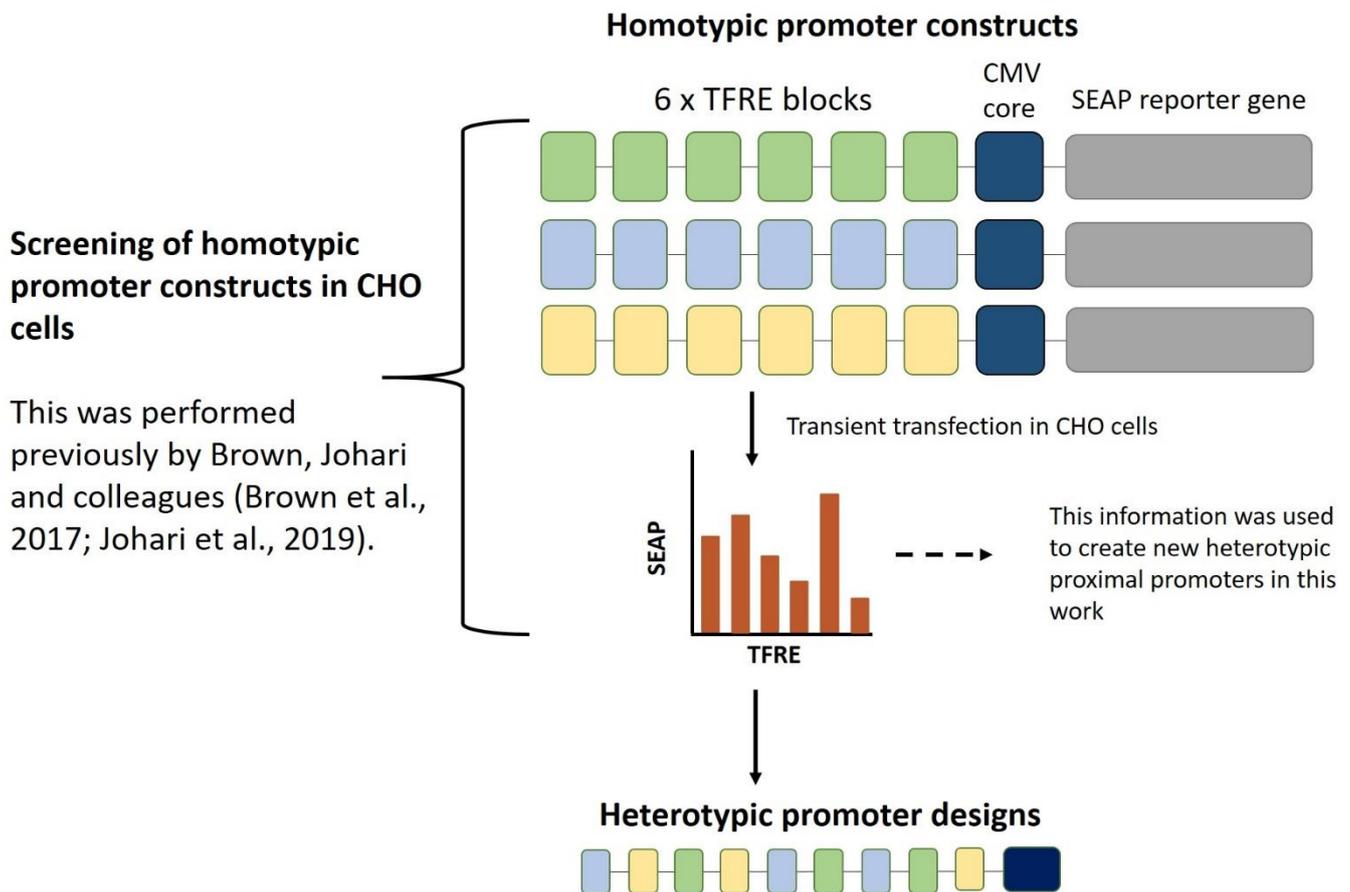


**Figure 3.1 Construction of heterotypic promoters through the TFRE 'building block' design strategy**

The TFRE building block design strategy was developed previously by Brown, Johari and colleagues. In these previous studies, they transiently transfected homotypic constructs containing six repeat copies of individual TFRE blocks into CHO cells. These TFRE repeats were upstream of a CMV core sequence and this whole promoter was driving the transcription of a SEAP (secreted alkaline phosphatase) reporter gene. Measurement of SEAP expression revealed the relative transcriptional activity levels of individual TFRE blocks in CHO cells. This information was used in the work presented in this manuscript to create new heterotypic synthetic promoters.

## 3.2 Results

### 3.2.1 Creation of new synthetic promoters to control LC and HC gene transcription in stably expressing CHO cells

In this work, synthetic promoters were designed to control LC and HC gene transcription. To construct synthetic promoter proximal regions that would then be positioned upstream of CMV core sequences, the TFRE 'building block' design strategy was utilised. The relative level of transcriptional output in CHO cells given by many different individual TFRE sequence blocks had already been determined previously by Brown, Johari and colleagues through screening of homotypic constructs in CHO cells (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019) (figure 3.1). This information was utilised to create heterotypic promoter sequences with different levels of predicted overall transcriptional output by combining different TFRE blocks together.

It would be advantageous if pairs of CHO cell synthetic promoters used alongside each other on a single expression vector, for example for the control of mAb LC and HC gene transcription, were designed to have no overlap in the TFRE blocks that they contained. This would ensure that the two promoters are not competing for binding of transcription factors within the nucleus and instead function as independent units. Furthermore, synthetic promoter pairs containing completely separate groups of TFREs will have very little sequence similarity, thereby ensuring there is minimal risk of homologous recombination events within the expression construct. Synthetic promoters designed previously through the TFRE 'building block' strategy, especially those that are relatively active, have substantial overlap in TFRE usage (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019).

Two important design criteria were therefore applied for the assembly of these proximal promoter sequences. Firstly, promoters had to vary in their predicted transcriptional activity levels to enable the expression of different LC:HC ratios. Secondly, LC and HC promoters were designed to have no overlap in the TFREs that they contain.

The proximal promoter design space comprised a list of TFRE blocks shown to be active in CHO cells across two previous analyses (Brown et al., 2017; Johari et al., 2019). These TFREs and their sequences are listed in table A1 of the appendix. The list of TFRE blocks was partitioned between LC and HC promoter design spaces, ensuring

there would be no overlap in TFRE usage between the two promoters that would be used in combination. TFREs were then assigned a transcriptional activity score by ranking their measured activity relative to the activity of NFKB, the most active TFRE block in both of these previous analyses. For example, if the homotypic construct for a particular TFRE generated 50% of the reporter gene expression shown by the NFKB homotypic construct then this particular TFRE would be assigned a transcriptional activity score of 50. This assignment of transcriptional activity scores to TFRE blocks then enabled them to very simply be assembled into heterotypic promoter designs of different predicted overall transcriptional activity. For example combination of two copies of a TFRE with a transcriptional activity score of 30, one copy of a TFRE with a transcriptional activity score of 85 and three copies of a TFRE with a transcriptional activity score of 45 would result in a heterotypic promoter with an overall predicted relative transcriptional activity of (30 x 2) + (1 x 85) + (3 x 45) = 190. TFRE compositions of synthetic proximal promoters used to control LC and HC transcription in this work are shown in appendix table A4 and their sequences are shown in table A2. A single, highly active promoter was designed for control of LC transcription while five different promoters with decreasing levels of predicted transcriptional activity were designed for control of HC transcription, creating predicted LC:HC transcriptional ratios of 1, 1.4, 2.3, 3.2 and 11.3.

TFRE blocks within these synthetic promoter proximal regions are separated by random 6nt spacer sequences. In order to ensure that additional TFREs were not inadvertently created at TFRE-spacer junctions, promoters were analysed using the online transcription factor binding site search tool MatInspector (part of the Genomatix software suite) with stringent search parameters to help minimise false positives (core similarity and matrix similarity values both set to 1.00).

The largest of the LC/HC synthetic promoter proximal regions created in this work, HC1, is 447nt in total length. For comparison, the hCMV proximal region used in this work is 598nt in length. DNA methylation at CpG dinucleotides has been shown to cause transcriptional silencing of the hCMV promoter in CHO cells (Kim et al., 2011). The hCMV proximal region used in this work contains 35 CpG dinucleotides while the highest amount in any of the LC/HC synthetic proximal regions is 33 in HC2 and HC3. The sequences of these LC/HC synthetic promoter proximal regions are listed in appendix table A2.

### 3.2.2 Creation of new synthetic promoters to control GS gene transcription in stably expressing CHO cells

Synthetic proximal promoters that would be positioned upstream of CMV core sequences and used to control GS transcription were designed. As with the design of the LC/HC synthetic proximal promoters, the TFRE 'building block' design strategy was used (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019). The GS synthetic proximal promoter design space comprised of TFRE blocks that a) showed low level activity in CHO cells in a previous analysis and b) were not contained in the LC/HC synthetic promoter designs (Johari et al., 2019). Avoiding TFRE blocks used in LC/HC synthetic promoter designs would ensure no competition for transcription factor binding between the three promoters when used together on the same expression vector. These TFREs were assigned a transcriptional activity score by ranking their measured activity relative to that of NFKB, the most active TFRE in this analysis (Johari et al., 2019). This assignment of transcriptional activity scores to TFRE blocks within the design space enabled promoters of different overall predicted activity to be constructed.

SV40 promoter transcriptional output from within a conventional vector containing hCMV driving LC and HC transcription would be influenced by the strength of the SV40 promoter as well as the attenuating effect of hCMV-SV40 TF binding competition. Synthetic GS promoters would need to give less transcriptional output than this in order to increase selection stringency. Previous studies performed in the James laboratory provide information that can be used to help estimate the level of SV40 transcriptional output in this situation. First, the SV40 promoter was found to be 4.3-fold weaker than hCMV (Johari et al., 2019). Second, comparison of co-transfecting an hCMV-driven marker alongside an SV40-driven marker against transfecting just the SV40-driven marker alone showed that hCMV reduced SV40 mRNA output by 2.69 to 5.15 fold (West, 2014). Taken together, this would mean that when alongside hCMV in the nucleus, SV40 output is ~11.6 - 22.15 fold less than hCMV. However, these hCMV/SV40 cotransfection experiments were performed with the two promoters at a 1:1 molar ratio. When a mAb expression vector containing two hCMV promoters per SV40 promoter is used, the reduction in SV40 promoter output will likely be even greater. Therefore, based on these previous findings, GS synthetic promoters were constructed by combination of different TFRE blocks to give overall predicted transcriptional activities of 1/15th, 1/20th, 1/30th, 1/40th, 1/50th and 1/60th the

transcriptional activity of hCMV.

TFRE compositions of the six synthetic proximal promoters designed to control GS gene transcription are shown in appendix table A5. The sequences of these proximal promoters are shown in appendix table A3. The sequences of all TFRE blocks used throughout this work are listed in appendix table A1. These GS proximal promoters contain the required TFRE blocks separated by random 6nt spacer sequences. In order to ensure that additional TFREs were not inadvertently created at TFRE-spacer junctions, promoters were analysed using the online TF binding site search tool Matinspector (part of the Genomatix software suite) with stringent search parameters to help minimise false positives (core similarity and matrix similarity values both set to 1.00).

### 3.2.3 Creation of 'synthetic SV40' promoters to control GS gene transcription in stably expressing CHO cells

The GS synthetic promoters described above were shown to create increased selection stringency on transfected cell pools (chapter 5) but this did not lead to improved mAb productivity, likely due to these promoters being too weak. The TFRE 'building block' design strategy was therefore ineffective in this instance. An alternative design strategy was then used to try and create synthetic GS promoters that gave a more effective level of selection stringency, leading to increased mAb productivity. This alternative strategy involved first identifying the key TFREs within the SV40 promoter that are likely to control the bulk of its transcriptional activity in CHO cells. Next these TFRE blocks would be used in synthetic promoter designs to create synthetic, attenuated and minimal versions of the SV40 promoter.

As explained in detail in the previous subsection, the presence of an hCMV promoter alongside an SV40 promoter in the nucleus of a CHO cell substantially reduces SV40 promoter output due to t binding competition (West, 2014). It was therefore reasoned that the key TFREs driving the bulk of SV40 promoter output in CHO are likely to be found within both SV40 and hCMV promoters. The transcription factor binding site search tool Matinspector was used to find predicted binding sites common to both promoters, generating a list of 33 different TFREs. This list of TFREs was then reduced based on whether each one had displayed activity in previous comprehensive analyses of homotypic constructs (Brown et al., 2017; Brown et al., 2014; Johari et al., 2019). From this analysis it was concluded that transcriptional

activity of the SV40 promoter in CHO cells is driven predominantly by TFREs corresponding to the transcription factors AP1, SP1, SP2 and HIVEP1. Using these TFRE blocks, two synthetic proximal promoters named SynSV40_1 and SynSV40_2 were constructed for the control of GS transcription.

The TFRE compositions of SynSV40_1 and SynSV40_2 proximal promoters are shown in appendix table A6 and the full sequences are shown in appendix table A3. Sequences of these TFRE blocks are included in appendix table A1. SynSV40_1 contains these four different TFRE blocks in the same quantities that are found in the SV40 promoter (11 blocks total). The 11 TFRE blocks in SynSV40_1 was reduced to 6 in SynSV40_2 in an effort to create a less transcriptionally active promoter. TFRE blocks within these synthetic proximal promoters are separated by random 6nt spacer sequences. In order to ensure that additional TFREs were not inadvertently created at TFRE-spacer junctions, promoters were analysed using Matinspector with stringent search parameters to help minimise false positives (core similarity and matrix similarity values both set to 1.00).

# Chapter 4

# Use of synthetic promoters to control LC and HC transcription

*This chapter describes, for the first time, the use of synthetic promoters to control mAb LC and HC transcription within stably expressing CHO cell pools. LC and HC synthetic promoters were constructed from transcription factor regulatory element (TFRE) sequence blocks to enable varying LC:HC transcriptional ratios. Stable pool expression of two different mAb molecules revealed synthetic promoter-containing stable pools to have improved characteristics (reduced recovery time from MSX selection and increased mean mAb titres) over conventional hCMV-containing stable pools. qRT-PCR and flow cytometry assays were carried out to further characterise these stable pools in terms of their LC:HC transcriptional ratio and cell-to-cell LC/HC expression variation respectively.*

## 4.1 Introduction

The expression of LC in excess of HC has been shown to be important for maximising production titres of some mAbs (Pybus et al., 2014). The molecular mechanisms behind why expressing LC in excess of HC can be beneficial in a mAb-specific manner are well understood. Following transcription and translation of the HC gene, the HC polypeptide is bound by the molecular chaperone known as BiP (binding immunoglobulin protein) in the lumen of the ER. The binding of BiP to HC represents a quality control checkpoint in mAb folding and assembly and HC will not be released from BiP until a LC polypeptide binds to this HC-BiP complex (Vanhove et al., 2001). HC polypeptides that remain bound to BiP for a prolonged period are eventually recognised by ER machinery as being unassembled/misfolded and build-up of misfolded polypeptides can eventually lead to activation of the ER stress response known as the Unfolded Protein Response (UPR) (Walter & Ron, 2011). One of the consequences of UPR activation is a global inhibition of mRNA translation, which prevents the flow of nascent polypeptides into the ER. When expressing mAbs from CHO cells, this translational inhibition will cause substantially reduced cell growth and mAb productivity (Pybus et al., 2014).

It appears that different HC polypeptides have a different temporal threshold for

when they must be bound by a LC. For example, some HC polypeptides may require LC to bind and relieve BiP relatively rapidly or else that HC will be recognised as misfolded, increasing the risk of UPR activation. For a mAb containing a HC with this profile, it may be important to express LC in excess of HC to ensure frequent LC-HC-BiP binding events, thereby reducing the risk of UPR activation. Conversely, other HC polypeptides may be capable of remaining bound to BiP for relatively prolonged periods before being bound by LC and they will not be recognised as misfolded, meaning the risk of UPR activation is relatively low. For a mAb containing a HC with this profile, it is likely to be most beneficial to express LC and HC at a balanced ratio.

An attempt has been made previously to develop expression vector components that enable the expression of different LC:HC ratios. Li et al placed mutant internal ribosome entry sites (IRES) within bicistronic mRNA sequences containing both the LC and HC coding sequence in an attempt to vary translation efficiencies of HC (Li et al., 2007). However, these mutant IRES elements did not display quantifiably variable translational efficiencies in CHO-K1 despite doing so in HEK293. Moreover, IRES-mediated translation is generally much less efficient in CHO compared to in HEK293 and plasmids with separate promoters controlling LC and HC expression units are favoured over bicistronic mRNAs for CHO-based mAb manufacture.

LC:HC expression ratios could also be manipulated at the transcriptional level through the use of synthetic promoters designed to give different levels of transcriptional output. In this chapter, pairs of synthetic promoters, the design of which is described in chapter 3, were used for the first time to control mAb LC and HC transcription in stably expressing CHO cell pools. Transient transfections in which the transfected LC:HC plasmid ratio was varied were first carried out to understand how different LC:HC expression ratios affected the assembly and expression of two different mAb molecules. Expression vectors were then created containing the synthetic promoter pairs that were designed to give varying LC:HC transcriptional ratios and these expression vectors were used to generate stable pools expressing two different mAb molecules. Synthetic promoter-containing pools showed reduced recovery times from MSX selection and in some cases higher mean mAb titres than conventional hCMV promoter-containing pools. qRT-PCR assays were carried out to examine the LC:HC transcriptional ratios given by these stable pools and flow cytometry was performed to examine their cell-to-cell variability in LC and HC expression.

**4.2 Results**

**4.2.1 Transient expression of LC and HC at various ratios**

Transient expression of genes of interest can rapidly provide information about how proteins deriving from those genes of interest are affecting the host cell system. In order to understand more about how different LC:HC expression ratios would affect monoclonal antibody production from CHO cells, it was therefore useful to transiently co-express plasmids that separately encoded the LC and HC of monoclonal antibodies at various different ratios.

**4.2.1.1 Test of Nucleofection transfection efficiency consistency**

Before transiently co-expressing plasmids encoding the LC and HC of monoclonal antibodies, a suitable transient transfection method needed to be identified. This transient transfection method would need to provide: rapid expression of genes of interest upon transfection; a high transfection efficiency; a high degree of consistency in transfection efficiency from transfection-to-transfection and minimal negative impact on cell viability. The method being high-throughput was also a desirable characteristic, as this would easily allow the testing of multiple different LC:HC expression ratios within a single experiment. The Nucleofection (Lonza) method was selected, in particular the 96 well Nucleocuvette plate system, based on purportedly high transfection efficiency in many cell lines and rapid expression of transgenes following transfection as well as it's high-throughput nature.

In order to examine the transfection efficiency across a full 96 well Nucleocuvette plate, CHO host cells were transfected with the Lonza pMAX-GFP (green fluorescent protein) plasmid. $2 \times 10^6$ viable cells were transfected with 400ng of GFP plasmid in each well of the plate. 24 hours after transfection the proportion of each transfected cell population that was positive for GFP expression was examined by flow cytometry. Use of TO-PRO-3 cell viability dye (ThermoFisher) allowed the impact of nucleofection on cell viability to be examined.

The mean % GFP positive cells from each well of the plate (n=3) is displayed in figure 4.1. The mean % GFP positive cells across the whole plate was 77.36% with a standard deviation of 2.23%. Moreover, the GFP median fluorescence intensity (MFI) was also calculated for each well (data not shown). The across-plate mean of each well's mean MFI (n=3) was $7.71 \times 10^4$ with a standard deviation $8.47 \times 10^3$. Taken together, these data show that the delivered transfection efficiency is highly consistent
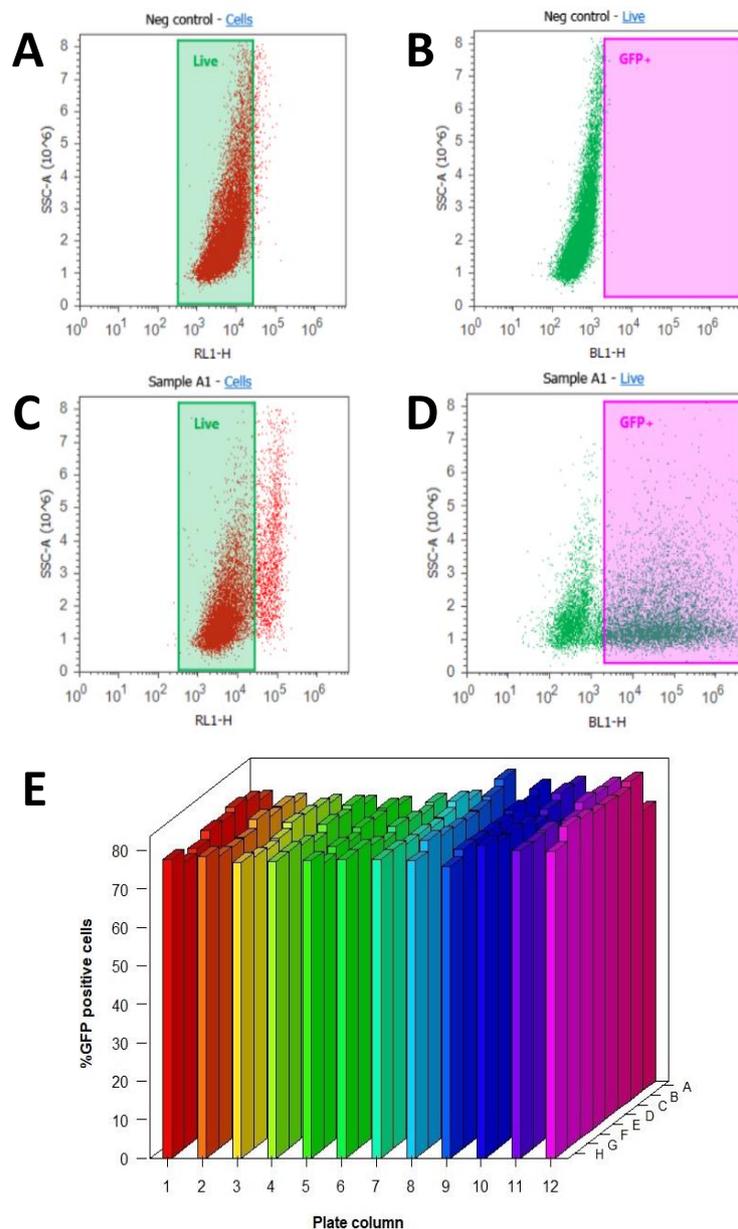
**Figure 4.1 – The 96 well Nucleocuvette plate gives consistently high transfection efficiency across the whole plate**

96 well Nucleocuvette plates were used to transfect 400ng of pMAX-GFP in to $2 \times 10^6$ viable CHO host cells per well. Mock transfections were also performed with 400ng of a non protein-coding plasmid. GFP expression and cell viability were measured 24h post-transfection by flow cytometry. **A)** and **C)** show gating of viable/non-viable mock transfected (**A**) and GFP transfected (**C**) cells based on TO-PRO-3 cell viability dye fluorescence. **B)** and **D)** show gating of GFP positive/GFP negative mock transfected (**B**) and GFP transfected (**D**) cells. **E)** displays the mean % GFP positive of cells transfected using each well of the 96 well Nucleocuvette plate. Bars represent the mean of three independent experiments. The gradient of colours for wells in different columns of the 96 well plate is purely to aid visualisation.

across the 96 well Nucleocuvette plate in terms of both the proportion of cells expressing the transgene and the amount of transgene taken up by each cell. Finally, the across-plate mean of each well's mean % viability (n=3) was 83.87% with a standard deviation of 3.45%, showing that culture viability was not excessively reduced by the transfection event and that viability was highly consistent across the 96 well plate. This experiment verified that the 96 well Nucleocuvette plate transfection method was capable of meeting all important criteria. The Nucleofection delivered high transfection efficiency within 24 hours and minimal cell death. Transfected cells showed high consistency in transfection efficiency and viability, meaning the effect of different LC:HC expression ratios on mAb production following transfection could be interpreted with confidence without confounding factors. This method was therefore used for transient transfection of mAb LC and HC plasmids at different ratios.

### 4.2.1.2 Test of 96 deep well plate cell growth consistency

The experiment described above confirmed that the 96 well Nucleocuvette plate transfection method gave high well-to-well consistency in transfection efficiency. The next objective was to ensure that transfected cells could be cultured for several days in a high-throughput plate format with consistent well-to-well growth rates. Any significant plate location-specific biases in cell growth, for example cells dividing more freely in the middle of the plate compared to the edges or vice versa, would obscure the analysis of how different transfected LC:HC plasmid ratios were affecting mAb production.

Nunc 96 deep well plates (ThermoFisher) were selected for culture of cells following Nucleofection due to them being a high-throughput plate format, facilitating the testing of multiple LC:HC expression ratios in a single experiment. Both the transfection plate and the culture plate being a 96 well format would enable easy transfer of cells following transfection. In order to assess the extent of well-to-well variation in post-transfection cell growth from these 96 deep well plates, $2 \times 10^6$ viable CHO host cells were transfected with 400ng of an equimolar mixture of mAb LC and HC plasmids per Nucleocuvette plate well across 21 wells. These transfected cells were then seeded into a 96 deep well plate to a VCD of around $0.65 \times 10^6$/ml and cultured for 5 days. Seeded wells covered both the edges and the centre of plate rows. The relative VCD of each well was measured immediately after transfection and then 5 days after transfection using PrestoBlue cell viability dye (ThermoFisher). Figure 4.2 displays the mean and standard deviation of day 5 - day 0 PrestoBlue signal across these 21 wells.

It is clear from these data that no substantial well-to-well variation in growth of transfected cells is seen when cultured for 5 days in Nunc 96 deep well plates. These plates were therefore used for culture of cells transiently transfected with mAb LC and HC plasmids at different ratios.
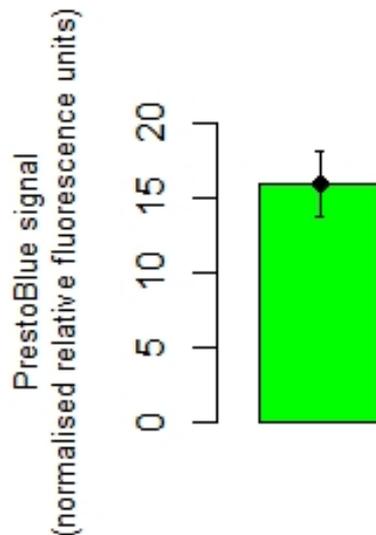


**Figure 4.2 – Transfected cells cultured in 96 deep well plates grow consistently from well-to-well**

Cells transfected using the 96 well Nucleocuvette plate method with 400ng of an equimolar mixture of mAb LC and HC plasmids were cultured for 5 days in Nunc 96 deep well plate wells. PrestoBlue cell viability dye was used to measure the relative viable cell density in each well immediately after transfection and then again 5 days after transfection. The bar represents the mean day 5 - day 0 PrestoBlue signal of 21 wells and error bars represent standard deviation of the mean.

### 4.2.1.3 Transient co-transfection of LC and HC plasmids at various ratios

In order to gain some understanding about how different LC:HC expression ratios would affect monoclonal antibody assembly and production titre, single gene plasmids separately encoding the LC and HC of two different mAbs were transiently co-transfected into CHO host cells at varying ratios. Transient expression would rapidly provide information that could then be used to inform more long-term stable expression-based experiments. Vector components controlling the expression output

from these LC and HC genes were the same in both plasmids, including hCMV promoters to control transcription.

LC and HC plasmids for two different mAbs were transiently co-transfected at LC:HC molar ratios of 1.0, 1.4, 2.1, 3.2, 5.4 and 12.0 while maintaining a constant overall DNA load. MAb 1 is an anti-interleukin-13 IgG1λ with a LC that had previously undergone a single amino acid mutation in the $V_L$ region to help engineer improved expression titre (Popovic et al., 2017). MAb 2 is also an IgG1λ molecule. These single gene plasmids were transiently co-transfected using the 96 well Nucleocuvette plate method described above and transfected cells were cultured for 5 days in the Nunc 96 deep well plates described above. After 5 days, cell culture supernatant was harvested and mAb titre was quantified using an IgG ELISA. The results of these experiments are displayed in figures 4.3 and 4.4 for mAb 1 and mAb 2 respectively.

Figure 4.3 shows that mean mAb 1 titre increases only slightly compared to the equimolar transfection control when LC is expressed in excess of HC up to a LC:HC ratio of 5.4. Mean mAb 1 titre is then reduced compared to the equimolar control when LC is expressed at a 12-fold excess of HC. Figure 4.4 shows that mean mAb 2 titre decreases steadily compared to the equimolar control as LC expression excess increases.

It was noted that the extent to which relative mAb titre decreased at high LC excess expression ratios was not proportional to the decrease in amount of HC gene being transfected. For example, the mean mAb 2 titre at the 12:1 LC:HC expression ratio is 45% that of the equimolar control. This is despite the amount of HC gene transfected being 17% that of the equimolar control. Therefore HC gene utilisation was calculated for all transfection ratios and for both mAbs to show how efficiently HC gene was converted to secreted antibody (Equation 4.1). Both figure 4.3 and figure 4.4 show that for mAb 1 and mAb 2 respectively HC gene is utilised more efficiently at higher LC:HC expression ratios.

$$Equation\ 4.1\ -$$

$$HC\ gene\ utilisation = \frac{Antibody\ titre\ (ng/ml)}{Transfected\ HC\ plasmid\ amount\ (fmol)}$$

## mAb 1

**A**

Relative mAb titre (fold change)

Molar LC:HC plasmid ratio

**B**

Heavy chain gene utilisation (fold change)

Molar LC:HC plasmid ratio
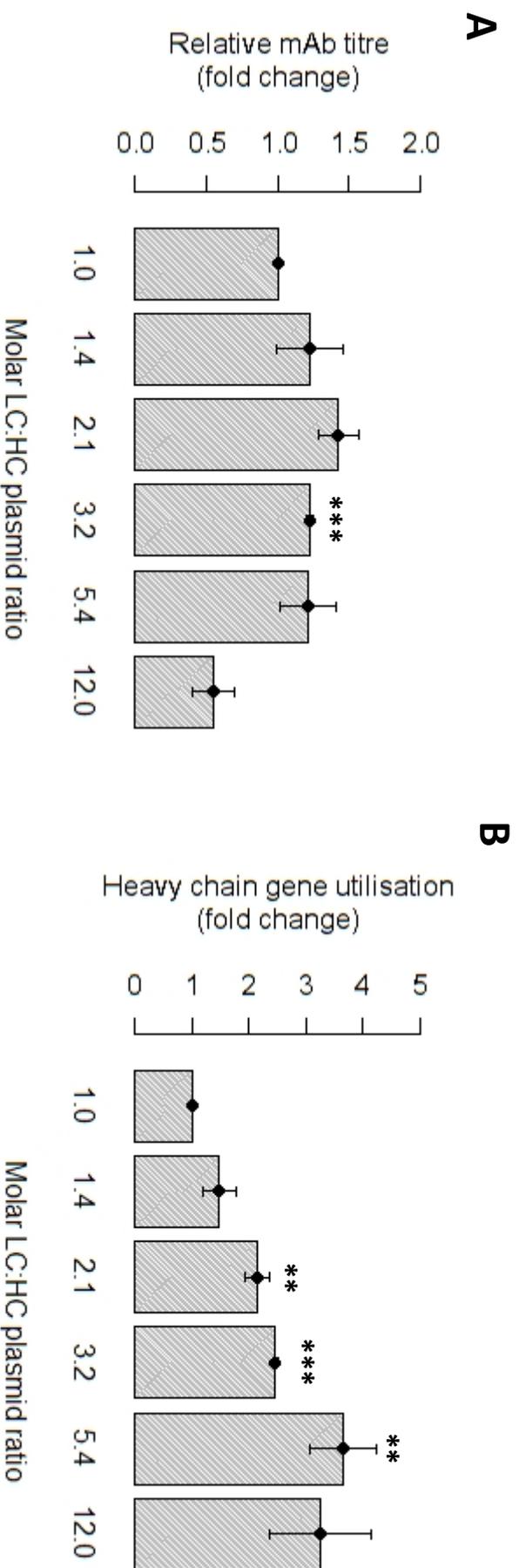
**Figure 4.3 – The effect of transiently transfecting various ratios of mAb 1 LC and HC plasmids**
CHO host cells were transiently transfected using 96 well Nucleocuvette plates with varying ratios of two plasmids that separately encoded the LC and HC polypeptides of an IgG1 monoclonal antibody. Antibody expression titre was measured 5 days post transfection using an IgG ELISA. **A)** shows relative antibody titre from cells transfected with different plasmid ratios. **B)** shows relative HC gene utilisation by cells transfected with different plasmid ratios. The equation used to calculate HC gene utilisation is displayed in the text (Equation 4.1). In both **(A)** and **(B)** data are expressed as fold change relative to the 1.0 LC:HC molar plasmid ratio. Bars represent the mean of three separate transfections and error bars represent standard error of the mean. Mean values significantly different (two-tailed Students t-test) from the 1.0 LC:HC molar plasmid ratio are indicated by asterisks ($*$ P$<$0.05, $**$P$<$0.025, $***$P$<$0.01).

**mAb 2**

**Figure 4.4 – The effect of transiently transfecting various ratios of mAb 2 LC and HC plasmids**
CHO host cells were transiently transfected using 96 well Nucleocuvette plates with varying ratios of two plasmids that separately encoded the LC and HC polypeptides of an IgG1 monoclonal antibody. Antibody expression titre was measured 5 days post transfection using an IgG ELISA. **A)** shows relative antibody titre from cells transfected with different plasmid ratios. **B)** shows relative HC gene utilisation by cells transfected with different plasmid ratios. The equation used to calculate HC gene utilisation is displayed in the text (Equation 4.1). In both **(A)** and **(B)** data are expressed as fold change relative to the 1.0 LC:HC molar plasmid ratio. Bars represent the mean of three separate transfections and error bars represent standard error of the mean. Mean values significantly different (two-tailed Students t-test) from the 1.0 LC:HC molar plasmid ratio are indicated by asterisks (* P<0.05, **P<0.025, ***P<0.01).

60

### 4.2.2 Effect of LC/HC synthetic promoters on mAb 1 stable pool generation

When making a stably expressing cell pool, manipulating the LC:HC expression ratio through use of synthetic promoters on a dual gene plasmid is a more efficient solution than altering the transfected ratio of single gene plasmids. This is because expression from one construct containing both LC and HC genes in tandem can easily be selected for using a single selection marker. Therefore synthetic promoters were designed (chapter 3) to control the transcription of LC and HC genes in stably expressing CHO cells.
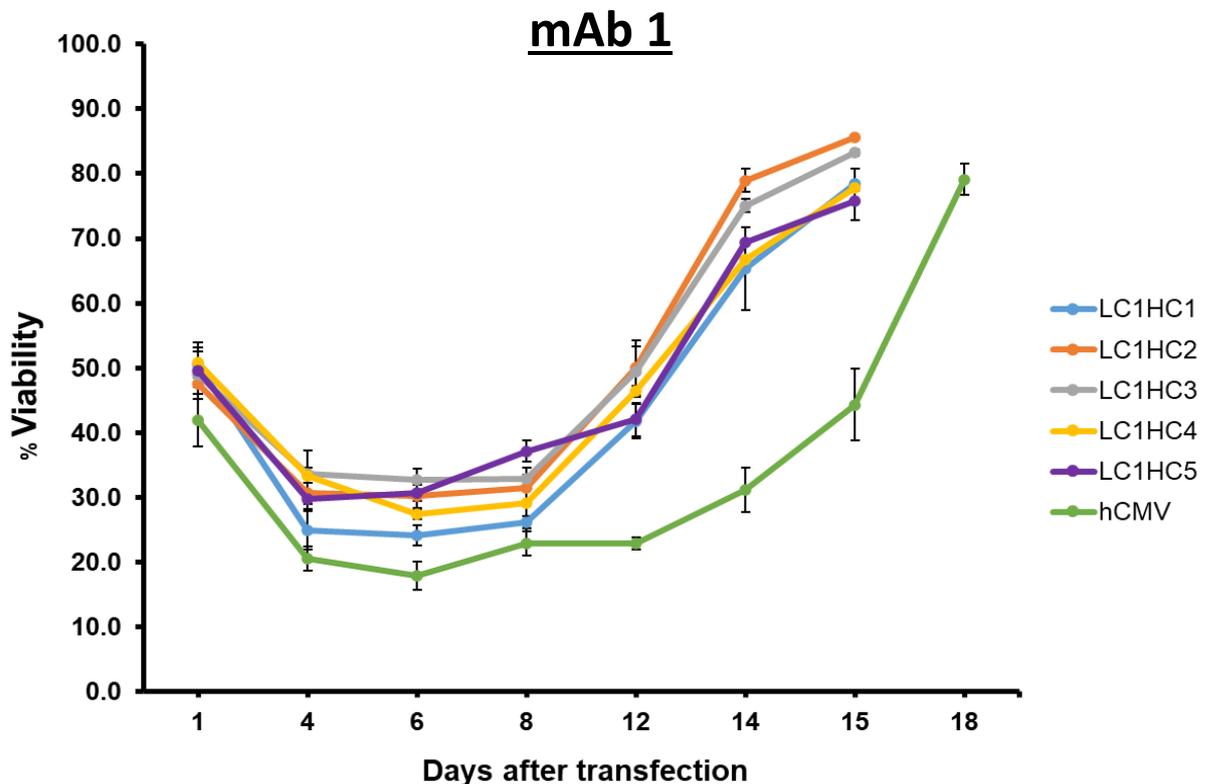


**Figure 4.5 – Effect of LC/HC synthetic promoters on recovery rate of mAb 1-expressing pools**

CHO host cells were transfected with stable expression vectors containing either different combinations of synthetic promoters to control LC and HC gene transcription or two hCMV promoters. Vector-derived expression was selected for by the addition of $50\mu$M MSX 24 hours post transfection. % viability was measured until full recovery of pools. Data points represent the mean of three separately transfected pools per vector at each time point. Error bars represent standard error of the mean.

Synthetic promoters designed for controlling LC and HC gene transcription were tested for their effect on stable mAb expression characteristics such as the rate of outgrowth of expressing cells and mAb productivity. Synthetic proximal promoters were cloned directly upstream of CMV core sequences in mAb 1 LC/HC expression units. These LC and HC expression units were then combined into a vector backbone containing a GS selection marker expression unit driven by an SV40 promoter. LC and HC synthetic promoters were used in the following combinations: LC1HC1, LC1HC2, LC1HC3, LC1HC4, LC1HC5. A control vector containing hCMV promoters to drive LC and HC transcription was also made. To generate stably expressing CHO pools using these expression vectors, CHO host cells were transfected and $50\mu$M MSX selection agent was added to transfected cultures 24 hours later. The recovery of pools from transfection and MSX selection was monitored (figure 4.5).

There was a clear and pronounced lag in the recovery of the hCMV pools compared to the recovery of pools generated using synthetic promoters. To reach a mean % viability > 70%, the hCMV pools required an extra 3 days compared with LC1HC1, LC1HC4 and LC1HC5 and an extra 4 days compared with LC1HC2 and LC1HC3. Differences in recovery rate among synthetic promoter pools were much smaller than the difference between synthetic promoter pools and the hCMV pools.

Following recovery of transfected cell pools from MSX selection they then went into 12 day fed batch overgrow cultures (FBOG) to assess their mAb productivity (figure 4.6). Pools generated with the expression vector LC1HC1 produced a slightly higher mean mAb 1 titre than pools generated with the hCMV expression vector, albeit this difference was not statistically significant (two-tailed Students t-test P value = 0.29). Pools generated using expression vectors LC1HC2, LC1HC3, LC1HC4 and LC1HC5 all show significantly reduced mean mAb 1 titre compared to the hCMV pools. The five synthetic promoter-containing pools show a negative correlation between the predicted LC:HC transcriptional ratio and mean mAb titre.

The hCMV pools had a significantly higher mean qP than all synthetic promoter-containing pools. As with titre, there was also a negative correlation between the predicted LC:HC transcriptional ratio and mean qP among the synthetic promoter-containing pools. In contrast to their increased qP, the hCMV pools showed the lowest mean IVCD of all six expression vectors.
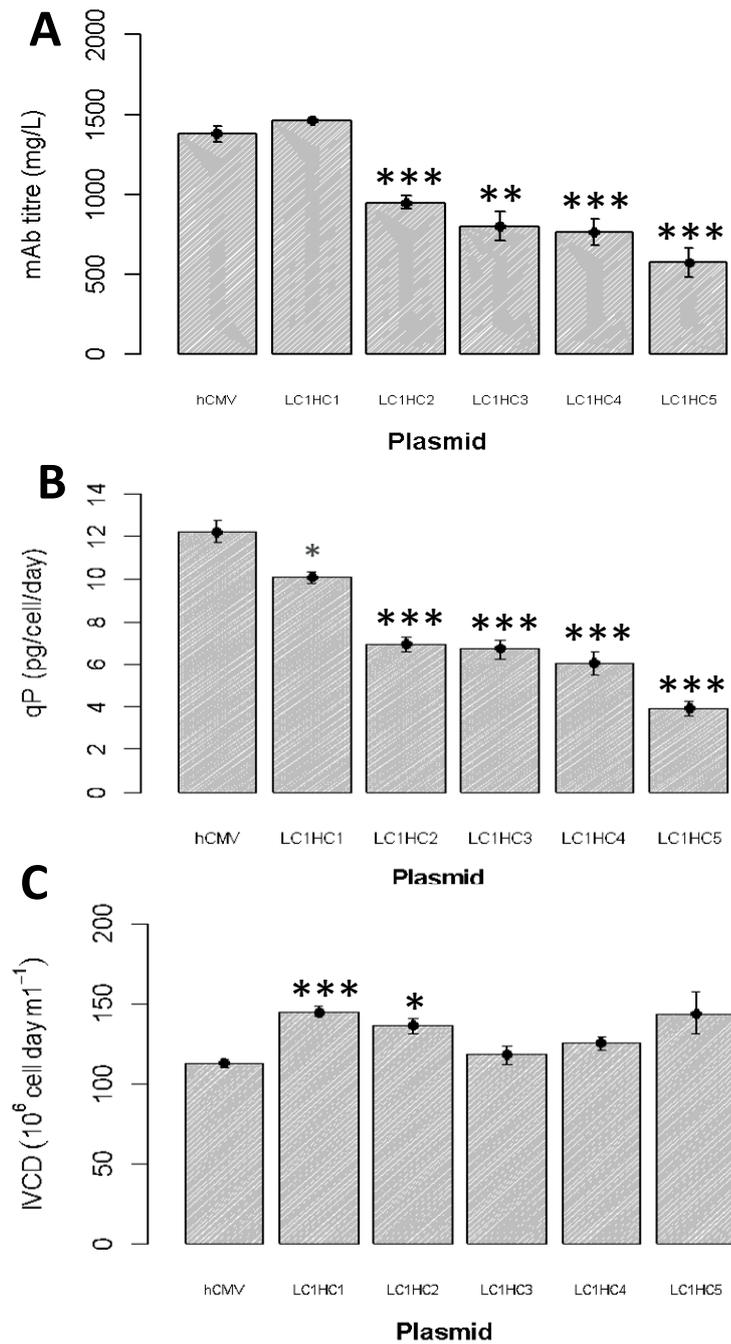
# mAb 1



**Figure 4.6 – Effect of LC/HC synthetic promoters on mAb productivity and IVCD of mAb 1-expressing pools**

CHO pools that had been transfected with synthetic promoter-containing mAb 1 expression vectors and selected using MSX went in to 12 day fed batch overgrow cultures. **A)** shows the volumetric mAb titre generated from pools transfected with each vector; **(B)** shows the cell specific productivity (qP) of pools transfected with each vector and **(C)** shows the integral of viable cell density (IVCD) across the 12 day culture period for pools transfected with each vector. Bars represent the mean of three separately transfected pools per vector and error bars represent standard error of the mean. Mean values significantly different (two-tailed Students t-test) from the mean hCMV value are indicated by asterisks (* P<0.05, **P<0.025, ***P<0.01).

### 4.2.3 Effect of LC/HC synthetic promoters on mAb 2 stable pool generation

LC/HC synthetic promoters were tested for their effect on stable expression of mAb 2. Expression vectors were constructed containing the same LC/HC promoter combinations used for stable expression of mAb 1 (LC1HC1, LC1HC2, LC1HC3, LC1HC4, LC1HC5 and hCMV) but ORF sequences were changed to those of mAb 2 LC and HC. These vectors contained a GS expression unit driven by an SV40 promoter. To generate stably expressing pools, CHO host cells were transfected with these vectors and 50$\mu$M MSX was added to transfected cultures 24 hours later. The recovery of pools from transfection and MSX selection was monitored (figure 4.7).
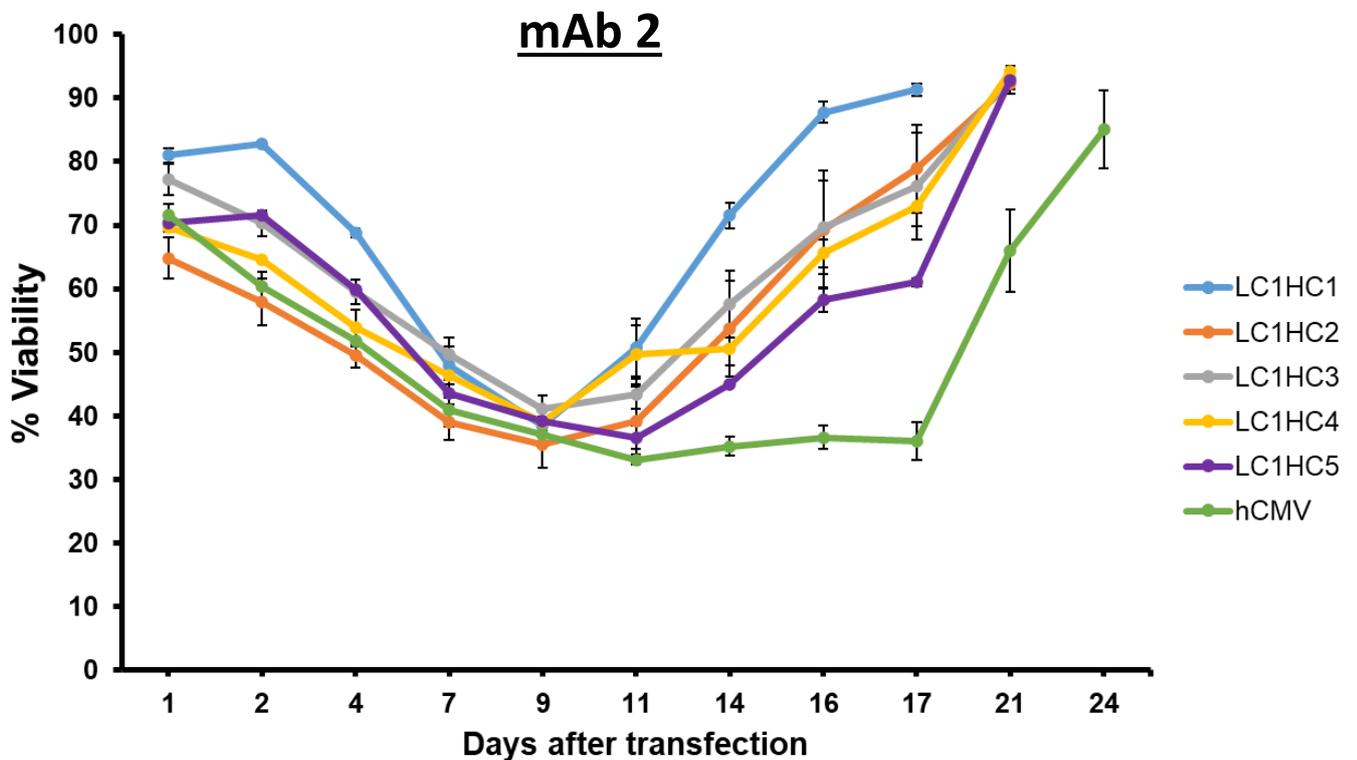


**Figure 4.7 – Effect of LC/HC synthetic promoters on recovery rate of mAb 2-expressing pools**
CHO host cells were transfected with stable expression vectors containing either different combinations of synthetic promoters to control LC and HC gene transcription or two hCMV promoters. Vector-derived expression was selected for by the addition of 50$\mu$M MSX 24 hours post transfection. % viability was measured until full recovery of pools. Data points represent the mean of two (LC1HC4 and LC1HC5) or three separately transfected pools per vector at each time point. Error bars represent standard error of the mean.
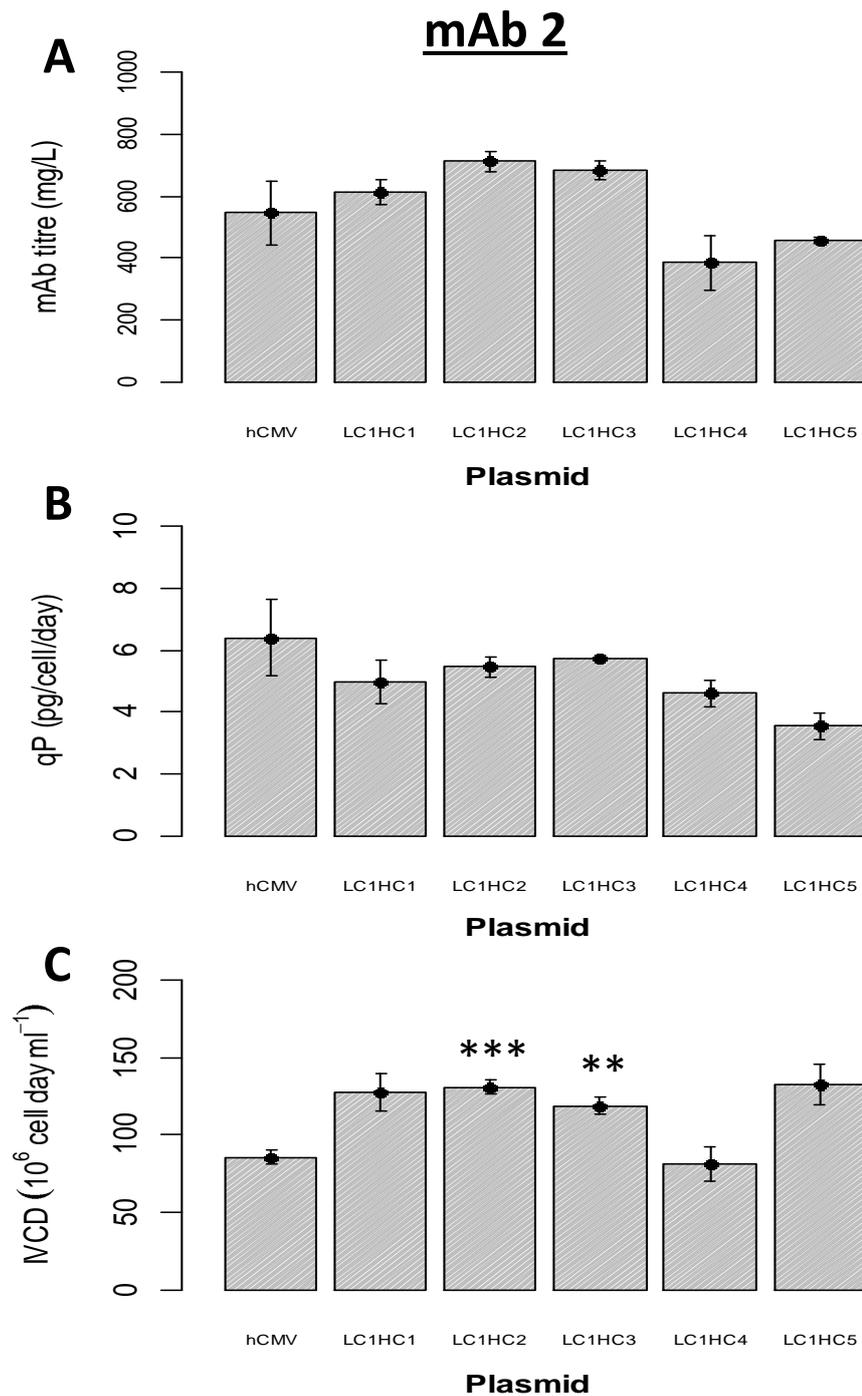
**Figure 4.8 – Effect of LC/HC synthetic promoters on mAb productivity and IVCD of mAb 2-expressing pools**

CHO pools that had been transfected with synthetic promoter-containing mAb 2 expression vectors and selected using MSX went in to 11 day fed batch overgrow cultures. **A)** shows the volumetric mAb titre generated from pools transfected with each vector; **(B)** shows the cell specific productivity (qP) of pools transfected with each vector and **(C)** shows the integral of viable cell density (IVCD) across the 11 day culture period for pools transfected with each vector. Bars represent the mean of two (LC1HC4 and LC1HC5) or three separately transfected pools per vector. Error bars represent standard error of the mean. Mean values significantly different (two-tailed Students t-test) from the mean hCMV value are indicated by asterisks (* $P<0.05$, **$P<0.025$, ***$P<0.01$).

These mAb 2-expressing pools show a delay in the recovery of the hCMV-containing pools compared to the synthetic promoter-containing pools, just as with mAb 1 stable expression. To reach a mean % viability > 70%, the hCMV pools required an extra 10 days compared with LC1HC1, an extra 7 days compared with LC1HC2, LC1HC3 and LC1HC4 and an extra 3 days compared with LC1HC5. There was slightly more variation in the recovery rate of synthetic promoter-containing mAb 2 pools compared to with mAb 1, with LC1HC1 pools recovering quicker than pools generated using the other four plasmids.

Following recovery of mAb 2-expressing pools from MSX selection, these pools were then seeded into 11 day FBOG cultures to assess their mAb productivity (figure 4.8). Vectors LC1HC1, LC1HC2 and LC1HC3 all generated higher mean mAb 2 expression titres than the hCMV vector. These increased volumetric titres compared to hCMV pools came through relative improvements in IVCD and despite relative reductions in qP. As with the mAb 1-expressing pools (figure 4.6), hCMV generated the highest mean qP of all six expression vectors. However, hCMV generated the second lowest mean IVCD of all six expression vectors with only LC1HC4 generating lower. This is similar to the mAb 1-expressing pools in which hCMV produced the lowest mean IVCD. It is notable that mAb 2 was generally much more difficult for CHO host cells to produce than mAb 1, with each LC/HC promoter combination generating a lower mean qP with mAb 2 than with mAb 1.

### 4.2.4 Effect of LC/HC synthetic promoters on stable LC:HC transcriptional ratios

In order to understand how LC/HC synthetic promoters were affecting LC and HC transcript levels in the mAb 1 and mAb 2-expressing stable pools described above, qRT-PCR assays were performed. RNA was extracted from stable pools and reverse transcribed into cDNA. These cDNA samples were used in qRT-PCR assays. LC and HC mRNA copy numbers were estimated using the absolute quantification standard curve method. Standard curves used for quantification are shown in the appendix. These copy numbers were then used to deduce the LC:HC transcriptional ratios of pools generated using different expression vectors (figure 4.9).

The data in figure 4.9 show that there is, broadly, a positive correlation between the predicted LC:HC transcriptional ratios given by synthetic promoter combinations (table A4) and the actual LC:HC transcriptional ratios. This is true for both mAb 1 and mAb 2-expressing pools, although the slope and trajectory of the red
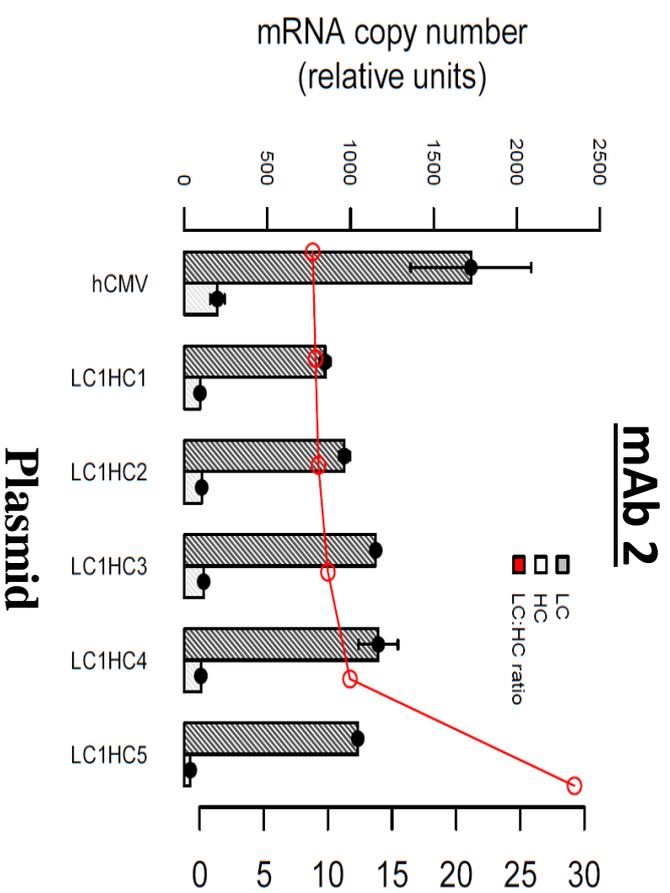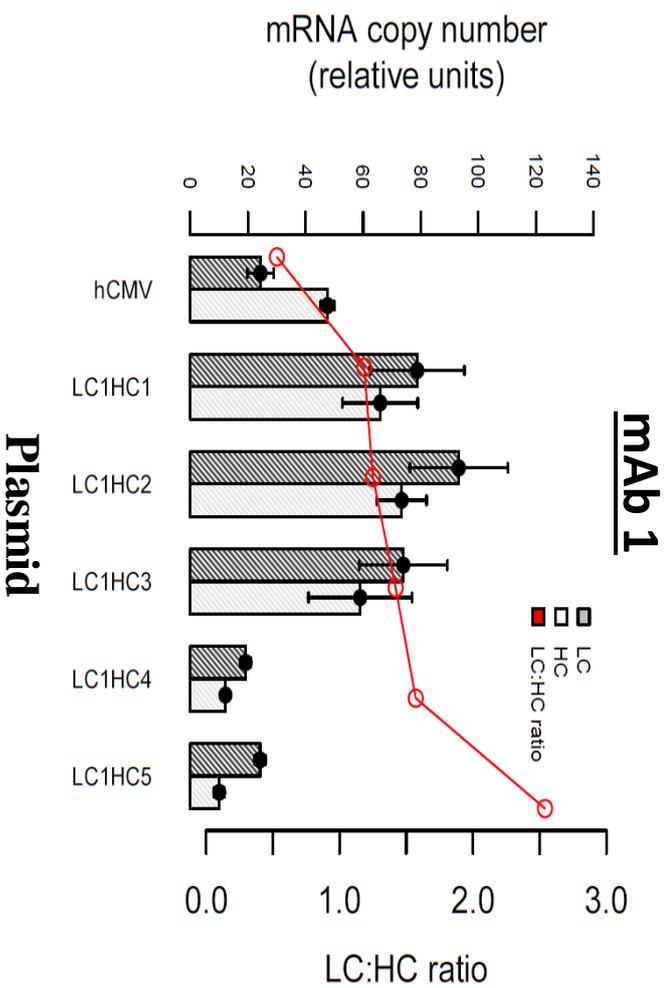
**Figure 4.9 – Effect of LC/HC synthetic promoters on LC and HC mRNA abundance in stably expressing pools**
RNA was extracted from stably expressing cell pools generated using expression vectors containing LC/HC synthetic promoters. RNA was reverse transcribed in to cDNA and qRT-PCR was run on cDNA samples. LC and HC mRNA copy numbers were estimated using the absolute quantification standard curve method. Bars represent the mean copy number of two (mAb 1 LC1HC2; mAb 1 LC1HC4; mAb 2 LC1HC4 and mAb 2 LC1HC5) or three stable pools per expression vector. Error bars represent standard error of the mean. Data points along the red LC:HC ratio line represent the mean LC:HC mRNA ratio for pools generated using that expression vector.

lines connecting mean LC:HC transcriptional ratio values is not the same for the two mAbs. The hCMV expression vector generated the lowest mean LC:HC transcriptional ratio for both mAbs.

There are substantial differences in the LC:HC transcriptional ratios of mAb 1 and mAb 2-expressing pools generated using the same synthetic promoter combination, with the ratio being consistently much greater for mAb 2. Furthermore, there is clear variation in the level of LC gene transcription among synthetic promoter-containing pools expressing the same antibody despite the same LC1 synthetic promoter being used to drive LC transcription. These results may suggest that the inherent LC:HC transcriptional ratio encoded by the promoter combination is capable of being moulded by the selective pressure on these pools towards a ratio that is closer to optimal for the particular antibody being expressed.

### 4.2.5 Effect of LC/HC promoter combination on cell-to-cell variability in LC and HC expression level

The experiments described above in subsection 4.2.4 identified the overall level of LC and HC gene transcription in each stably expressing cell pool as a whole. The variability in LC and HC expression level from cell-to-cell within each of these cell pools was yet to be revealed. Ultimately, clinical supply of mAb molecules must come from monoclonally-derived cell populations. It would be expected generally that relatively productive stable pools would lead to the isolation of relatively productive clones. However, assessing the cell-to-cell variability in LC/HC expression within stable pools may give some indication of how variable the performance of clones taken from that pool would be. A pool in which cell-to-cell performance is highly variable may require the screening of more clones in order to find one that is highly productive.

In order to assess cell-to-cell variability in LC and HC expression, stable pool cell samples were stained with fluorescent antibody conjugates against LC and HC polypeptides and analysed on a cell-by-cell basis using flow cytometry. mAb 1 and mAb 2 expressing pools containing the LC/HC synthetic promoter combination LC1HC1 were chosen for this analysis due to them producing higher mean titres than hCMV pools for both mAbs. hCMV-containing pools were also analysed for comparison. Mean LC and HC expression cell-to-cell variability (expressed as % coefficient of variation) is displayed in figure 4.10.

With respect to expression of mAb 1 LC, the LC1HC1 pools had higher mean
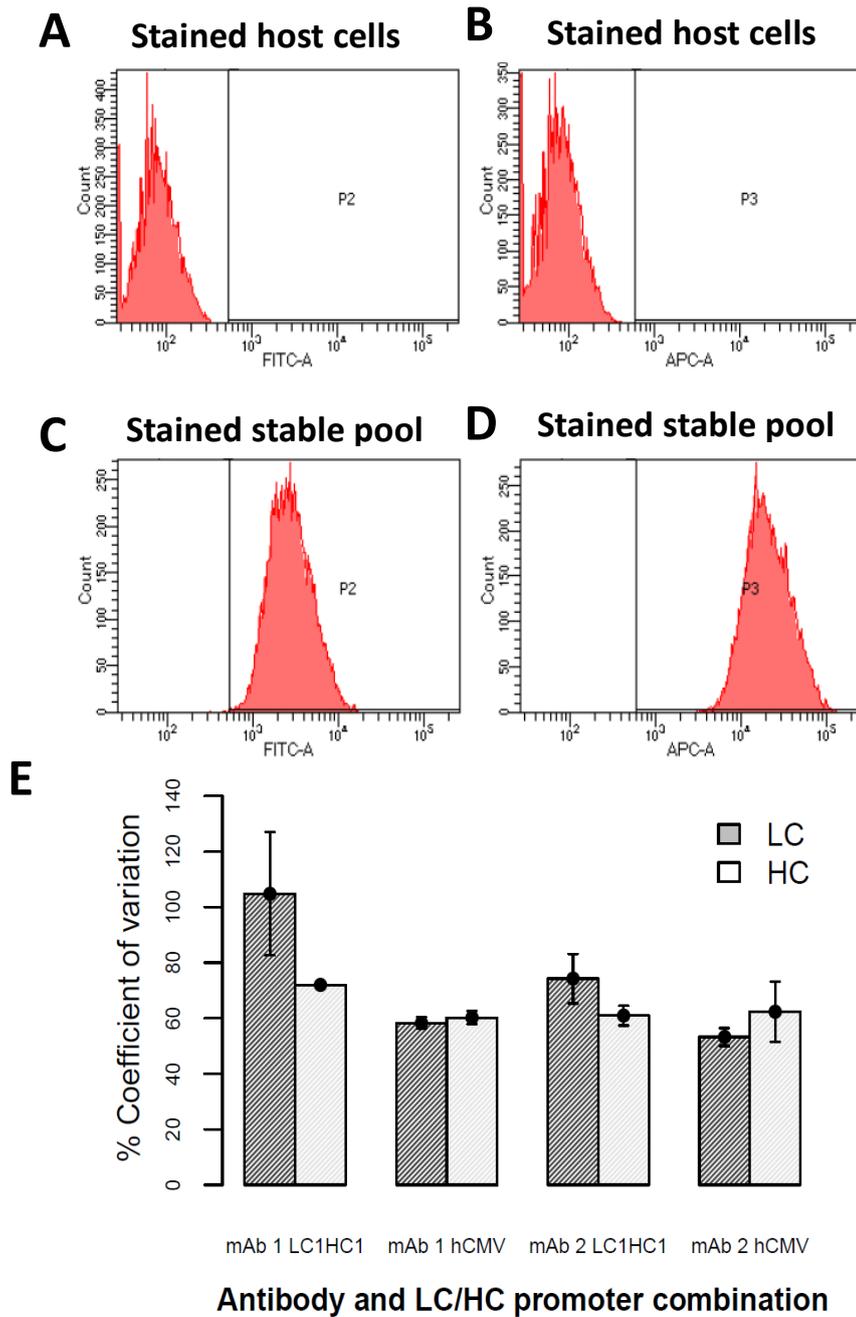
**Figure 4.10 – Cell-to-cell variation in LC and HC expression levels of stable pools**

Cell samples taken from pools stably expressing mAbs 1 and 2 were stained with anti-LC antibodies conjugated to FITC and anti-HC antibodies conjugated to APC. FITC and APC fluorescence levels were analysed on a cell-by-cell basis using flow cytometry, revealing the cell-to-cell variation in expression level of LC and HC respectively for each of these stable pools. Gate P2 identifies FITC/LC positive cells and gate P3 identifies APC/HC positive cells. **A)** and **C)** show example gating of FITC positive/negative untransfected host cells **(A)** and stable pool cells **(C)**. **B)** and **D)** show example gating of APC positive/negative untransfected host cells **(B)** and stable pool cells **(D)**. **E)** displays the % coefficient of variation in LC and HC expression levels of different stable pools. Bars represent the mean % coefficient of variation of three separately transfected pools per antibody and LC/HC promoter combination. Error bars represent standard error of the mean.

cell-to-cell variability than the hCMV pools, albeit this difference was not statistically significant (two-tailed Students t-test P value = 0.16). Likewise the mean cell-to-cell variability in mAb 1 HC expression was significantly higher in LC1HC1 pools than in hCMV pools (two-tailed Students t-test P value = 0.017). With respect to expression of mAb 2 LC, again the mean cell-to-cell variability is higher in LC1HC1 pools than in hCMV pools (two-tailed Students t-test P value = 0.14). Finally, the mean cell-to-cell variability in mAb 2 HC expression was very similar between LC1HC1 and hCMV pools. Overall, these data suggest that the LC1HC1 promoter combination may lead to increased cell-to-cell variability in LC/HC expression compared to hCMV. This could mean that LC1HC1-containing pools will have greater variance in clone productivity compared to hCMV pools.

## 4.3 Discussion

This results chapter describes, for the first time, the use of synthetic promoter combinations to stably express the LC and HC of monoclonal antibodies. These synthetic promoter combinations displayed improved stable expression characteristics over conventionally-used hCMV promoters. Namely, synthetic promoters were capable of producing higher mean mAb titres than hCMV while recovering from MSX selection more quickly. These reduced recovery times will help to shorten cell line development timelines, an extremely desirable outcome for a company that manufactures biopharmaceuticals from CHO cells.

### 4.3.1 Expressing LC in excess of HC may be more important in a stable expression mode than in a transient expression mode

The experiments described above showed that expressing LC in excess of HC meant that HC gene product was more efficiently utilised in assembly of full antibody structures. This greater HC utilisation may be more beneficial towards mAb expression titre in a stably expressing cell compared to in a transiently expressing cell. If, for example at more balanced LC:HC expression ratios, the level of HC utilisation is relatively low then this implies that some HC polypeptide is not being assembled into full mAb structures within the ER and instead remains unassembled. Build-up of this unassembled HC polypeptide may eventually trigger induction of the UPR (Vanhove et al., 2001).

In a stably expressing cell that has been transfected with a LC/HC construct and

high vector-derived expression has been selected for using a selection marker such as the GS system, the level of HC gene expression will generally be substantially higher compared to in a cell that has been transiently transfected with a LC/HC construct. If a relatively balanced LC:HC expression ratio means that $x$% of HC polypeptide expressed is remaining unassembled, this $x$% will be a greater amount of HC in absolute terms in a stably expressing cell compared to in a transiently expressing cell. This means that the stable expression of this LC:HC ratio is more likely to lead to UPR induction. Therefore, in a stable expression mode, it may be important to ensure that almost all HC polypeptide is assembled into full mAb structures by expressing LC in excess of HC. This may prevent UPR induction and thus ultimately ensure greater mAb expression titres. This idea is summarised in figure 4.11.
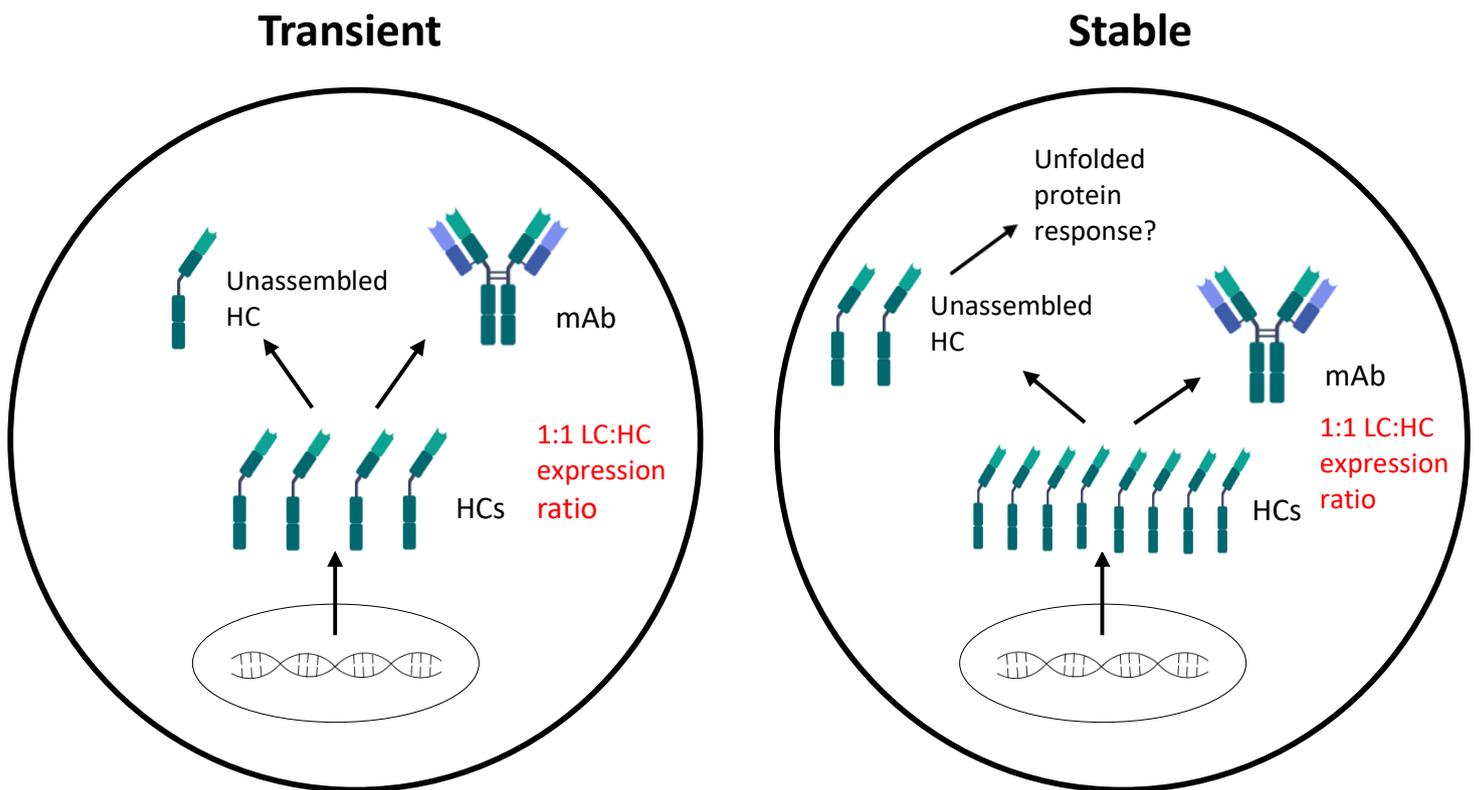


**Figure 4.11 – Stable expression of balanced LC:HC ratios may be more likely to lead to UPR induction than transient expression of the same balanced ratio**

In this example the 1:1 LC:HC expression ratio causes 25% of HC polypeptide expressed to remain unassembled.

### 4.3.2 Comparison with a previous study

Data shown in this chapter revealed vastly different optimal LC:HC transcriptional ratios for the two mAbs that were stably expressed. It is unsurprising that the optimal transcriptional ratio is mAb-specific: a previous study transiently transfected different LC:HC plasmid ratios in CHO and found that the ratios giving highest volumetric titre for three different mAb molecules were 9:1, 4:1 and 2.3:1 (Pybus et al., 2014). The mAb with an optimal transiently transfected LC:HC plasmid ratio of 9:1 was mAb 2 used in this chapter. This 9:1 figure is very similar to the 9.3:1 mean LC:HC transcriptional ratio that was found to give the highest mean volumetric mAb 2 titre in this chapter (given by LC1HC2 stable pools).

Although the results from Pybus et al support the mAb 2 stable expression results acquired here, the beneficial effect on titre of transiently transfecting LC in excess of HC that Pybus et al found was not replicated in the mAb 2 transient transfections shown in figure 4.4 of this chapter. This is likely due to differences in experimental scale and set up. The transfected DNA load in the experiment shown in figure 4.4 was ~ $0.4\mu g/1x10^6$ viable cells while in the experiment performed by Pybus et al it was ~ $1.25\mu g/1x10^6$ viable cells. Moreover, unlike in the experiment shown in figure 4.4, Pybus et al used a host cell line that has been engineered for optimal transient production by EBNA1-based plasmid retention (Daramola et al., 2014). Although plasmid sizes, efficiency of transfection methods and expression vector components used would need to be taken into account for a true comparison between experiments, it is highly likely based on the use of a substantially higher DNA load and an optimised cell line that the amount of HC being expressed in each cell is higher in the Pybus et al experiment than in the experiment shown in figure 4.4. This greater HC expression in the Pybus et al experiment may mean there is a much greater risk of UPR induction, compared to in the experiment in figure 4.4, when expressing LC and HC at a relatively balanced ratio. It may therefore be important to express LC in excess of HC in the Pybus et al experiment, but not in the figure 4.4 experiment, to avoid UPR induction and ensure high levels of productivity.

The transient transfections from figure 4.3 and figure 4.4 of this chapter showed that transfecting LC in excess of HC caused an increase in HC utilisation efficiency for both mAb 1 and mAb 2 (although this did not necessarily lead to increased mAb titre). This phenomena was also shown in previous studies - Schlatter et al showed that the

efficiency of conversion of HC polypeptide to mAb increased as the amount of LC gene transfected was increased around a constant HC transfected gene load (Schlatter et al., 2005). Furthermore, mathematical modelling of mAb folding and assembly revealed that assembly time decreased as the LC:HC ratio increased (Gonzalez et al., 2002). This is likely due to an excess of LC polypeptides in the ER ensuring rapid release of HC polypeptides from BiP (Vanhove et al., 2001).

### 4.3.3 LC/HC synthetic promoters may allow more flexible control over the LC:HC transcriptional ratio than hCMV

qRT-PCR assays performed in this chapter on mAb 1 and mAb 2 stable pools suggest that the selective pressure on these pools was, along with the LC/HC promoter combination being used, influential over the LC:HC transcriptional ratio achieved. It appears that the inherent LC:HC transcriptional ratio encoded by the LC/HC promoter combination is pushed by the selective pressure towards a ratio that is closer to optimal for the particular mAb being expressed.

It may be that because LC and HC synthetic promoters have no overlap in their TFRE usage, the LC:HC transcriptional ratios given by synthetic promoter combinations are more flexible than those given by two copies of hCMV. One way in which two cells that contain the same LC/HC promoter combination could transcribe LC and HC at different ratios is if the expression level of key TFs varies between cells. For example, a mAb 2-expressing cell that has relatively high expression of TFs driving the LC promoter will produce an elevated LC:HC transcriptional ratio and may therefore be selected for within the population. In contrast, when hCMV is used to transcribe both LC and HC the selection process does not have access to this mechanism of altering the LC:HC transcriptional ratio as any alteration in TF abundance would affect both promoters equally. It has long been known that CHO cell populations are extremely genetically diverse, especially following transfection to create a recombinantly-expressing pool (Derouazi et al., 2006; Wurm & Wurm, 2017). Although no direct evidence can be found for variation within cell populations in expression level of the TFs that drive LC/HC synthetic promoter output, there is no reason not to believe that this general genetic diversity would lead to variation in the expression of these TF genes.

### 4.3.4 LC/HC synthetic promoters may give greater transfection efficiency than hCMV

The synthetic promoter sequences used in this chapter for LC and HC gene transcription are all shorter in nucleotide length than hCMV. This is not unexpected, since these promoters were designed specifically for CHO cell activity by assembly of TFRE blocks known to be active in CHO. There is therefore likely to be very little functional redundancy in these sequences. Viral promoters like hCMV, on the other hand, have evolved to enable infection of a range of host cell types by utilising their different sets of expressed TFs (Sinzger et al., 2008). As a result, the hCMV sequence has substantial functional redundancy in CHO cells with many TFREs in hCMV known to be inactive in CHO (Brown, 2014). The reduction in size of the synthetic promoters compared to hCMV will lead to a marginal improvement in transfection efficiency, as plasmid copy number per mass of plasmid DNA transfected will be higher.

### 4.3.5 hCMV is likely to create greater selective pressure on cell pools than LC/HC synthetic promoters

The data in this chapter show the effect of using hCMV promoters to control LC and HC transcription compared to using synthetic promoters to be consistent across stable expression of two different mAbs. Specifically, hCMV-containing pools showed slower recovery from MSX selection than synthetic promoter-containing pools but had higher qP once recovered. This consistent effect may be explained by the impact of LC/HC promoters on SV40 promoter activity and therefore the different levels of selective pressure on transfected cell populations. The hCMV promoter is known to have substantial overlap with the SV40 promoter in the TFREs that they contain, leading to competition between the two promoters for binding of TFs within the nucleus. Analysis of the SV40 promoter sequence using the TF binding site search tool Matinspector found predicted binding sites corresponding to 57 different TF families. Analysis of the hCMV promoter sequence found that 33 of these 57 TFRE families were also predicted to bind within hCMV. The result of this TF binding competition is a weakening of SV40 promoter activity - comparison of co-transfecting an hCMV-driven marker alongside an SV40-driven marker against transfecting just the SV40-driven marker alone showed that the presence of hCMV reduced SV40 mRNA output by 2.69 to 5.15 fold (West, 2014). Furthermore, in this experiment hCMV and SV40 promoters were co-transfected at a 1:1 copy number ratio. When a mAb expression vector containing

two hCMV promoters per SV40 promoter is used, the reduction in SV40 promoter activity will likely be even greater.

In comparison, the LC/HC synthetic promoters used in this chapter have little overlap with the SV40 promoter in TFRE use. Analysis using Matinspector showed that of the 57 different TF families predicted to bind within the SV40 promoter, 3 of these were predicted to bind within the LC1 promoter and 4 of these were predicted to bind within the HC synthetic promoter designs.

This reduction in SV40 promoter activity by hCMV but not by LC/HC synthetic promoters will lead to more stringent selective pressure on cell pools transfected with the hCMV-containing expression vector. In order to reach the threshold level of GS expression required to overcome the MSX inhibition and allow cell division, hCMV-transfected cells will need to have higher GS production capability than LC/HC synthetic promoter-transfected cells. This higher GS production capability could come from: the vector construct integrating into a highly transcriptionally active genomic site; genomic integration of many copies of the vector construct or superior capability in a downstream protein production step such as translation. Any of these cellular mechanisms of higher GS production capability should extend to higher LC and HC production capability. Therefore, the increased selection stringency on hCMV-transfected pools may explain why they generate higher mean qP than synthetic promoter-transfected pools. Additionally, this increased selection stringency is likely to explain the slower recovery of hCMV pools from MSX selection, as a smaller proportion of the hCMV-transfected cell population will be capable of meeting the required GS expression threshold than in the LC/HC synthetic promoter-transfected cell population.

# Chapter 5
# <u>Use of synthetic promoters to control GS transcription</u>

*This chapter describes, for the first time, the use of synthetic promoters to control GS gene transcription within stably expressing CHO cell pools. GS synthetic promoters were designed to be less transcriptionally active than the conventionally-used SV40 promoter and were utilised in the generation of stable pools expressing a mAb molecule. Stable pools containing the GS synthetic promoter SynSV40_2 alongside synthetic promoters to control LC/HC transcription showed improved qP compared to those containing SV40 promoters alongside the same LC/HC synthetic promoters. Results within this chapter also indicate that transfected cell populations may be capable of adapting to MSX inhibition through mechanisms other than high vector-derived GS expression. qRT-PCR assays were performed and results suggest that one of these mechanisms may be upregulation in expression of the endogenous CHO host cell GS gene.*

## 5.1 Introduction

Ensuring highly stringent selective pressure on stably transfected cell populations can be an effective way of enriching this population with cells that have high level expression from the vector construct. If very high expression of a vector-derived selection marker gene is required in order for a transfected cell to survive and divide in culture then the expression level of the recombinant product genes, such as mAb LC and HC genes, should also be very high. Any genetic event that might ensure high selection marker expression, such as incorporation of the vector construct into a transcriptionally active genomic site or incorporation of many vector copies into the genome, would extend to high recombinant product gene expression and therefore high qP. There has therefore been considerable research focus on the development of methods to increase selection stringency on stably transfected CHO cell populations.

One factor that will influence the stringency of selection on transfected cell populations is the level of endogenous selection marker gene expression. For example with the GS/MSX selection system, the contribution of any endogenous host cell GS

expression will reduce the level of vector-derived GS expression required to overcome the MSX inhibition. To this end, biallelic GS-KO cell lines have been developed to ensure that the requirement for vector-derived expression is greater (Fan et al., 2012).

Various vector engineering approaches have been taken to try to increase selection stringency by restricting the expression of functional selection marker protein. For example one study codon de-optimised the DHFR selection marker gene so that it was translated less efficiently in CHO, resulting in greater recombinant product titres (Westwood et al., 2010). However another study found that codon-deoptimising the DHFR gene did not lead to a reduction in DHFR protein, with the authors speculating that cell populations had adapted by altering tRNA expression levels such that the transcript was translated with high efficiency again, reversing the effect of the codon deoptimisation (Chin et al., 2015). Another approach taken has been to introduce destabilising motifs such as PEST domains and AU-rich elements to ensure quicker degradation of selection marker protein and mRNA respectively (Chin et al., 2015; Ng et al., 2007). Both of these destabilising motifs were successfully shown to increase product titres when introduced into DHFR genes. Single amino acid mutations that reduce selection marker protein activity, including in GS and the NPT selection marker, have also been effective at increasing recombinant product titre of transfectants (Lin et al., 2019; Sautter & Enenkel, 2005).

Attempts have also been made to increase selection stringency on transfected CHO populations by weakening the transcriptional activity of the SV40 promoter driving GS transcription (Fan et al., 2013). A series of mutant SV40 promoters were created by Fan et al through the deletion of 5' portions of the wild type promoter. Controlling GS transcription with these mutant SV40 promoters led to improvement in recombinant product titres.

Although these crude methods of weakening functional selection marker protein expression have been successful in increasing product titres, a technology that enables the precise manipulation of selection marker expression may allow identification of the optimal level that fully maximises recombinant product titre. The design of synthetic promoters to control selection marker gene transcription may be capable of meeting this requirement. Use of a synthetic promoter in place of the conventional SV40 promoter also has the added benefit of being likely to marginally reduce expression vector size and thereby increase transfection efficiency.

In this chapter, synthetic promoters were used for the first time to control GS

selection marker gene transcription in stably expressing CHO cell pools. Adaptation of mAb-expressing stable pools to a range of MSX concentrations was first carried out to understand how these different levels of selection stringency would affect mAb productivity. Synthetic promoters (the design of which is described in chapter 3) predicted to be less transcriptionally active than the SV40 promoter were used to control GS transcription. Expression vectors containing these promoters were transfected into CHO host cells to generate mAb-expressing stable pools. One GS synthetic promoter, named SynSV40_2, was shown to increase qP of stable pools compared to the SV40 promoter but only when combined with LC/HC synthetic promoters. Interestingly, results within this chapter also indicate that transfected cell populations may be capable of adapting to MSX inhibition by mechanisms other than high vector-derived GS expression and qRT-PCR assays performed suggest that one of these mechanisms may be upregulation in expression of the endogenous CHO host cell GS gene.

## 5.2 Results

### 5.2.1 Adaptation of mAb 1-expressing stable pools to greater concentrations of MSX

hCMV promoter-containing mAb 1 stable pools displayed a significantly greater mean qP and a reduced mean IVCD compared to LC/HC synthetic promoter-containing mAb 1 stable pools (figure 4.6). The perceived explanation for this, as discussed in chapter 4, is the increased selective pressure on hCMV pools compared to LC/HC synthetic promoter-containing pools due to hCMV-SV40 TF binding competition weakening GS selection marker transcription. Experiments were therefore carried out that attempted to increase the level of selective pressure on LC/HC synthetic promoter-containing pools. It was hypothesised that increasing selective pressure on these pools may improve their qP, albeit possibly reduce their growth rate.

One method of increasing the selective pressure on cell populations that have been transfected with a vector containing the GS selection marker is to increase MSX concentration in culture media. Only cells expressing enough GS enzyme to overcome the MSX inhibition will be capable of dividing. The two most productive triplicate sets of mAb 1 stable pools, containing hCMV and LC1HC1 promoters for LC/HC transcription, were seeded into culture media containing MSX at concentrations of $50\mu M$, $75\mu M$, $100\mu M$ and $125\mu M$. The VCD of these cultures was monitored over 7 days in order to observe differences in rate of outgrowth. These different rates of

outgrowth are representative of different levels of selective pressure - a more stringent selective pressure means a smaller proportion of the cell population is capable of dividing and therefore outgrowth is slower. The mean 7 day IVCD of stable pools cultured in different MSX concentrations is shown in figure 5.1.

There are two conclusions from the data shown in figure 5.1. Firstly, stable pools containing the LC1HC1 LC/HC promoter combination were more easily able to divide in the presence of increased MSX concentrations compared to hCMV stable pools. At each MSX concentration, the LC1HC1 pools gave a higher mean IVCD than the corresponding hCMV pools (two-tailed Students t-test P values of 0.016, 0.18, 0.14 and 0.06 for the $50\mu$M, $75\mu$M, $100\mu$M and $125\mu$M comparisons respectively). Secondly, there is a negative overall correlation between MSX concentration and mean IVCD, with the hCMV $50\mu$M MSX pools being the only condition not conforming to this trend. This indicates that greater MSX concentrations successfully created more stringent selective pressure.

The cultures shown in figure 5.1 were then passaged a further four times in culture media containing the different concentrations of MSX to ensure that cell populations were fully adapted. Following thorough adaptation of cell pools to these different levels of selective pressure they then went into 11 day fed batch overgrow cultures (FBOG) to assess their mAb productivity (figure 5.2).

The data in figure 5.2 show that there is a positive correlation between the level of selective pressure applied to these pools and their mean qP. A Pearson correlation test was run with the X variable being mean 7 day IVCD values of pools in response to different MSX concentrations in culture media (figure 5.1 - a representation of differing levels of selective pressure) and the Y variable being mean FBOG qP (figure 5.2). This test revealed a statistically significant correlation with an r value of -0.89 (P<0.01). Conversely, the data in figure 5.2 show that there is a negative correlation between the level of selective pressure applied to these pools and their mean IVCD. The contrasting effects of increased MSX concentration on qP and IVCD result in mean volumetric titres that are not significantly different among pools containing the same LC/HC promoter combinations but adapted to different MSX concentrations.
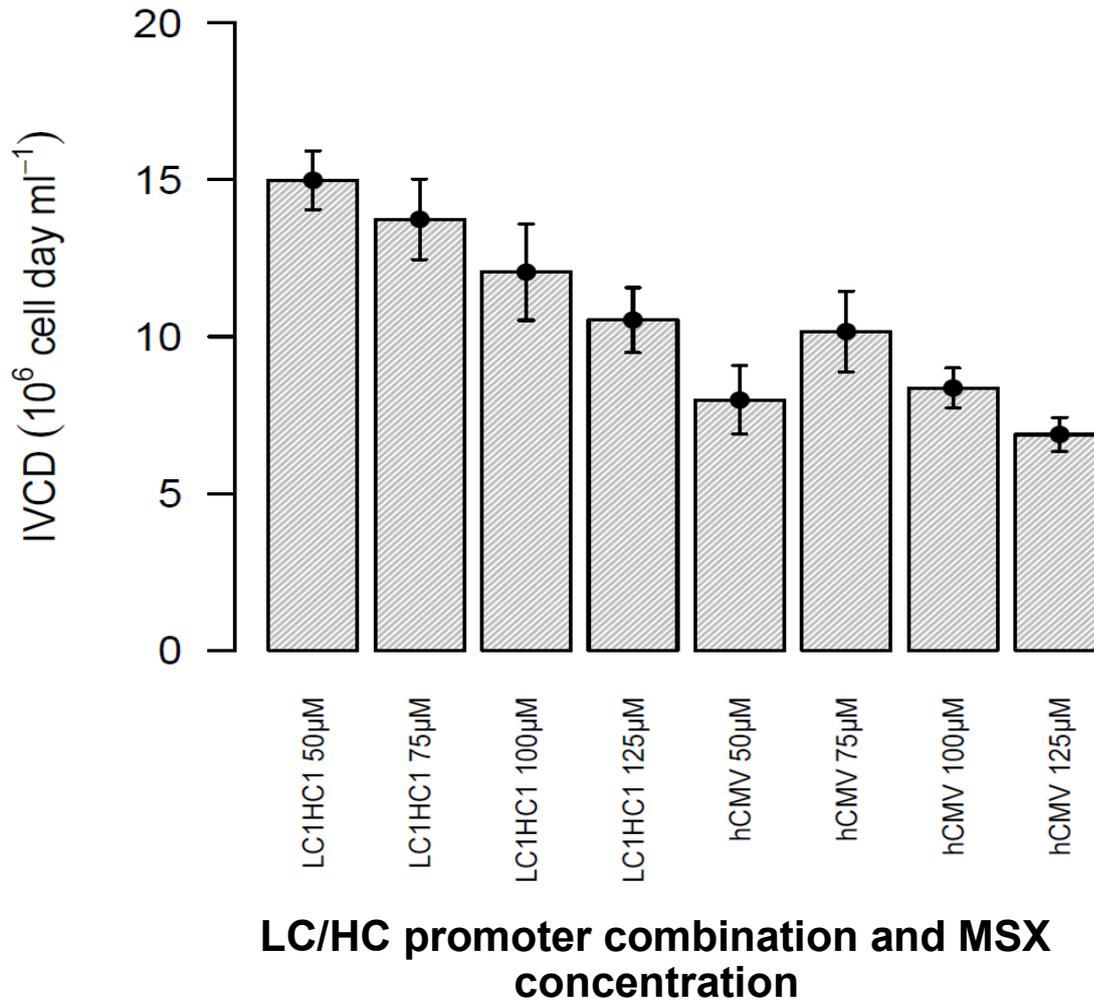
**Figure 5.1 – Adaptation of stable pools to greater MSX concentrations increases the stringency of selective pressure on them**

Cell pools stably expressing mAb 1 through either LC1HC1 or hCMV LC/HC promoter combinations were seeded in to culture media containing MSX at concentrations of 50µM, 75µM, 100µM or 125µM. VCD of cultures was then measured at different time points up to 7 days post seed in order to assess the different rates of outgrowth/different levels of selective pressure. Bars represent the mean day 7 IVCD for 3 pools per LC/HC promoter combination and MSX concentration. Error bars represent standard error of the mean.
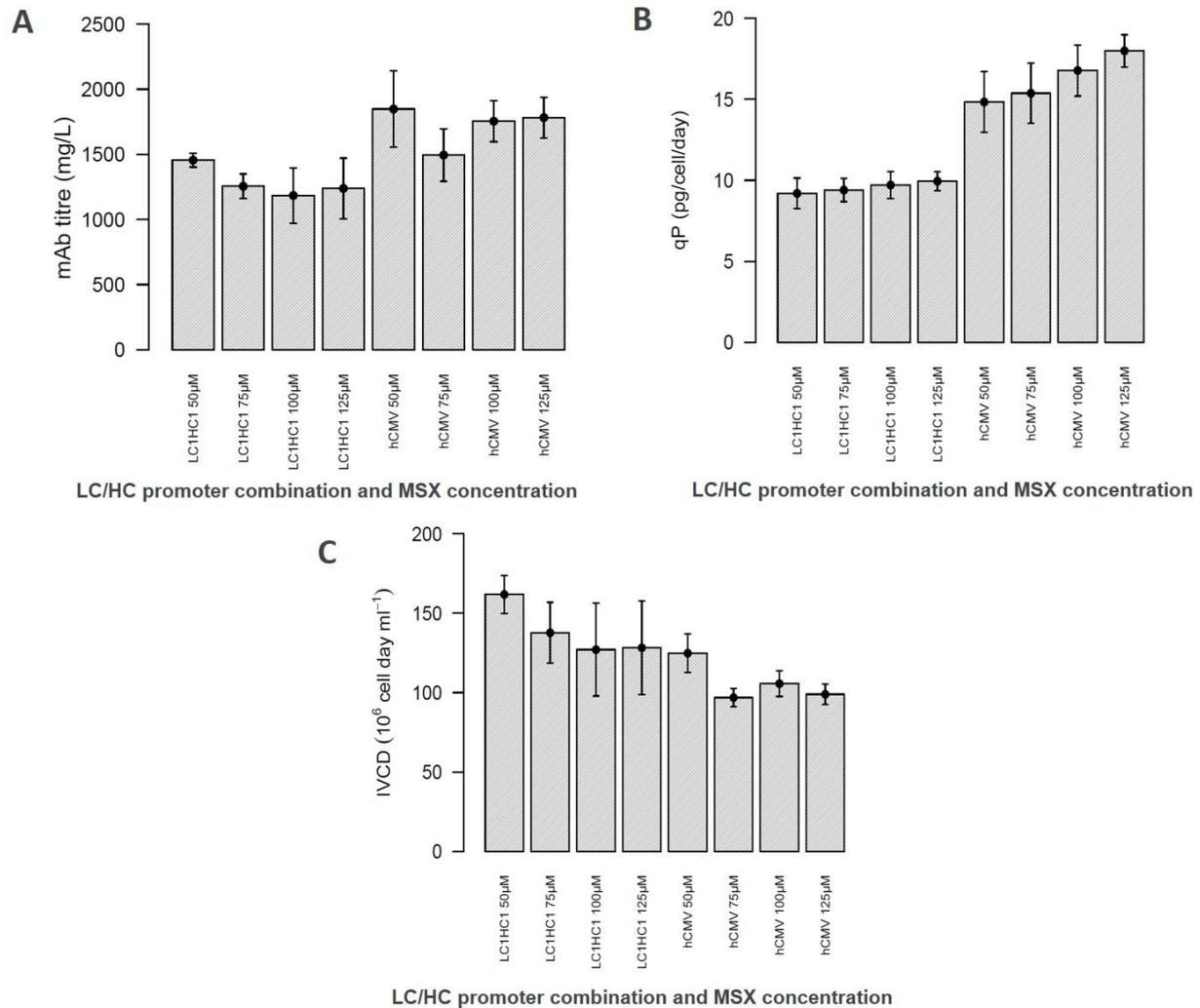
**Figure 5.2 – Effect of adaptation of stable pools to greater MSX concentrations on mAb productivity and IVCD**

Cell pools stably expressing mAb 1 through either LC1HC1 or hCMV LC/HC promoter combinations were adapted to MSX concentrations of 50µM, 75µM, 100µM or 125µM and then went in to 11 day fed batch overgrow cultures. **A)** shows the volumetric mAb titre generated from pools; **(B)** shows the cell specific productivity (qP) of pools and **(C)** shows the integral of viable cell density (IVCD) of pools across the 11 day culture period. Bars represent the mean of three pools per LC/HC promoter combination and MSX concentration. Error bars represent standard error of the mean.

**5.2.2 Effect of GS synthetic promoters on mAb 1 stable pool generation**

In the experiments described above, the selective pressure on stably expressing cell pools was made more stringent by increasing the concentration of MSX in culture media. However, the use of synthetic promoters to replace the SV40 promoter in controlling GS gene transcription may be a more elegant way of increasing selective pressure on cell pools. GS synthetic promoters (the design of which is described in chapter 3) were therefore tested for their effect on stable mAb expression characteristics including the rate of outgrowth of expressing cells and mAb productivity. Synthetic proximal promoters were cloned directly upstream of CMV core sequences and these promoters were used to control GS gene transcription in mAb 1 stable expression vectors. These vectors were tested against a conventional mAb 1 stable expression vector containing an SV40 promoter to control GS gene transcription. hCMV promoters were used to control mAb 1 LC and HC transcription in all vectors to ensure that solely the effect of varying GS transcription levels was being tested. To generate stably expressing pools using these expression vectors, CHO host cells were transfected and 50$\mu$M MSX was added to transfected cultures 24 hours later. The recovery of pools from transfection and MSX selection was monitored (figure 5.3).

The data in figure 5.3 show a clear difference in the recovery rate of SV40 promoter-containing pools compared to pools containing GS synthetic promoters, with the former recovering much more rapidly than the latter. The slower recovery of GS synthetic promoter-containing pools suggests that these promoters did indeed create increased selection stringency on cell populations compared to the SV40 promoter.

Following recovery of transfected cell pools from MSX selection they then went into 11 day FBOG cultures to assess their mAb productivity (figure 5.4). The data in figure 5.4 show that the increased selection stringency conferred by the GS synthetic promoters did not lead to increased mAb productivity of pools. All three SV40 promoter-containing pools produced higher volumetric titres than any of the GS synthetic promoter-containing pools. There is substantial variation in volumetric titres even among pools generated with the same expression vector, which is why the data are displayed as one bar per pool rather than displaying the mean value for each expression vector. GS synthetic promoter pools that were relatively productive compared to the others (GS2 pool 1 and 3; GS4 pool 1 and GS5 pool 2) may contain a heterogenous mixture of subpopulations that have undergone the expected adaptation to the selective
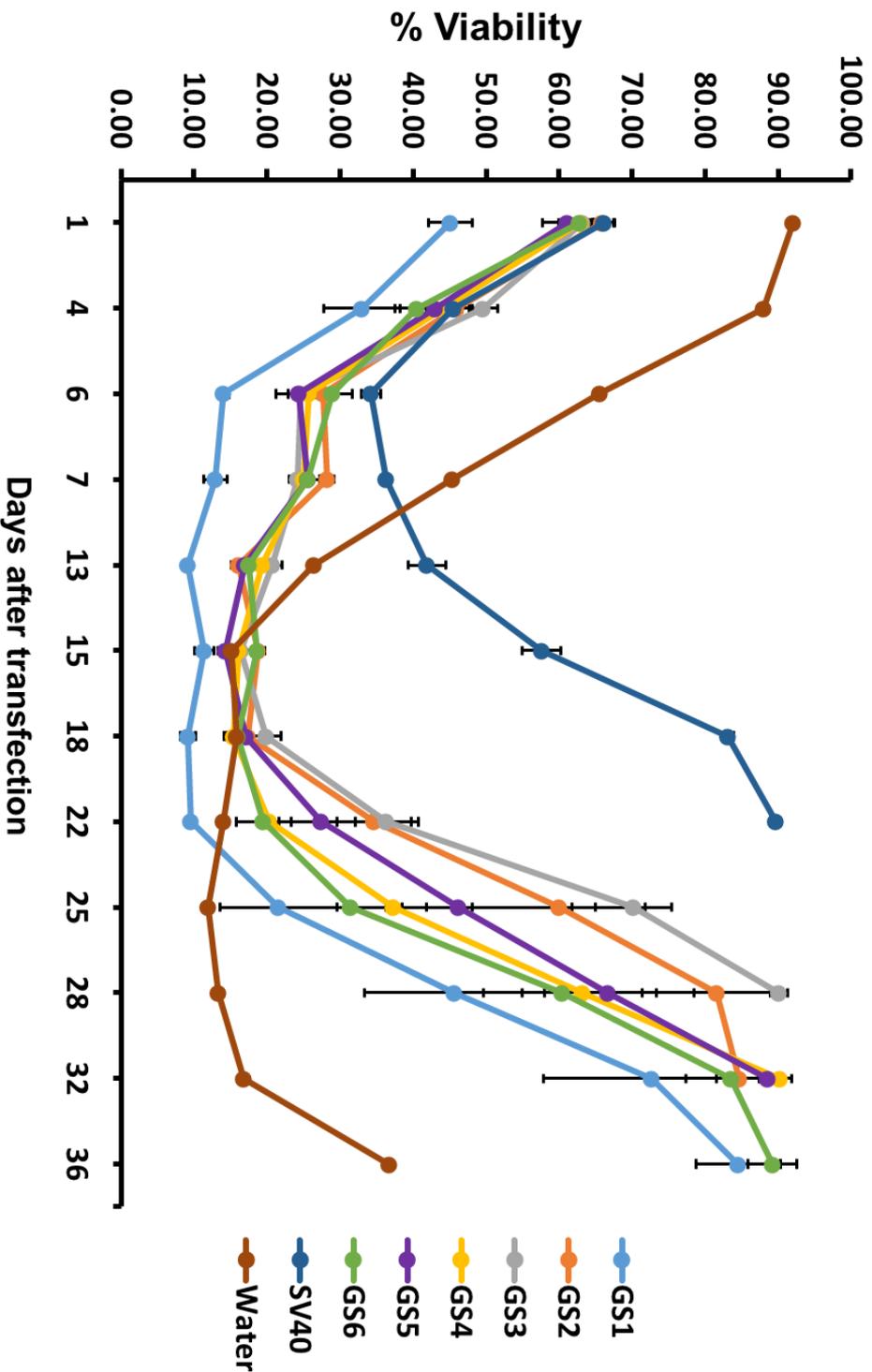
**Figure 5.3 – Effect of GS synthetic promoters on recovery rate of mAb 1-expressing pools** CHO host cells were transfected with stable expression vectors containing either synthetic promoters to control GS gene transcription or an SV40 promoter. hCMV promoters were used to control LC and HC transcription in all vectors. A negative control pool transfected with water in place of plasmid DNA was also included for comparison. Vector-derived expression was selected for by the addition of 50μM MSX 24 hours post transfection. % viability was measured until full recovery of pools. Data points represent the mean of two (GS1) or three separately transfected pools per vector at each time point. Error bars represent standard error of the mean.
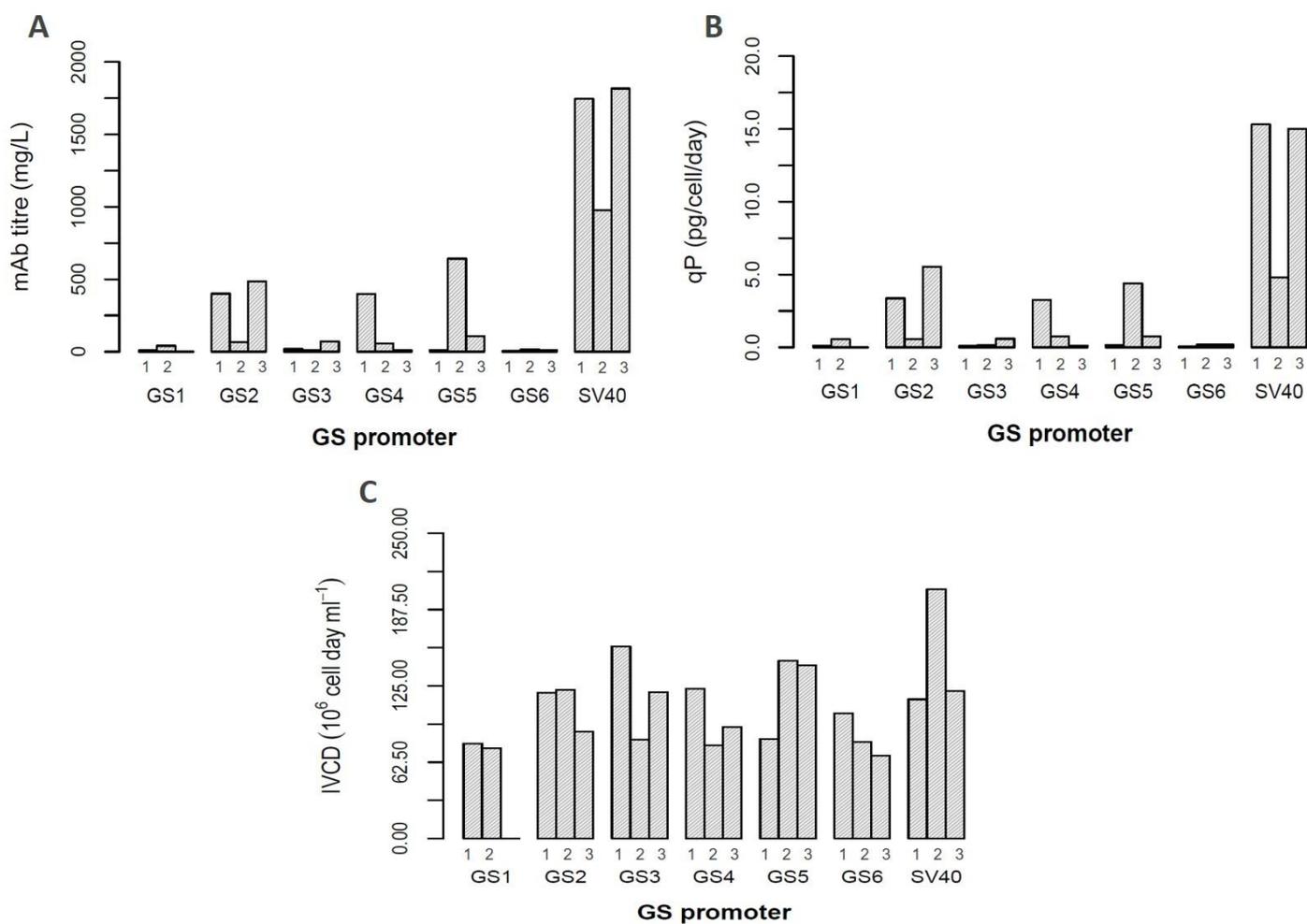
**Figure 5.4 – Effect of GS synthetic promoters on mAb productivity and IVCD of mAb 1-expressing pools**

CHO pools that had been transfected with GS synthetic promoter-containing mAb 1 expression vectors and selected using MSX went in to 11 day fed batch overgrow cultures. **A)** shows the volumetric mAb titre generated from pools transfected with each vector; **(B)** shows the cell specific productivity (qP) of pools transfected with each vector and **(C)** shows the integral of viable cell density (IVCD) across the 11 day culture period for pools transfected with each vector. Each bar represents an individually transfected pool.

pressure, i.e. upregulation of vector-derived GS expression and subpopulations that have instead undergone some alternative adaptation. Calculation of qP and IVCD generated by each stable pool reveals that differences in volumetric titre are mostly explained by differences in qP.

### 5.2.3 Endogenous GS mRNA levels in GS synthetic promoter-containing stable pools

The data in figures 5.3 and 5.4 show that, while GS synthetic promoter-containing pools were capable of recovering from the MSX selection, this adaptation did not appear to have come from the expected mechanism of upregulating vector-derived GS expression (along with LC and HC gene expression). It may be that these GS synthetic promoters were so transcriptionally inactive that transfected cells were unable to overcome MSX inhibition from vector-derived GS expression, regardless of how transcriptionally active genomic integration sites were or how many vector copies were integrated. This therefore suggests that the cell populations transfected with GS synthetic promoter vectors may have undergone an alternative adaptation to the MSX selection. Furthermore, the negative control water transfected cell pool was seemingly starting to recover towards the end of the culture period in figure 5.3 and this may also be indicative of possible alternative adaptations. An obvious alternative adaptation might be upregulation of expression of the endogenous host cell GS gene.

In order to examine the endogenous GS mRNA levels in GS synthetic promoter-containing pools, qRT-PCR assays were performed. RNA was extracted from stable pools and reverse transcribed into cDNA and these cDNA samples were used in qRT-PCR assays. The $\Delta\Delta$CT method was used to calculate fold changes in endogenous GS mRNA levels relative to the pool that produced the highest volumetric mAb titre - SV40 pool 3. The TaqMan probes used in this assay were shown to selectively amplify endogenous GS cDNA but not the vector-derived GS cDNA. Appendix figure A7 shows amplification of a serial dilution of cDNA generated from untransfected CHO host cell RNA with almost perfect linearity. The same TaqMan probes were unable to amplify a serial dilution of a linearised plasmid containing the selection marker GS gene (data not shown). The different endogenous GS mRNA levels of individual pools are shown in figure 5.5. Sufficient quantities of RNA could not be extracted from GS1 pool 1 and GS6 pool 1 and so qRT-PCR assays could not be run for these pools.

The data in figure 5.5 show that, similarly to the titre results displayed in figure
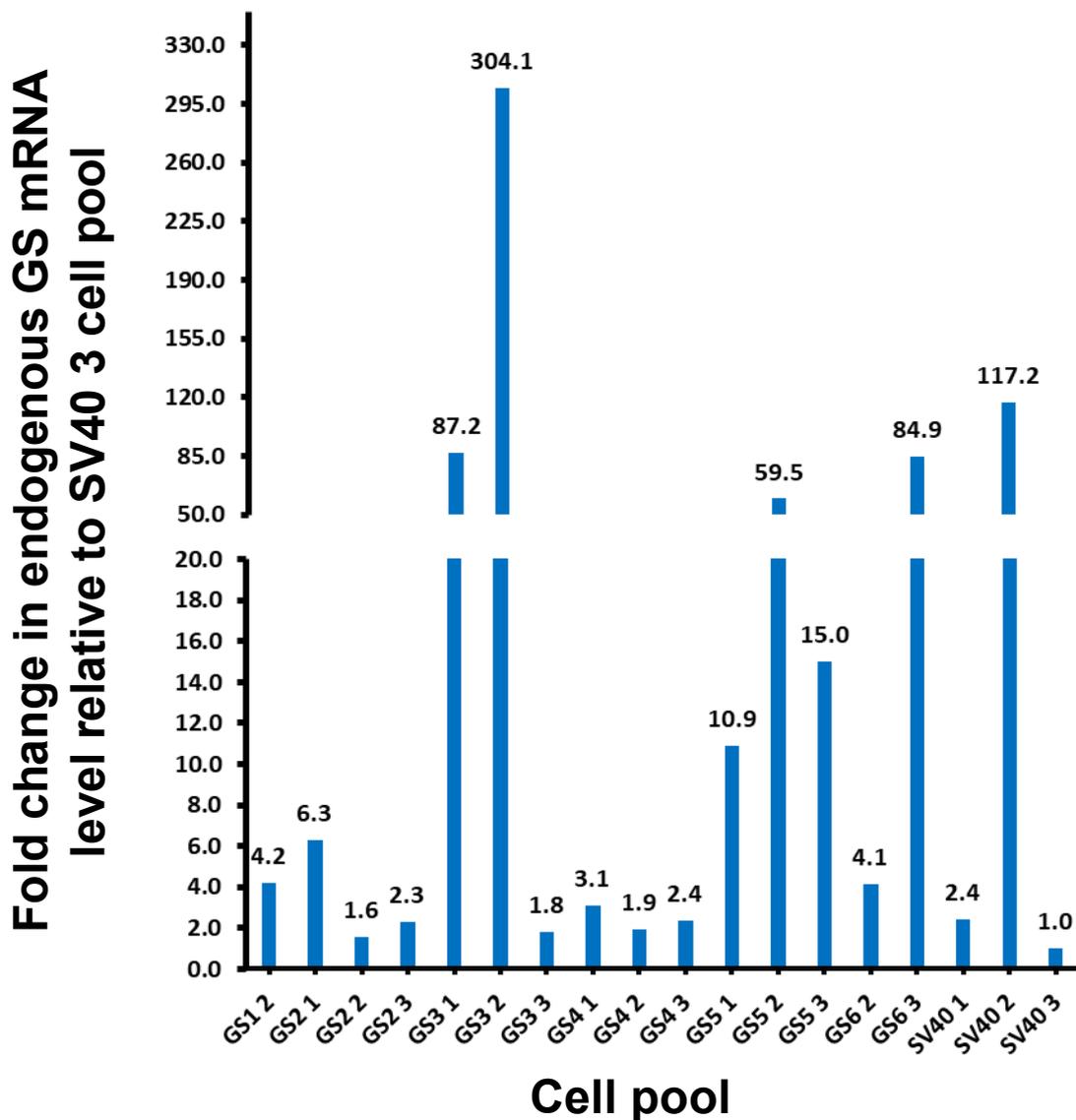
**Figure 5.5 – Endogenous GS mRNA levels from GS synthetic promoter-containing stable pools**

RNA was extracted from stably expressing cell pools generated using expression vectors containing GS synthetic promoters. RNA was reverse transcribed in to cDNA and qRT-PCR was run on cDNA samples. The ΔΔCT method was used to calculate fold change relative to the SV40 pool 3 sample. Each bar represents an individually transfected pool.

5.4, there is substantial variation between pools in endogenous GS mRNA levels, even among pools generated using the same expression vector. The recovery of some pools from MSX selection is perhaps explained by an upregulation in endogenous GS expression. For example GS3 pool 2 showed very low mAb productivity (figure 5.4) which is indicative of low vector-derived GS expression but showed extremely high endogenous GS mRNA levels, the highest of any stable pool. Conversely, the recovery of some pools from MSX selection can not be explained by an upregulation in endogenous GS expression. For example, GS2 pool 2 showed low mAb productivity but also has low endogenous GS mRNA levels, the lowest of any GS synthetic promoter-containing pool. Pools such as these must have made an alternative adaptation to expressing high amounts of either vector-derived or endogenous GS.

It is interesting to note that SV40 pool 2 displayed much lower mAb productivity than the other two SV40 promoter-containing pools and also had much higher endogenous GS mRNA levels. This may suggest that upregulation of endogenous GS expression can occur as an alternative adaptation to high vector-derived GS expression even in conventional pools containing SV40 promoters to drive GS transcription. These results may advocate the use of GS-KO CHO host cell lines, as the lack of an endogenous GS gene would improve the fidelity of the GS/MSX selection system.

## 5.2.4 Effect of 'synthetic SV40' promoters and LC/HC synthetic promoters on mAb 1 stable pool generation

The use of the GS synthetic promoters described above in stable pool creation led to increased selection stringency on transfected cell pools but this did not lead to improved mAb productivity, likely due to these promoters being too weak. The TFRE 'building block' design strategy was therefore ineffective in this instance. An alternative design strategy (described in detail in chapter 3) was then used to try and create synthetic GS promoters that gave a more effective level of selection stringency that would lead to increased mAb productivity. This alternative strategy involved the assembly of TFRE blocks identified to be providing the bulk of SV40 promoter transcriptional activity in CHO cells. Through this strategy, two 'synthetic SV40' promoters were created, named SynSV40_1 and SynSV40_2. These promoters were designed to function as synthetic, attenuated and minimal versions of the SV40 promoter.

SynSV40_1 and SynSV40_2 were tested alongside synthetic promoters to

control LC and HC transcription for their effect on stable mAb expression characteristics, including the rate of outgrowth of expressing cells and mAb productivity. SynSV40_1 and SynSV40_2 proximal promoters were cloned directly upstream of CMV core sequences and these promoters were used to control GS gene transcription in mAb 1 stable expression vectors. In some expression vectors mAb 1 LC and HC transcription was controlled by the LC1HC1 synthetic promoter combination, which was shown to give the highest mean mAb 1 titre in figure 4.6. In other expression vectors mAb 1 LC and HC transcription was controlled by a synthetic promoter known as 100RPU1, which was developed in a previous study and was shown to be very transcriptionally active (Brown et al., 2017). Conventional SV40 and hCMV promoters were also used to drive GS and LC/HC transcription respectively in some expression vectors to allow comparison with synthetic promoter alternatives. Nine different stable expression vectors were therefore used in this experiment. To generate stably expressing pools using these nine vectors, CHO host cells were transfected and $50\mu$M MSX was added to transfected cultures 24 hours later. The recovery of pools from transfection and MSX selection was monitored (figure 5.6).

The data in figure 5.6 show that all three SV40 promoter-containing pools recovered more quickly than all of the pools containing either SynSV40_1 or SynSV40_2 controlling GS transcription. This indicates that these synthetic promoters were successful in creating more stringent selective pressure on cell populations. However, the data do not show that SynSV40_2 conferred a greater selection stringency than SynSV40_1 despite being designed to be less transcriptionally active. Of the three cell pools transfected with the expression vector SynSV40_1 CMV only one pool fully recovered. Since the recovery rate of the other two pools was behind that of the negative control water transfected pool, these other two pools were not taken forward into further experiments.

Following recovery of transfected cell pools from MSX selection they then went into 11 day FBOG cultures to assess their mAb productivity (figure 5.7). The data in figure 5.7 show that there is an increase in mean qP when SynSV40_2 is used alongside either 100RPU1 or LC1HC1 compared to when the SV40 promoter is used alongside these LC/HC synthetic promoters (two-tailed Students t-test P values of 0.044 and 0.29 for the 100RPU1 and LC1HC1 comparisons respectively). Conversely, SynSV40_2 caused a reduction in IVCD when combined with 100RPU1 or LC1HC1 compared to when the SV40 promoter was combined with these LC/HC synthetic promoters.
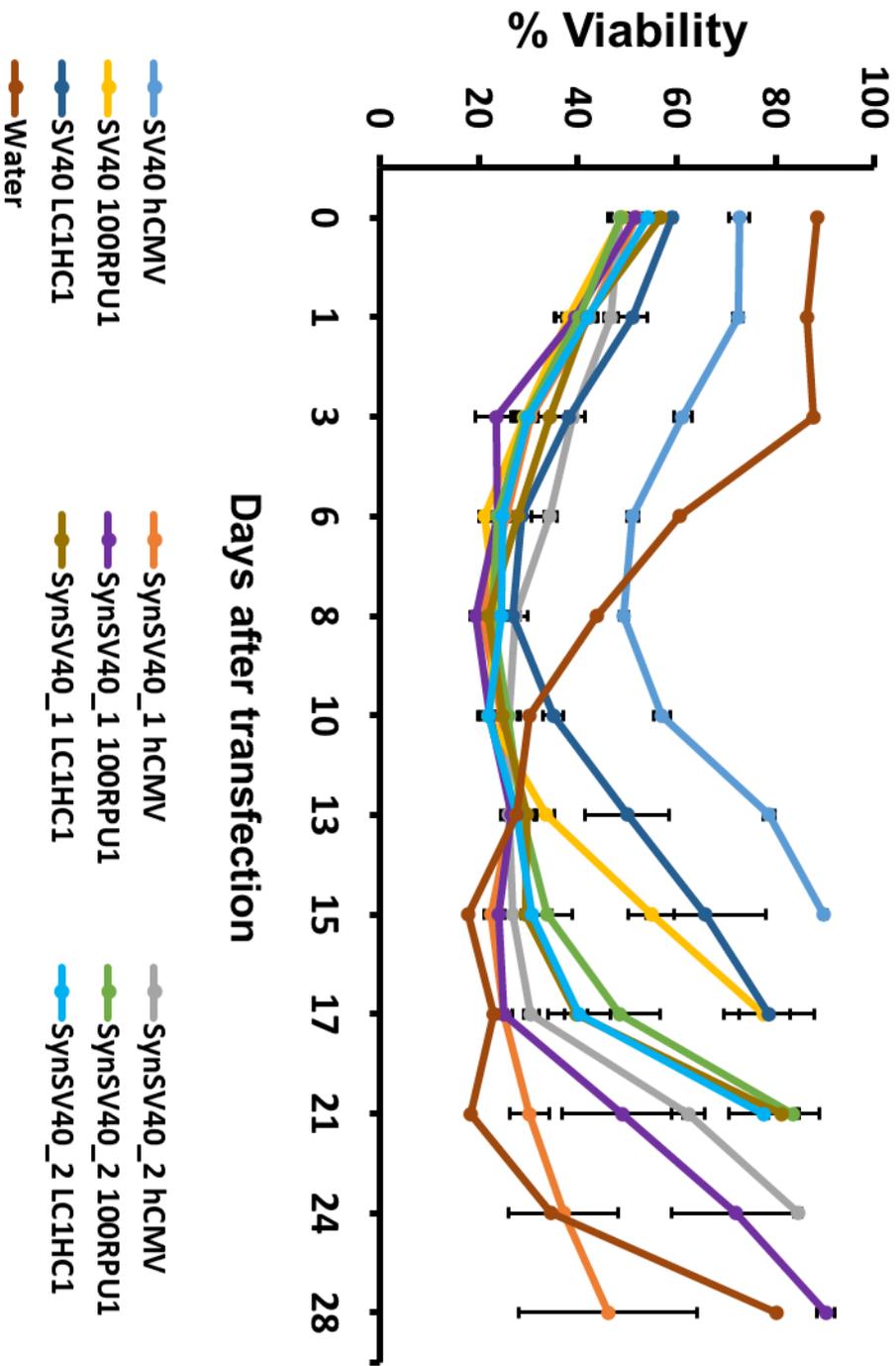
**Figure 5.6 – Effect of GS/LC/HC synthetic promoters on recovery rate of mAb 1-expressing pools**
CHO host cells were transfected with stable expression vectors containing either the SV40 promoter or one of two synthetic promoters to control GS transcription and either hCMV or synthetic promoter combinations to control LC and HC transcription. Vector-derived expression was selected for by the addition of $50\mu M$ MSX 24 hours post transfection. % viability was measured until full recovery of pools. Data points represent the mean of two (SynSV40_1 100RPU1) or three separately transfected pools per vector at each time point. Error bars represent standard error of the mean.
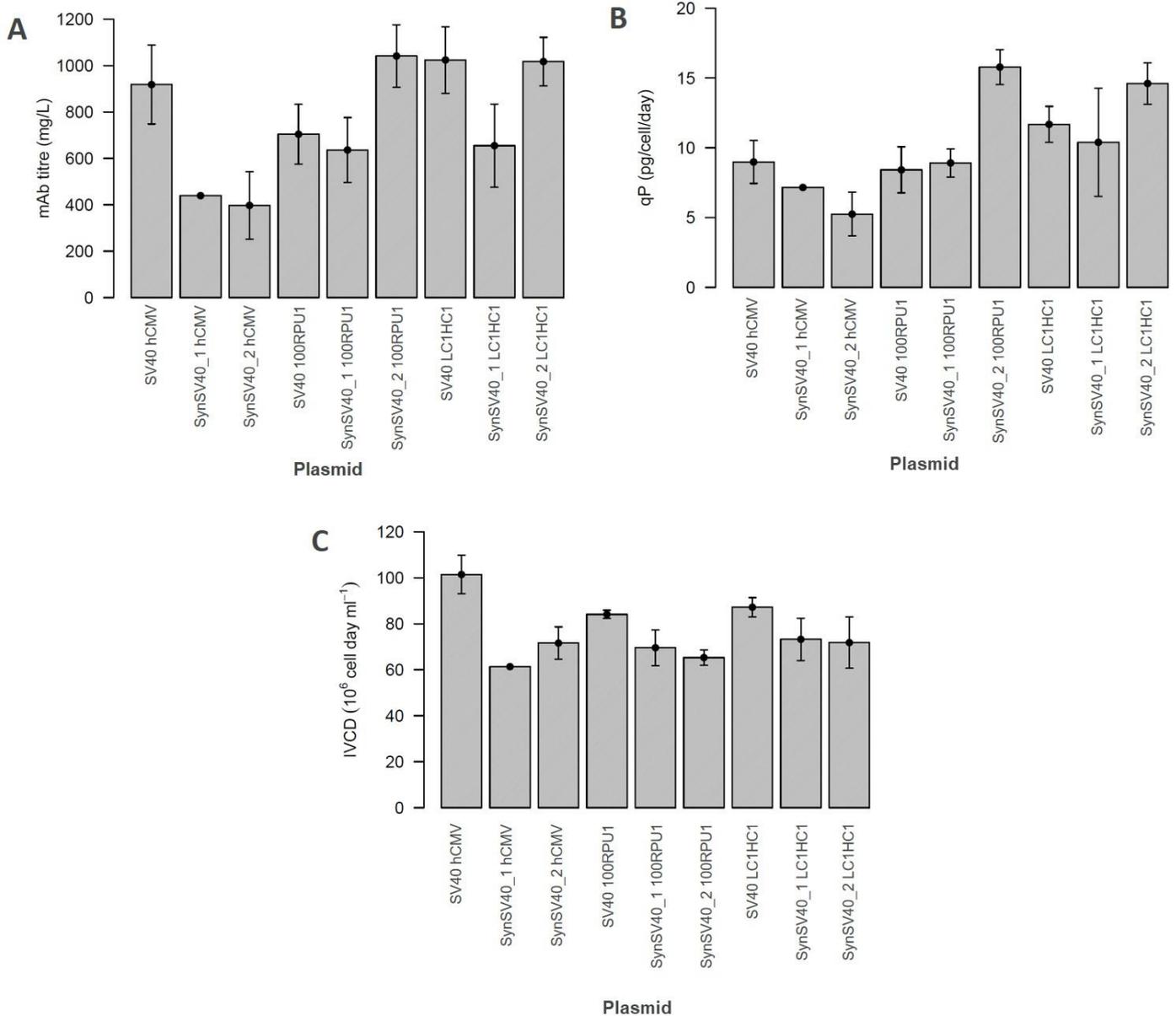
**Figure 5.7 – Effect of GS/LC/HC synthetic promoters on mAb productivity and IVCD of mAb 1-expressing pools**

CHO pools that had been transfected with mAb 1 expression vectors containing either SV40, SynSV40_1 or SynSV40_2 promoters to control GS transcription and either hCMV, 100RPU1 or LC1HC1 promoters to control LC and HC transcription were selected using MSX and then went in to 11 day fed batch overgrow cultures. **A)** shows the volumetric mAb titre generated from pools transfected with each vector; **(B)** shows the cell specific productivity (qP) of pools transfected with each vector and **(C)** shows the integral of viable cell density (IVCD) across the 11 day culture period for pools transfected with each vector. Bars represent the mean of two (SynSV40_1 100RPU1) or three separately transfected pools per vector, except for SynSV40_1 hCMV for which only one of the three pools recovered from MSX selection. Error bars represent standard error of the mean.

SynSV40_2 100RPU1 generated a 1.5-fold increase in mean volumetric titre compared to SV40 100RPU1 (two-tailed Students t-test P value of 0.21). SynSV40_2 LC1HC1 produced a very similar mean volumetric titre to SV40 LC1HC1. Producing more or a similar overall amount of mAb with a lower cell density is highly desirable due to the reduction in cell-derived impurities such as host cell protein that must be purified out in downstream processing steps.

Neither SynSV40_1 or SynSV40_2 had a positive effect on mAb 1 production when combined with hCMV, with both producing lower mean qP, lower mean IVCD and lower mean volumetric titre compared to when SV40 was combined with hCMV. Given that hCMV reduces SV40 promoter transcriptional output considerably due to transcription factor binding competition (as explained in detail in subsection 3.2.2) and that SynSV40_1/2 promoters are simply attenuated versions of SV40 that were constructed from the same key TFRE blocks, it is likely that hCMV further reduces transcriptional output from SynSV40_1/2 due to transcription factor binding competition. This may mean that when combined with hCMV for control of LC/HC transcription, SynSV40_1/2 are too transcriptionally inactive for transfected cell populations to adapt to MSX selection via high level vector-derived expression. Indeed, two of the three SynSV40_1 hCMV-transfected cell pools were not able to adapt to the MSX selection at all.

## 5.3 Discussion

This chapter describes the use of synthetic promoters to control transcription of the GS selection marker gene in stably expressing CHO cells. The results presented in this chapter have shown that increasing selection stringency on transfected cell populations can lead to increased qP at the cost of reduced IVCD. Crucially, this was shown to be possible via a synthetic GS promoter. When compared to the SV40 promoter, SynSV40_2 created more stringent selective pressure and this led to elevated qP and either equal or greater overall titre when combined with synthetic promoters for the control of LC and HC transcription. Even a relative increase in qP that leads to the same amount of mAb being produced from a smaller cell density is a hugely beneficial outcome due to a reduction in burden on downstream purification processes. Downstream processing is estimated to account for up to 80% of total mAb production cost (Goey et al., 2018). In particular, cell-derived impurities such as host cell protein can be challenging and costly to remove from CHO cell culture supernatant. Any

innovation that can reduce the amount of cell-derived impurities produced is therefore highly cost-effective. Finally, results within this chapter showed that if stringency of the GS/MSX selection system is increased too far then alternative, non-productive adaptations can be made by cell populations instead of adapting via high level vector-derived expression.

Recent studies in which stably expressing CHO cells have been adapted to different MSX concentrations are supportive of the results from similar experiments presented within this chapter. In one study, mAb-expressing CHO-K1 derived clones were selected at 0, 25 and 50µM MSX and those selected at 50µM showed the highest qP but the lowest growth rate (Noh et al., 2018). Another study adapted three cell lines each expressing different mAbs to MSX concentrations of 6.25µM, 25µM and, for one of the three cell lines, 75µM before seeding these cells into 14 day fed batch bioreactor cultures (Tian et al., 2020). The results showed that across all three cell lines adaptation to higher MSX concentrations led to greater qP. VCD was monitored throughout culture for one of the three cell lines and it was shown that adaptation to 25µM MSX created a slight reduction in IVCD compared with adaptation to 6.25µM MSX. Results presented in this chapter support this general trend whereby increasing selection stringency through culturing stably expressing cells in greater MSX concentrations causes an increase in qP but a reduction in IVCD and show that this trend continues at even greater MSX concentrations (50µM, 75µM, 100µM and 125µM).

Use of the initial round of six GS synthetic promoters in stable pool generation showed that they were successful in creating more stringent selective pressure but this did not lead to increased mAb productivity. It is likely that these promoters were too weak to enable the expected adaptation to occur - the upregulation of vector-derived GS expression to a threshold level that overcame the MSX inhibition. This might suggest that use of the TFRE 'building block' design strategy was on this occasion ineffective at these very low levels of transcriptional output. These promoters were constructed from TFRE blocks that had shown low level activity in homotypic constructs previously but had not been tested in heterotypic constructs (Johari et al., 2019). Although generally the activity level displayed by a TFRE in a homotypic construct is indicative of how it will perform in a heterotypic construct, occasionally this is not the case (Brown et al., 2017). One potential explanation for these exceptions is that some TFs form homodimers or multimers and these may bind to concatenated TFRE blocks within a homotypic construct but when that TFRE is flanked by different TFREs within a

heterotypic architecture this TF dimer or multimer may no longer bind. For example the TF STAT3, for which the corresponding TFRE is a key building block in these GS synthetic promoter designs, is known to form both homodimers and homotetramers and these tetramers have been shown to bind closely spaced sites within promoters (Hu et al., 2015; Sgrignani et al., 2018).

The finding that increasing stringency of the GS/MSX selection system too far can cause alternative, non-productive adaptations to be made has also been shown with the DHFR/methotrexate (MTX) selection system (Kim et al., 2001). Kim and colleagues found that when 23 clones were isolated from a stably expressing CHO pool and subjected in parallel to selection with a high concentration of MTX, only 1/23 clones experienced an increase in qP. The main alternative adaptation made here was shown to be an impairment in MTX cell surface transport while another study found that CHO cells can also adapt to produce DHFR with reduced MTX affinity when under MTX selective pressure (Flintoff & Essani, 1980). Results in this chapter show that the adaptation of some transfected cell pools to highly stringent MSX selection can not be explained by either high vector-derived expression or upregulation in expression level of the endogenous GS gene. Therefore it may be that similar alternative adaptations to those shown with the DHFR/MTX system could also be possible with the GS/MSX selection system. Investigation of these alternative adaptations would require molecular level interrogation.

# Chapter 6

# High-throughput screening of genes acting within the ER for their effect on mAb production

*In this chapter high-throughput transient transfections were performed to screen genes acting within the ER for their effect on mAb production. Genes mostly from the PDI and ERdj families as well as others that are functionally related such as Ero1, PRDX4, FICD and pERP1 were transfected into CHO cells stably expressing two different mAbs. Transfection of pERP1 was found to consistently increase the qP of cells producing both mAbs without decreasing growth rate. pERP1 therefore represents a novel target for CHO host cell engineering.*

## 6.1 Introduction

Disulphide bonds are a type of covalent bond that are integral to the structure of many proteins (Wiedemann et al., 2020). These disulphide bonds form between sulfhydryl (SH) groups of cysteine amino acid residues through oxidation reactions. IgG mAb molecules contain an internal disulphide bond within each domain of the antibody as well as a set of disulphide bonds connecting the two HCs (known as the hinge region) and a single disulphide bond joins the LC $C_L$ domain with the HC $CH_1$ domain (Liu & May, 2012). In mammalian cells, the process of disulphide bond formation is catalysed by the PDI family of enzymes within the ER (Ellgaard & Ruddock, 2005).

Chapter 4 of this manuscript was centred around the expression of IgG LC and HC polypeptides at different ratios in order to find the ratio at which mAb productivity was optimised. As explained in introductory subsection 4.1 of chapter 4, for some antibodies it is important to express LC in excess of HC to ensure frequent LC-HC binding events and release of BiP from the $CH_1$ domain of the HC, thus preventing UPR induction and ensuring mAb titres are high. An alternative approach to try and ensure frequent LC-HC binding events is to overexpress the ER machinery involved in the joining of LC and HC, i.e. PDI enzymes.

Overexpression of PDI family members in CHO has been shown previously to have beneficial effects on mAb productivity (Cartwright et al., 2020). However, PDI family members will vary in their substrate binding properties (Ellgaard & Ruddock,

2005). Different PDI family members have been demonstrated to be capable of interacting with IgG HC substrates (Jessop et al., 2009). Moreover, the level of enzymatic activity may vary between different PDI enzymes. For example PDI (P4HB) has two active sites, whereas PDIA4 (ERp72) and TXNDC5 (ERp46) have three and so may be expected to catalyse more disulphide bond formation per single copy of the enzyme (Ellgaard & Ruddock, 2005). Taken together, these findings justify a comprehensive comparison of a range of PDI family members for their effect on mAb production.
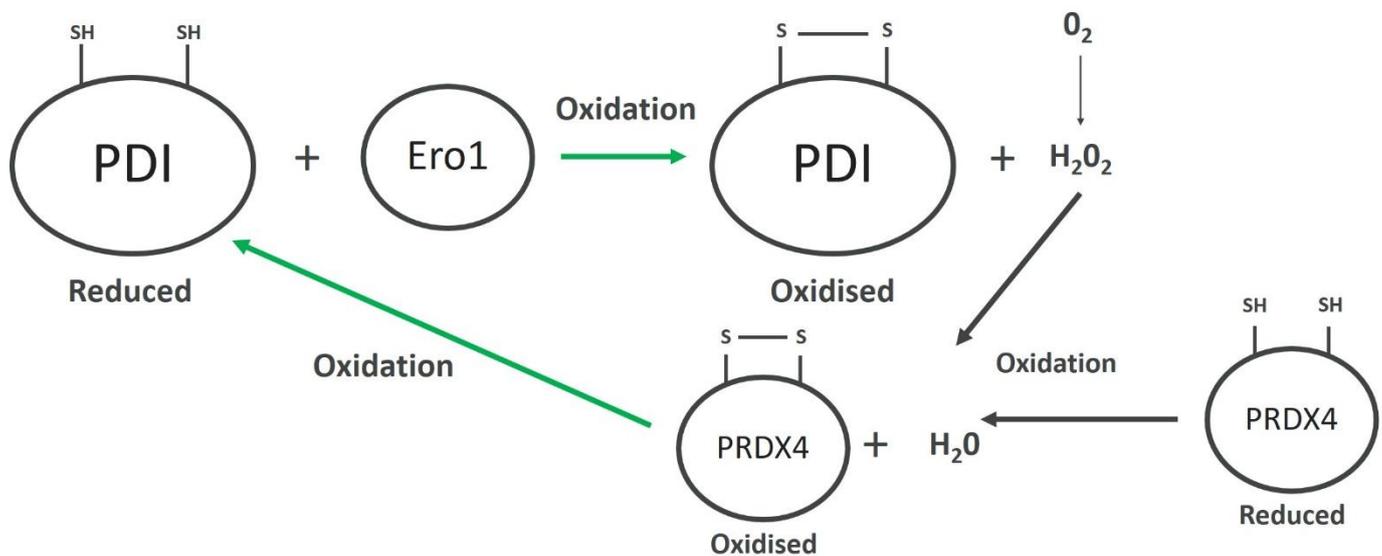


**Figure 6.1 – Redox regulation of PDI enzymes by Ero1 and PRDX4**

Ero1 will oxidise reduced PDI enzymes, transferring two electrons to molecular oxygen in the process and creating hydrogen peroxide ($H_2O_2$) as a by-product. Reduced PRDX4 is oxidised by the transfer of two electrons to hydrogen peroxide, generating water as a by-product, and this oxidised PRDX4 can then oxidise PDIs.

In order for PDI enzymes to catalyse disulphide bond formation, the CXXC active site motif of this PDI enzyme must itself be oxidised. The expression level of other ER-localised enzymes that control the redox state of PDIs (the ratio of oxidised and reduced active sites) is therefore an important consideration. Ero1 has long been known to oxidise PDI enzymes and this oxidation event transfers two electrons from reduced PDIs to molecular oxygen, creating hydrogen peroxide ($H_2O_2$) as a by-product (Appenzeller-Herzog & Ellgaard, 2008). There is evidence for Ero1-mediated oxidation

of PDI family members PDIA3 (ERp57), PDI, PDIA6 (P5) and TXNDC5 (Jessop et al., 2009). In addition to Ero1, another enzyme known as PRDX4 was later discovered to be capable of oxidising PDIs (Zito et al., 2010). Reduced PRDX4 is oxidised by the transfer of two electrons to hydrogen peroxide, generating water as a by-product, and this oxidised PRDX4 can then oxidise PDIs. In a cell that is expressing both Ero1 and PRDX4 at sufficient levels, there is therefore a positive feedback loop created whereby PDIs can be oxidised twice for every molecular oxygen molecule that is eventually converted to water via hydrogen peroxide. This positive feedback loop is illustrated in figure 6.1. PDI genes were therefore screened alongside Ero1, PRDX4 and both in this chapter.

The gene pERP1 (MZB1) is hugely upregulated in expression during differentiation of B cells into antibody-producing plasma cells, more so than all of the PDI variants and key ER-localised folding chaperones such as BiP (van Anken et al., 2009). pERP1 is not expressed in CHO cells and is limited only to lymphocyte expression. Interestingly, pERP1 expression was shown to trigger formation of the internal disulphide bond within the $CH_1$ domain of an IgG HC without LC expression in the mouse plasmacytoma cell line Ag8 (Shimizu et al., 2009). Normally, LC expression is required to release BiP binding from the $CH_1$ domain so that this internal disulphide bond can form (Vanhove et al., 2001). This led to the theory that perhaps pERP1 is a novel chaperone that functions to help trigger BiP release from Ig domains and thus enable folding and internal disulphide bond formation to occur. Knockout of pERP1 expression in mouse models caused impaired antibody secretion from plasma cells (Rosenbaum et al., 2014). pERP1 is not a PDI family member and was recently identified to belong to the CNPY group of proteins based on sequence/structural similarity (Schildknegt et al., 2019). Nevertheless, pERP1 was screened in this chapter alongside PDI proteins due to its perceived function being broadly similar, i.e. involvement in ER-localised events that cause BiP to be released from HCs.

The ERdj family of genes were also screened within this chapter for their effect on mAb production. The general function of this family of genes is to regulate the various activities of BiP within the ER. Overexpression of ERdj genes in CHO cells has been shown previously to increase mAb titres (Swindley, 2021). Different ERdj genes regulate BiP in different ways. For instance ERdj1 and ERdj2 are integral ER membrane proteins that, through interactions with both ribosomes and BiP, serve to regulate translation and translocation in a BiP-dependent manner, ensuring BiP is

available to engage nascent polypeptide chains as they enter the ER. Other ERdj family members, on the other hand, interact directly with unfolded polypeptides within the ER lumen through their substrate binding domains and transfer these substrates to BiP (Pobre et al., 2019). In addition to ERdj family members, the enzyme FICD is also capable of modulating BiP activity through post translational modification (PTM). Under conditions of low demand, FICD AMPylates BiP. AMPylation is a PTM that maintains BiP in its ATP-bound form, which has low affinity for substrates. FICD is also reported to be capable of de-AMPylating BiP when demand for BiP increases, i.e. when there is a lot of unfolded polypeptide within the ER (Preissler et al., 2017). The de-AMPylation of BiP enables ATP hydrolysis to occur, converting BiP to its ADP-bound form which has high affinity for substrates. FICD was therefore screened alongside ERdj genes for its effect on mAb production.

In this chapter, PDI family genes and pERP1 were screened by high-throughput transient transfection both individually and in combination with Ero1, PRDX4 or both. These effector genes/effector gene combinations were transfected into a CHO clonal cell line stably expressing a mAb referred to as mAb 3 as well as a CHO cell pool stably expressing mAb 1. ERdj family genes and FICD were also screened by high-throughput transient transfection into the mAb 3-expressing CHO clonal cell line. Results showed pERP1 to consistently and robustly increase the qP of cells producing both mAbs without reducing their growth rate.

## 6.2 Results

Human protein coding versions of the genes listed in table 6.1 were screened by transient expression within this chapter for their effect on mAb production. These human protein coding sequences were taken from the CCDS project (https://www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi). The PDI family member PDIA3 was set to be included within the PDI gene screens but had to be discarded due to consistently inadequate plasmid yields from *E.coli* culture. This may suggest that the gene sequence was in some way toxic to the *E.coli*.

The 96 well Nucleocuvette plate form of Nucleofection (Lonza) was used throughout this chapter as the method of transient transfection. Importantly, this high-throughput method was shown in figure 4.1 (described in subsection 4.2.1.1), through transient transfection of a GFP plasmid into a CHO host cell line, to generate high transfection efficiency within 24 hours that was very consistent across an entire plate.

Nunc 96 deep well plates (ThermoFisher) were also used throughout this chapter for culture of transfected cells. It was shown in figure 4.2 (described in subsection 4.2.1.2) that cell growth within these plates following transfection by the 96 well Nucleocuvette plate method was highly consistent from well-to-well. Taken together, these results ensured that when performing high-throughput gene screen experiments there would be no plate location-specific biases introduced, for example genes being transfected more efficiently or transfected cells dividing more easily in the middle of plate rows than the edges or vice versa.

An "empty" plasmid vector that did not code for a transgene was used throughout this chapter as a control to which the effect of transgene expression could be compared. To create this empty plasmid vector, an effector gene plasmid was digested with a single restriction enzyme that had recognition sites flanking the effector gene expression unit and the resulting backbone fragment was then self-ligated. The sequence of the empty vector therefore matches the backbone sequence of the effector gene vectors.

**Table 6.1 - Effector genes screened by transient transfection within this chapter**

| Gene name | Alternative name |
| --- | --- |
| **PDI genes** | |
| PDI | P4HB |
| PDIA2 | PDIp |
| TXNDC5 | ERp46 |
| pERP1 | MZB1 |
| Ero1α | |
| PDIA4 | ERp72 |
| PRDX4 | |
| PDIA6 | P5 |
| **ERDJ genes** | |
| ERdj1 | DNAJC1 |
| ERdj2 | DNAJC23 |
| ERdj3 | DNAJB11 |
| ERdj4 | DNAJB9 |
| ERdj5 | DNAJC10 |
| ERdj6 | DNAJC3 |
| ERdj7 | |
| FICD | |

### 6.2.1 Effect of transient expression of PDI family and related genes on mAb 3 production

Genes from the PDI family, as well as related genes acting within the ER such as pERP1, Ero1α (referred to as Ero1 herein) and PRDX4 were screened for their effect on production of a mAb molecule referred to herein as mAb 3. To perform this screen, effector genes were transiently transfected via Nucleofection into a CHO cell clone stably expressing mAb 3. Transfected cell populations were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. Figure 6.2 shows the effect of these genes on mAb productivity and cell division.

The data in figure 6.2 show that none of the eight effector genes increased volumetric mAb titre compared to transfection of the empty vector control. These differences in volumetric mAb titre are largely controlled by differences in IVCD of the different transfectants, as can be seen from bar chart B. Differences in IVCD are in turn largely controlled by differences in the day 0 VCD (bar chart C). These day 0 VCD measurements were taken immediately post transfection. Differences in day 0 % viability are similar to these differences in day 0 VCD, with variation among triplicate sets being similarly as tight (data not shown). This indicates that the different effector genes had varying levels of cell toxicity upon transfection. Transfection of the empty vector and PDI had the least toxic effect and these transfectants produced the highest mean volumetric titres while transfection of pERP1 was the most toxic and pERP1 transfectants had the lowest mean volumetric titre.

In order to assess growth rate while normalising for these variations in starting VCD, the effect of these genes on doubling time of cells was calculated (bar chart D). Mean doubling times were relatively consistent across cells transfected with the different effector genes. pERP1 transfected cells had the lowest mean doubling time although this was not statistically significantly lower than that of the empty vector control (two-tailed Students t-test P value = 0.14).

The effect of differences in IVCD (which are explained largely by differences in day 0 VCD) on mAb titre can be normalised for by calculation of qP (titre/IVCD) (bar chart E). Six of the eight effector genes tested generated a significant improvement in qP compared to the empty vector control, with PDI and PDIA6 being the exceptions.

Interestingly, pERP1 generated the lowest mean doubling time and the highest mean qP of all transfectants, albeit it appeared to have the most toxic effect upon
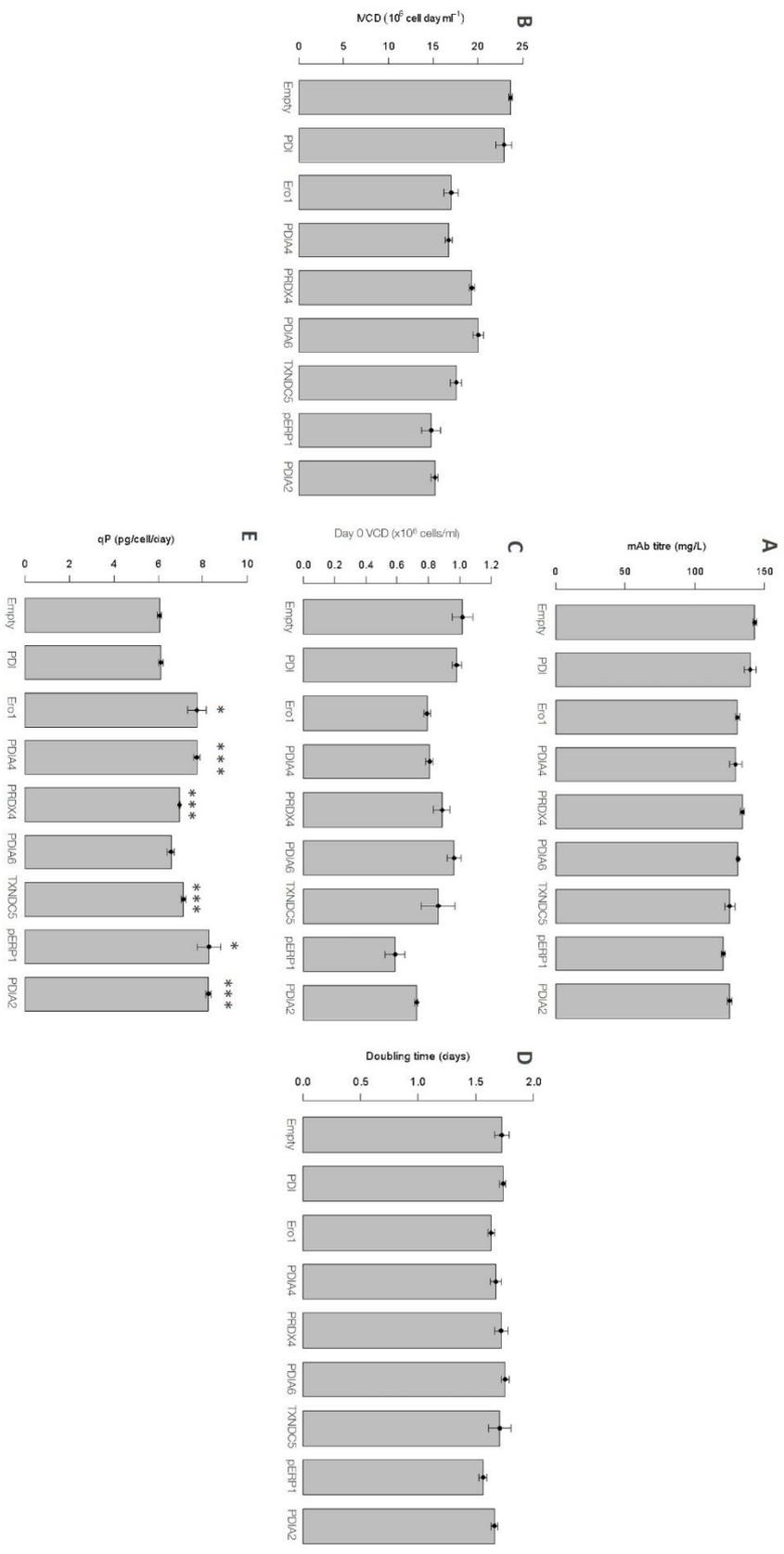
**Figure 6.2 - Effect of transient expression of various PDI effector genes on mAb 3 production**

Various effector genes as well as an "empty" plasmid vector were transiently transfected via Nucleofection into a CHO cell line stably expressing mAb 3. Transfected cells were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. **A)** shows the volumetric mAb titre generated from cells transfected with different effector genes; **(B)** shows the integral of viable cell density (IVCD) across the 5 day culture period for cells transfected with different effector genes; **(C)** shows the day 0 viable cell density (VCD) of cells transfected with different effector genes; **(D)** shows the doubling time of cells transfected with different effector genes and **(E)** shows the cell specific productivity (qP) of cells transfected with different effector genes. Bars represent the mean of three separate transfections per effector gene. Error bars represent standard error of the mean. Mean values significantly greater (two-tailed Students t-test) than the mean empty vector value are indicated by asterisks (* P<0.05, **P<0.025, ***P<0.01).

transfection. When developing a new host cell line that stably overexpresses an effector gene, any reduction in VCD upon transfection of the effector gene is unlikely to be a major problem as VCD would be reduced anyway as part of the selection process. PDIA2 and Ero1 make up the three best performers after pERP1 in terms of both mean doubling time and mean qP.

## 6.2.2 Effect of transient expression of PDI family and related gene combinations on mAb 3 production

Genes from the PDI family, as well as related genes acting within the ER such as pERP1 were screened in combination with Ero1, PRDX4 or both for their effect on production of mAb 3. To provide controls to which transfection of effector genes plus Ero1/PRDX4/both could be compared, effector genes were also co-transfected alongside the empty plasmid vector. In these effector gene plus empty co-transfection controls, the effector gene plasmid load matched that used in the corresponding effector gene plus Ero1/PRDX4/both transfections ($400ng/2x10^6$ cells in the effector gene plus Ero1/PRDX4 transfections and $267ng/2x10^6$ cells in the effector gene plus both Ero1 and PRDX4 transfections). This set up is illustrated at the top of figure 6.3.

These effector gene combinations were transiently transfected via Nucleofection into a CHO cell clone stably expressing mAb 3. Transfected cell populations were cultured for 5 days before mAb titre analysis and VCD measurements were taken on days 0, 3 and 5 of culture. Figure 6.3 shows the effect of these gene combinations on mAb productivity and cell division.

The data in figure 6.3 show that differences in volumetric mAb titre generated by the transfectants were predominantly explained by differences in IVCD. As with the single gene conditions shown in figure 6.2, differences in IVCD were largely explained by differences in day 0 VCD (data not shown), with some gene combinations causing greater cell toxicity upon transfection than others. An untransfected control had the highest mean % viability of all conditions on day 0 (data not shown) and this lack of transfection-induced toxicity is the primary reason for the untransfected control producing the highest mean IVCD. To assess growth rates of transfectants while normalising for differences in day 0 VCD, cell doubling time was calculated. These results show that co-expression of Ero1, PRDX4 or both alongside any of the effector genes did not cause a significant reduction in mean cell doubling time compared to effector gene expression alone.

**A**

| | | 33.3% | 33.3% | 33.3% |
|---|---|---|---|---|
| Effector gene | | | Ero1 | PRDX4 |

| | 50% | 50% |
|---|---|---|
| Effector gene | | Ero1/PRDX4/Empty |

| | 33.3% | 66.6% |
|---|---|---|
| Effector gene | | Empty |

**B**

**C**

**D**

**E**
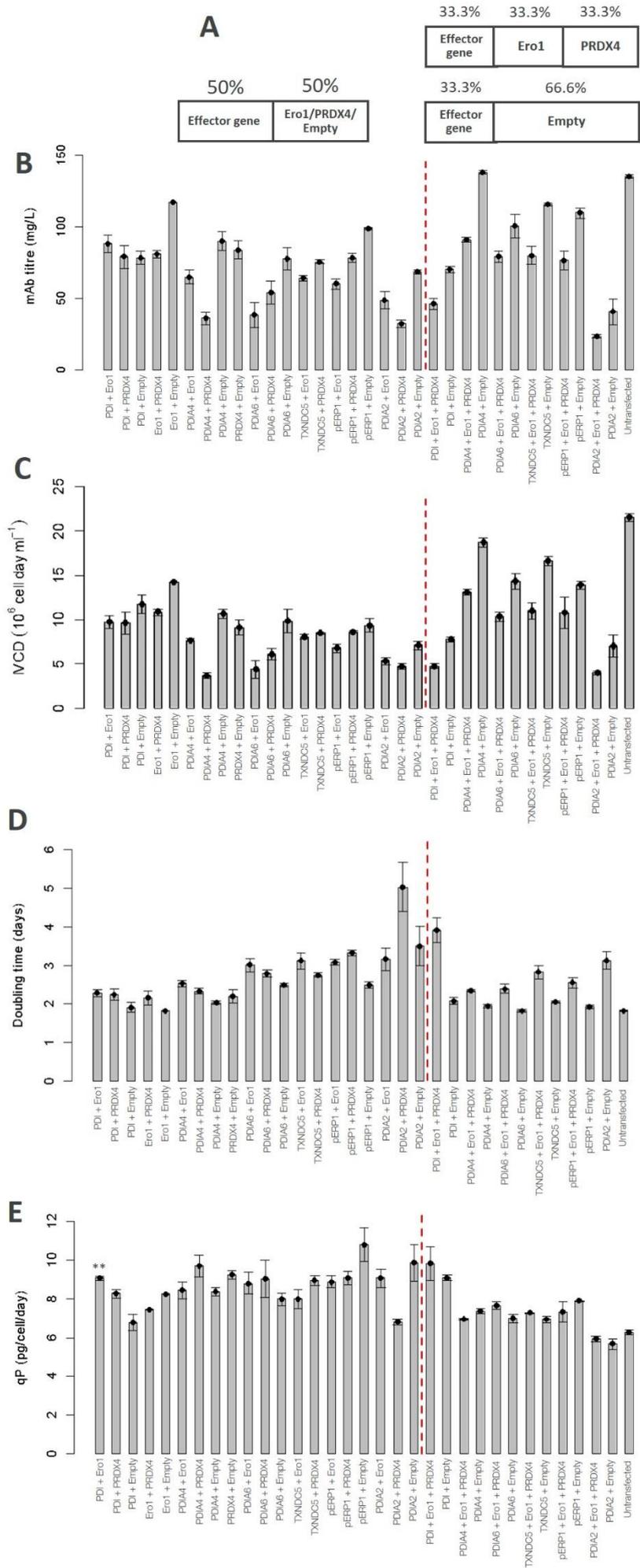
**Figure 6.3 - Effect of transient expression of various PDI effector gene combinations on mAb 3 production**

Various effector gene combinations were transiently transfected via Nucleofection into a CHO cell line stably expressing mAb 3. Transfected cells were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. **A)** shows the proportion of overall transfected DNA load taken up by effector genes or empty, non protein-coding plasmids in the double and triple gene transfections. **B)** shows the volumetric mAb titre generated from cells transfected with different effector gene combinations; **(C)** shows the integral of viable cell density (IVCD) across the 5 day culture period for cells transfected with different effector gene combinations; **(D)** shows the doubling time of cells transfected with different effector gene combinations and **(E)** shows the cell specific productivity (qP) of cells transfected with different effector gene combinations. Bars represent the mean of one (Ero1 + Empty), two (Ero1 + PRDX4; TXNDC5 + PRDX4 and PDIA2 + Empty within the triple gene transfectants) or three separate transfections per effector gene combination. Error bars represent standard error of the mean. Mean values significantly greater (two-tailed Students t-test) than the mean value for the corresponding "effector gene + empty" control are indicated by asterisks (* $P<0.05$, **$P<0.025$, ***$P<0.01$).

Calculating qP of transfectants revealed that co-expression of Ero1 alongside PDI caused a significant increase in mean qP compared to the "PDI + Empty" transfections at the same PDI gene load. PRDX4 co-expression alongside PDI also caused an increase in mean qP, albeit this difference to "PDI + Empty" was not quite statistically significant (two-tailed Students t-test P value = 0.07). Ero1/PRDX4/both did not generate a significant increase in mean qP when co-expressed alongside any of the other effector genes. Comparing each condition to the untransfected control by two-tailed Students t-test shows that all conditions generated a significant increase in mean qP except from "PDI + Empty" within the 50:50 co-transfections, "pERP1 + Ero1 + PRDX4", "PDIA2 + Ero1 + PRDX4" and "PDIA2 + Empty" within the triple gene transfections. Among the 50:50 co-transfections, the condition displaying the highest mean qP was "pERP1 + Empty". This is comparable to the results of the single gene transfections in figure 6.2, in which pERP1 produced the highest mean qP. Among the triple gene transfection conditions, "pERP1 + Empty" produced the third highest mean

qP. Table 6.2 shows the three highest performing effector genes/effector gene combinations in terms of both mean cell doubling time and mean qP for the single gene, double gene and triple gene mAb 3 transfections.

**Table 6.2 – The 3 highest performing effector genes/effector gene combinations in terms of both mean cell doubling time and mean qP for single gene, double gene and triple gene mAb 3 transfections**

| | Single gene transfections | | Double gene transfections | | Triple gene transfections | |
|---|---|---|---|---|---|---|
| | Mean cell doubling time | Mean qP | Mean cell doubling time | Mean qP | Mean cell doubling time | Mean qP |
| **1** | pERP1 | pERP1 | Ero1 + Empty | pERP1+ Empty | PDIA6 + Empty | PDI + Ero1 + PRDX4 |
| **2** | Ero1 | PDIA2 | PDI + Empty | PDIA2 + Empty | pERP1 + Empty | PDI + Empty |
| **3** | PDIA2 | Ero1 | PDIA4 + Empty | PDIA4 + PRDX4 | PDIA4 + Empty | pERP1 + Empty |

### 6.2.3 Effect of transient expression of PDI family and related genes on mAb 1 production

In order to assess whether the effect on mAb production of genes from the PDI family, as well as related genes acting within the ER such as pERP1, Ero1 and PRDX4 was likely to be generic or mAb-specific, these genes were screened for their effect on production of a second mAb molecule – mAb 1. To perform this screen, effector genes were transiently transfected via Nucleofection into a CHO cell pool stably expressing mAb 1. Transfected cell populations were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. Figure 6.4 shows the effect of these genes on mAb productivity and cell division.

Figure 6.4 shows that none of these effector genes increased mean volumetric mAb 1 titre when transfected compared to transfection of the empty plasmid. These differences in volumetric titre were predominantly controlled by differences in IVCD. As with transfection of these genes into mAb 3-expressing cells, these differences in IVCD were mostly explained by differences in day 0 VCD caused by varying levels of cell toxicity upon transfection (data not shown). Calculating cell doubling time of transfectants revealed that none of the effector genes generated a significant reduction in mean doubling time compared to the empty vector control. The effect of some of these genes on doubling time/IVCD was considerably variable between replicates, as
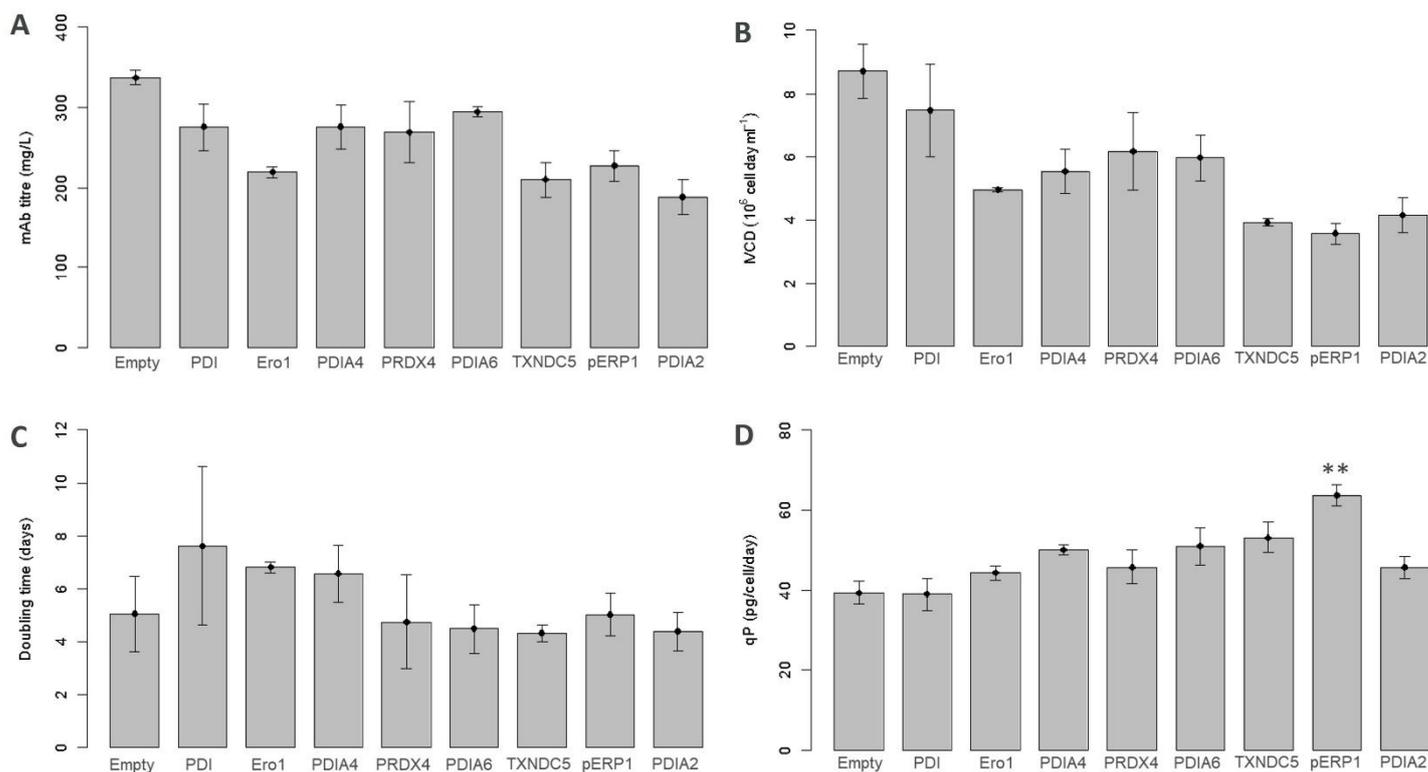
**Figure 6.4 - Effect of transient expression of various PDI effector genes on mAb 1 production**

Various effector genes as well as an "empty" plasmid vector were transiently transfected via Nucleofection into a CHO cell pool stably expressing mAb 1. Transfected cells were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. **A)** shows the volumetric mAb titre generated from cells transfected with different effector genes; **(B)** shows the integral of viable cell density (IVCD) across the 5 day culture period for cells transfected with different effector genes, **(C)** shows the doubling time of cells transfected with different effector genes and **(D)** shows the cell specific productivity (qP) of cells transfected with different effector genes. Bars represent the mean of two (Empty and PDIA4) or three separate transfections per effector gene. Error bars represent standard error of the mean. Mean values significantly greater (two-tailed Students t-test) than the mean empty vector value are indicated by asterisks (* $P<0.05$, **$P<0.025$, ***$P<0.01$).

indicated by the large error bars. This is in contrast to the effect of these genes on doubling time/IVCD of mAb 3-expressing cells (figure 6.2), which was very consistent among replicates. This difference may be due to mAb 3 expression coming from a monoclonally-derived cell line whereas mAb 1 expression was from a heterogenous cell pool, meaning there was more variability in the cell background.

All effector genes except from PDI generated an increase in mean qP compared to the empty vector control. pERP1 generated the highest mean qP and this was the only

condition for which the difference to the empty vector control was statistically significant. These results are comparable to the effect of these genes on mAb 3 qP (figure 6.2), in which PDI was the only gene not to generate a substantial increase compared to the empty vector control and pERP1 caused the highest mean qP.

## 6.2.4 Effect of transient expression of PDI family and related gene combinations on mAb 1 production

The above PDI and similar genes were screened in combination with Ero1, PRDX4 or both for their effect on production of mAb 1 through transient transfection via Nucleofection into a CHO cell pool stably expressing mAb 1. As with transfection of these effector gene combinations into mAb 3-expressing cells, effector genes were also co-transfected alongside the empty plasmid vector to provide controls to which effector genes plus Ero1/PRDX4/both could be compared. In these effector gene plus empty co-transfection controls, the effector gene plasmid load matched that used in the corresponding effector gene plus Ero1/PRDX4/both transfections ($400ng/2x10^6$ cells in the effector gene plus Ero1/PRDX4 transfections and $267ng/2x10^6$ cells in the effector gene plus both Ero1 and PRDX4 transfections). This set up is illustrated at the top of figure 6.5. Transfected cell populations were cultured for 5 days before mAb titre analysis and VCD measurements were taken on days 0, 3 and 5 of culture. Figure 6.5 shows the effect of these gene combinations on mAb productivity and cell division.

The data in figure 6.5 show that once again the different mean volumetric mAb 1 titres are explained primarily by differences in mean IVCD from the conditions. Differences in mean IVCD are again predominantly determined by differences in mean day 0 VCD due to differing levels of cell toxicity caused by transfection of the effector genes (data not shown). Calculating doubling time of transfectants revealed that co-expression of either Ero1 or PRDX4 alongside PDI caused a notable reduction in mean cell doubling time compared to PDI expression alone, although these reductions wernot statistically significant (two-tailed Students t-test P values of 0.09 and 0.22 for "PDI + Ero1" and "PDI + PRDX4" respectively). Co-expression of Ero1, PRDX4 or both alongside any of the effector genes did not significantly reduce the mean cell doubling time compared to effector gene expression alone.

The effector gene combinations "PDIA6 + Ero1", "PDIA2 + Ero1" and "TXNDC5 + Ero1 + PRDX4" all generated notable increases in mean qP compared to their respective "Effector gene + Empty" co-transfection control, albeit these increases
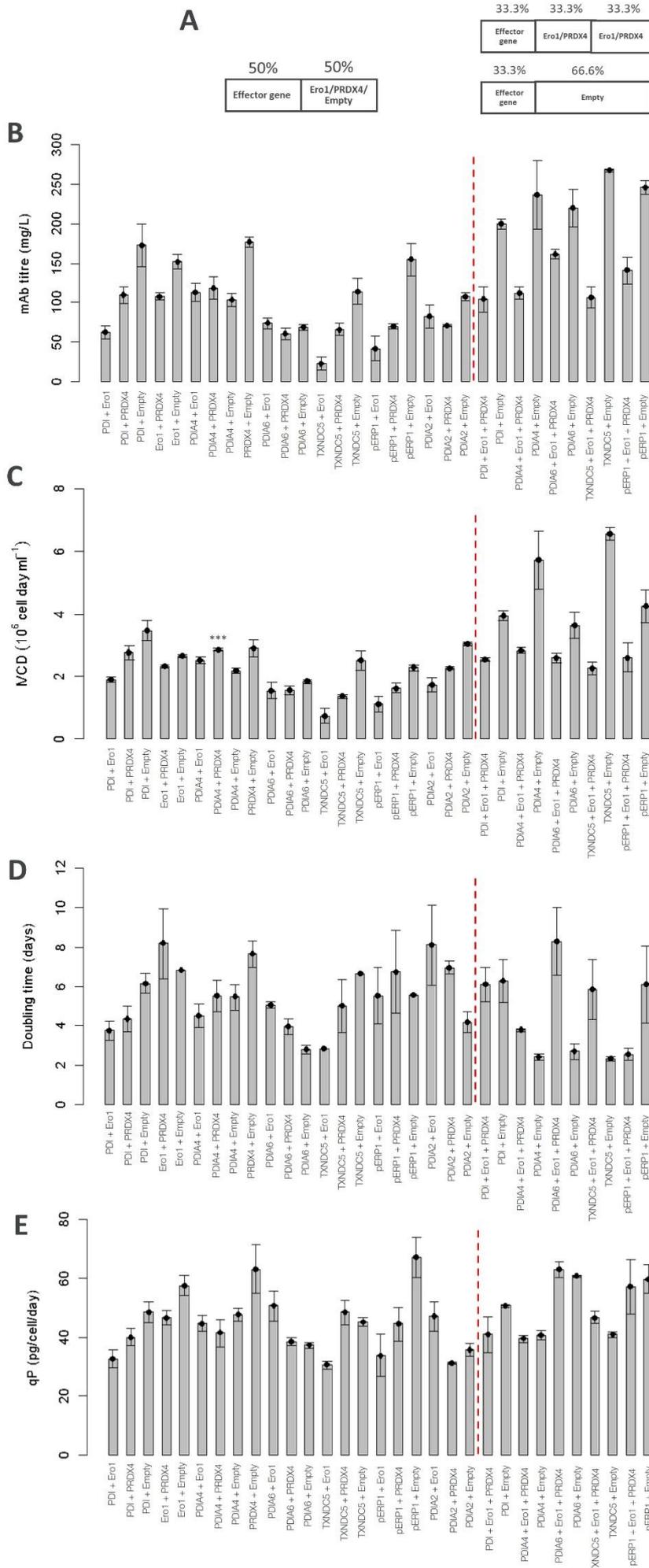
**Figure 6.5 - Effect of transient expression of various PDI effector gene combinations on mAb 1 production**

Various effector gene combinations were transiently transfected via Nucleofection into a CHO cell pool stably expressing mAb 1. Transfected cells were cultured for 5 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 5 of culture. **A**) shows the proportion of overall transfected DNA load taken up by effector genes or empty, non protein-coding plasmids in the double and triple gene transfections. **B**) shows the volumetric mAb titre generated from cells transfected with different effector gene combinations; **(C)** shows the integral of viable cell density (IVCD) across the 5 day culture period for cells transfected with different effector gene combinations; **(D)** shows the doubling time of cells transfected with different effector gene combinations and **(E)** shows the cell specific productivity (qP) of cells transfected with different effector gene combinations. Bars represent the mean of two (PDIA2 + PRDX4; PDI + Empty within the triple gene transfectants; PDIA4 + Ero1 + PRDX4 and PDIA6 + Empty within the triple gene transfectants) or three separate transfections per effector gene combination. Error bars represent standard error of the mean. Mean values significantly greater (two-tailed Students t-test) than the mean value for the corresponding "effector gene + empty" control are indicated by asterisks (* P<0.05, **P<0.025, ***P<0.01).

**Table 6.3 – The 3 highest performing effector genes/effector gene combinations in terms of both mean cell doubling time and mean qP for single gene, double gene and triple gene mAb 1 transfections**

| | Single gene transfections | | Double gene transfections | | Triple gene transfections | |
|---|---|---|---|---|---|---|
| | Mean cell doubling time | Mean qP | Mean cell doubling time | Mean qP | Mean cell doubling time | Mean qP |
| **1** | TXNDC5 | pERP1 | PDIA6 + Empty | pERP1 + Empty | TXNDC5 + Empty | PDIA6 + Ero1 + PRDX4 |
| **2** | PDIA2 | TXNDC5 | TXNDC5 + Ero1 | PRDX4 + Empty | PDIA4 + Empty | PDIA6 + Empty |
| **3** | PDIA6 | PDIA6 | PDI + Ero1 | Ero1 + Empty | pERP1 + Ero1 + PRDX4 | pERP1 + Empty |

were not quite statistically significant (two-tailed Students t-test P values of 0.11, 0.16 and 0.11 respectively). Among the 50:50 co-transfection conditions, "pERP1 + Empty" produced the highest mean qP, just as it did in the mAb 3 50:50 co-transfections (figure 6.3). Among the triple gene transfection conditions, "pERP1 + Empty" produced the third highest mean qP. Table 6.3 shows the three highest performing effector genes/effector gene combinations in terms of both mean cell doubling time and mean qP for the single gene, double gene and triple gene mAb 1 transfections.

### 6.2.5 Effect of transient expression of ERdj family genes on mAb 3 production

Genes from the ERdj family as well as FICD were screened for their effect on production of mAb 3. To perform this screen, effector genes were transiently transfected via Nucleofection into a CHO cell clone stably expressing mAb 3. These effector genes were screened at 3 different transfected DNA loads: the 'full' DNA load ($800ng/2x10^6$ cells), a 'half' DNA load ($400ng/2x10^6$ cells) and a $1/3^{rd}$ DNA load ($266.6ng/2x10^6$ cells). In the half and the $1/3^{rd}$ effector gene DNA load conditions, the empty vector was also co-transfected to take the overall DNA load to $800ng/2x10^6$ cells. Transfected cell populations were cultured for 4 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 4 of culture. Figure 6.6 shows the effect of these genes on mAb productivity and cell division.

The data in figure 6.6 show that none of the effector gene transfections generated an increase in mean volumetric mAb 3 titre compared to the empty vector control. These ERdj effector genes caused less drastic and more consistent cell toxicity upon transfection compared to transfection of the PDI family and related effector genes (as judged by comparison of day 0 VCD values, data not shown). Nevertheless, cell doubling time was calculated to assess growth rate while normalising for this relatively small amount of variation in day 0 VCD. None of the effector gene conditions generated a reduction in mean cell doubling time compared to the empty vector control. ERdj7 and "FICD + Empty" within the half DNA load conditions had the two highest mean cell doubling times of all conditions and this explains why they also had the two lowest mean IVCD values of all conditions.

Comparing qP generated by the different conditions reveals that "ERdj4 + Empty" within the $1/3^{rd}$ DNA load transfections produced the highest mean qP. However this mean qP was not statistically significantly greater than that of the empty vector control (two-tailed Students t-test P value = 0.17). None of the conditions produced a
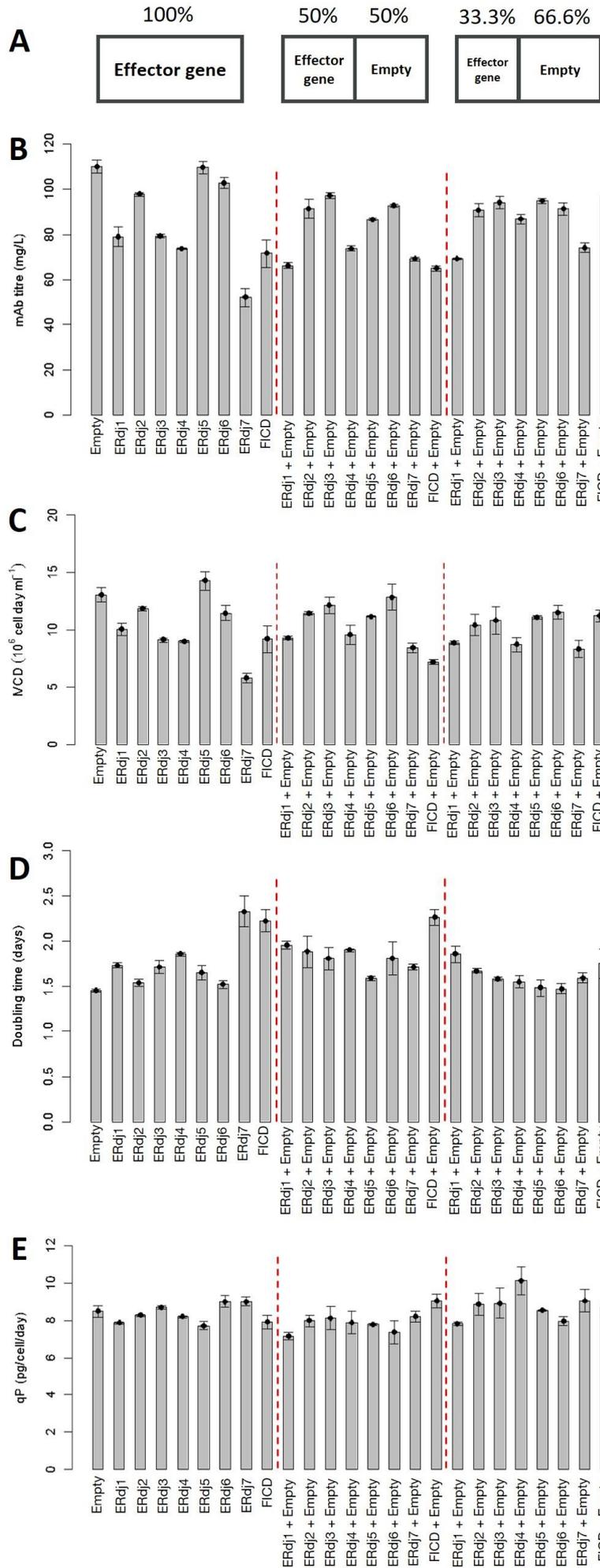
**Figure 6.6 - Effect of transient expression of various ERdj effector genes on mAb 3 production**

Various effector genes were transiently transfected at different DNA loads via Nucleofection into a CHO cell line stably expressing mAb 3. Transfected cells were cultured for 4 days before supernatant was harvested for mAb titre analysis. Viable cell density measurements were taken on days 0, 3 and 4 of culture. **A)** shows the proportion of overall transfected DNA load taken up by effector genes or empty, non protein-coding plasmid. **B)** shows the volumetric mAb titre generated from cells transfected with different effector genes; **(C)** shows the integral of viable cell density (IVCD) across the 4 day culture period for cells transfected with different effector genes; **(D)** shows the doubling time of cells transfected with different effector genes and **(E)** shows the cell specific productivity (qP) of cells transfected with different effector genes. Bars represent the mean of three separate transfections per condition. Error bars represent standard error of the mean.

significant increase in mean qP compared to the empty vector control. "ERdj5 + Empty" at the 1/3$^{rd}$ DNA load produced a significantly higher mean qP compared to ERdj5 transfected at the full DNA load (two-tailed Students t-test P value = 0.04). This may indicate that this protein activates different signalling pathways or contributes to different processes within the ER depending on its abundance.

## 6.3 Discussion

In this chapter high-throughput gene screening revealed pERP1 expression to be a novel target for CHO host cell engineering. pERP1 was shown robustly and consistently to increase the qP of cells producing two different mAbs without inhibiting their growth rate. Co-expression of Ero1 alongside PDI caused a significant increase in mAb 3 qP compared to PDI expression alone. Apart from this however, co-expression of Ero1, PRDX4 or both alongside PDI effector genes did not yield significant improvements in either mean cell doubling time or mean qP. Likewise, high-throughput screening of ERdj effector genes did not reveal any of them to significantly improve either mean cell doubling time or mean qP.

The beneficial effect of some PDI enzymes on mAb production may be dependent upon co-expression of partner proteins that were not included in this screen. For example, PDIA4 in combination with cyclophilin B increased the *in vitro* rate of

assembly of a $CH_1$-$C_L$ heterodimer compared to PDIA4 or cyclophilin B alone (Jansen et al., 2012). In another study, co-overexpression of PDIA3 alongside ERp27 significantly increased titre of an IgG mAb whereas overexpression of PDIA3 or ERp27 alone did not (Berger et al., 2020).

Co-expression of Ero1, PRDX4 or both alongside PDI effector genes was largely unsuccessful at increasing mAb productivity. It may be that redox regulation of PDI enzymes within the ER is highly sensitive to the precise expression level of multiple factors and that this level of precision was beyond the scope of these screens. For example, it was discovered that there is a negative feedback mechanism under highly oxidising conditions whereby PDI and TXNDC5 can catalyse formation of a regulatory disulphide within Ero1 that inactivates it, preventing it from oxidising PDI family members (Shepherd et al., 2014). This negative feedback mechanism may have been triggered by overexpression of Ero1, PRDX4 or both in the screens within this chapter. Similarly, PDI family members have been shown to be capable of oxidising each other (Oka et al., 2015). It could be that overexpression of some of the PDI variants screened in this chapter helps to oxidise other PDI variants and catalysis of disulphide bond formation by these other variants has a negative overall effect on mAb production.

In addition to catalysis of disulphide bond formation, other subsidiary functions of PDIs within the ER have been discovered. For instance, PDIA6 expression was found to attenuate the IRE1α branch of UPR signalling and another study showed that phosphorylation of PDI by the kinase Fam20C triggers a conformational change that is associated with a functional switch from disulphide bond formation to chaperone activity (Coelho & Feige; Eletto et al., 2014). Although these particular subsidiary PDI functions are not necessarily likely to hinder mAb production, this may be the case for others. For example, PDIA4 was shown in CHO to be involved in retention of misfolded proteins inside the ER, preventing their retrotranslocation to the cytosol for degradation (Forster et al., 2006). Therefore PDIA4 overexpression might increase the risk of UPR induction. There may be other such subsidiary PDI functions that hinder high level mAb production which are as yet unidentified.

# Chapter 7
# Conclusions and future work

## 7.1 Chapter 4 conclusion

Chapter 4 describes, for the first time, the use of synthetic promoter combinations to stably express the LC and HC of monoclonal antibodies. These synthetic promoter combinations displayed improved stable expression characteristics over conventionally-used hCMV promoters. Namely, synthetic promoters were capable of producing higher mean mAb titres than hCMV (up to 1.3-fold higher with mAb 2 – see figure 4.8) while recovering from MSX selection more quickly (up to 7 days more quickly for mAb 2 – see figure 4.7). These reduced recovery times will help to shorten cell line development timelines, an extremely desirable outcome for CHO cell-based mAb manufacturers.

## 7.2 Chapter 4 future work

There are more experiments that could be carried out to help further evaluate the performance of the LC/HC synthetic promoters shown in this chapter. First, single cell clones could be isolated from mAb 1 and mAb 2-expressing stable pools generated in this chapter and assessed for key characteristics such as growth rate, qP and product quality attributes. The results of flow cytometry assays shown in this chapter revealed that stable pools containing the LC1HC1 synthetic promoter combination, which generated higher mean mAb 1 and mAb 2 titres than hCMV-containing pools, had increased cell-to-cell variability in LC and HC expression compared to hCMV pools. This increased variability may lead to more variable clone performance, necessitating the screening of more clones to find one that shows high performance.

Second, the production stability of clones isolated from stable pools generated in this chapter could be assessed. Production instability, defined as the relative loss of qP of a cell line over many subcultures, is a significant problem in CHO-based recombinant mAb manufacture (Kim et al., 2011). The work of Kim et al identified two main causes of production instability: methylation-induced transcriptional silencing of the hCMV promoter and recombinant gene copy loss. Transcriptionally repressive DNA methylation occurs at CpG dinucleotides and all LC/HC synthetic promoters used in this chapter contain fewer CpG dinucleotides than the hCMV promoter. Kim et al found that loss of recombinant LC gene copies was more common than loss of recombinant

HC gene copies, likely due to homologous recombination (HR) events occurring between the two identical hCMV promoter sequences flanking the LC ORF. When using two different synthetic promoters to control LC and HC gene transcription, with the promoters having minimal sequence similarity due to being constructed from separate partitions of TFRE blocks, the risk of such HR events should be greatly reduced. It may therefore be hypothesised that assessing production stability of clones containing either LC/HC synthetic promoters or hCMV would reveal greater production stability conferred by the synthetic promoters. These assessments usually involve regular measurement of mAb productivity throughout long term culture that can last around 30 passages or 100 cell doublings (Bailey et al., 2012; Noh et al., 2018).

## 7.3 Chapter 5 conclusion

Chapter 5 describes the use of synthetic promoters to control transcription of the GS selection marker gene in stably expressing CHO cells. The results presented in this chapter have shown that increasing selection stringency on transfected cell populations can lead to increased qP at the cost of reduced IVCD. Crucially, this was shown to be possible via a synthetic GS promoter. When compared to the SV40 promoter, SynSV40_2 created more stringent selective pressure. When combined with synthetic promoters for the control of LC and HC transcription, SynSV40_2 led to elevated qP compared to the use of the SV40 promoter (up to 1.9-fold, see figure 5.7) and either equal or greater overall titre. Even a relative increase in qP that leads to the same amount of mAb being produced from a smaller cell density is a hugely beneficial outcome due to a reduction in burden on downstream purification processes.

## 7.4 Chapter 5 future work

There are more experiments that could be carried out to help further evaluate the performance of the GS synthetic promoters shown in this chapter. Firstly, single cell clones could be isolated from GS synthetic promoter-containing stable pools, such as those generated using expression vectors SynSV40_2 100RPU1 and SynSV40_2 LC1HC1, and assessed for key characteristics (growth rate, qP, product quality attributes and production stability).

Another experiment that may provide useful findings would be to use GS synthetic promoters described within this chapter for stable pool generation in the absence of MSX. For example, the use of GS1 - GS6 synthetic promoters alongside

50µM MSX appeared to create selective pressures that were too harsh to enable productive adaptations to occur. However if cells transfected with GS1 - GS6 were cultured in the absence of glutamine supplementation but without the added MSX inhibition then the level of selection stringency may be tuned to an effective level. Previous studies have shown that highly productive stable CHO cells can be generated in the absence of MSX by increasing selection stringency in other ways, i.e. through use of a mutant GS selection marker with reduced activity or use of a GS-KO host cell line (Lin et al., 2019; Noh et al., 2018). Using MSX in stable cell line generation can be problematic because cell lines selected in this way often see a reduction in productivity when MSX is removed from culture, as shown in the above study by Lin et al for example. Removal of MSX is necessary during the expansion of cell cultures prior to seeding industrial, large bioreactor scale fed batch runs.

## 7.5 Chapter 6 conclusion

In chapter 6 high-throughput transient transfection gene screening revealed pERP1 expression to be a novel target for CHO host cell engineering. pERP1 was shown robustly and consistently to increase the qP of cells producing two different mAbs (up to 1.6-fold, see figure 6.4) without inhibiting their growth rate. Co-expression of Ero1 alongside PDI caused a significant increase in mAb 3 qP compared to PDI expression alone. Apart from this however, co-expression of Ero1, PRDX4 or both alongside PDI effector genes did not yield significant improvements in either mean cell doubling time or mean qP. Likewise, high-throughput screening of ERdj effector genes did not reveal any of them to significantly improve either mean cell doubling time or mean qP.

## 7.6 Chapter 6 future work

Since results within this chapter showed pERP1 to increase mAb productivity, the development of CHO host cell lines that are stably expressing pERP1 is an important next step. This would involve transfection with a pERP1 expression vector, selection of cells stably expressing pERP1 and subsequent isolation of clones from this heterogenous stable pool. Selection would need to be performed using an alternative system to the GS selection system. This would ensure that the GS selection system can still be used to select for mAb-expressing cells upon transfection of the newly developed pERP1-expressing host with a mAb construct.

Following isolation of pERP1-expressing clones they would then need to be

tested for their mAb productivity. This could be assessed by transient transfection of these clonal populations with a mAb construct followed by measurement of mAb titre and VCD. The optimal pERP1 expression level cannot be deduced from the screening performed within this chapter. However, if the pERP1 construct was randomly integrated into the host genome upon transfection then each clone would display a different pERP1 expression level. Testing a large number of clones for their mAb productivity would thereby enable the optimal pERP1 expression level to be deduced.

These stable pERP1-expressing clones could also be used to help investigate the molecular mechanism of action of pERP1. As explained in the introductory subsection 6.1 of this chapter, Shimizu and colleagues identified that pERP1 expression allowed the internal disulphide bond of an IgG HC $CH_1$ domain to be formed even in the absence of LC expression (Shimizu et al., 2009). To discover this they first used siRNA to establish stable subclones of the mouse plasmacytoma cell line Ag8 with significantly reduced pERP1 expression. HC was then expressed in both pERP1-expressing and siRNA cells, immunoprecipitated and analysed by SDS-PAGE. HCs with different disulphide bond configurations could be identified because formation of disulphide bonds within a polypeptide alters its migration on an SDS-PAGE gel. Similar experiments could be performed in CHO host cell lines expressing pERP1, to establish whether pERP1 expression enables the internal $CH_1$ disulphide bond to be formed in the absence of LC expression in CHO.

Product quality attributes of mAb produced from pERP1-expressing host cells would also need to be assessed. The release of BiP from the $CH_1$ domain only upon LC-HC binding is viewed as a quality control step of the mAb folding and assembly process. If pERP1 does function to release BiP from $CH_1$ in the absence of LC-HC binding then this quality control step will be bypassed and so any potential impact on product quality will need to be evaluated. Furthermore, this supposed pERP1-induced removal of the requirement for LC-HC binding in order to release BiP would mean that free, unassembled HC can exist within the ER lumen without being recognised as misfolded due to prolonged BiP binding. As a result, HC dimers may be formed in addition to full mAb structures. It has been shown that HC dimers are capable of being assembled and secreted in CHO cells (Carl et al., 2020). The production of HC dimers from pERP1-expressing host cells would therefore need to be examined and quantified, as this could provide challenges during downstream mAb purification processes.

Finally, within this manuscript two engineering solutions designed to increase

the rate of BiP dissociation from HC have been described. Namely, the use of synthetic promoters to express LC in excess of HC, ensuring frequent LC-HC binding events (chapter 4) and the expression of pERP1 during mAb production. It may be interesting to test whether these two solutions have an additive effect when combined or whether they solve the same problem mechanistically and therefore there is no added effect of combining these solutions. For example, synthetic promoter-containing cells stably expressing a high LC:HC ratio, such as those expressing mAb 2 in chapter 4 (see figure 4.9), could be transiently transfected with a pERP1 vector as well as an empty vector control to compare the effect on mAb production.

# Reference list

Adamson, L., & Walum, E. (2007). Insulin and IGF-1 mediated inhibition of apoptosis in CHO cells grown in suspension in a protein-free medium [Article; Proceedings Paper]. *Atla-Alternatives to Laboratory Animals*, *35*(3), 349-352. https://doi.org/10.1177/026119290703500301

Agostinetto, R., Rossi, M., Dawson, J., Lim, A., Simoneau, M. H., Boucher, C., . . . Dey, A. K. (2022). Rapid cGMP manufacturing of COVID-19 monoclonal antibody using stable CHO cell pools. *Biotechnol Bioeng*, *119*(2), 663-666. https://doi.org/10.1002/bit.27995

Appenzeller-Herzog, C., & Ellgaard, L. (2008). The human PDI family: Versatility packed into a single fold. *Biochimica et biophysica acta. Molecular cell research*, *1783*(4), 535-548. https://doi.org/10.1016/j.bbamcr.2007.11.010

Arnosti, D. N., & Kulkarni, M. M. (2005). Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *J Cell Biochem*, *94*(5), 890-898. https://doi.org/10.1002/jcb.20352

Backliwal, G., Hildinger, M., Kuettel, I., Delegrange, F., Hacker, D. L., & Wurm, F. M. (2008). Valproic acid: A viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures [Article]. *Biotechnology and Bioengineering*, *101*(1), 182-189. https://doi.org/10.1002/bit.21882

Bailey, L. A., Hatton, D., Field, R., & Dickson, A. J. (2012). Determination of Chinese hamster ovary cell line stability and recombinant antibody expression during long-term culture. *Biotechnology and Bioengineering*, *109*(8), 2093-2103. https://doi.org/10.1002/bit.24485

Bandaranayake, A. D., & Almo, S. C. (2014). Recent advances in mammalian protein production. *FEBS Lett*, *588*(2), 253-260. https://doi.org/10.1016/j.febslet.2013.11.035

Baruah, G. L., & Belfort, G. (2004). Optimized recovery of monoclonal antibodies from transgenic goat milk by microfiltration [Article]. *Biotechnology and Bioengineering*, *87*(3), 274-285. https://doi.org/10.1002/bit.20112

Berger, A., Le Fourn, V., Masternak, J., Regamey, A., Bodenmann, I., Girod, P. A., & Mermod, N. (2020). Overexpression of transcription factor Foxa1 and target genes remediate therapeutic protein production bottlenecks in Chinese hamster ovary cells. *Biotechnol Bioeng*, *117*(4), 1101-1116. https://doi.org/10.1002/bit.27274

Black, C. B., Duensing, T. D., Trinkle, L. S., & Dunlay, R. T. (2011). Cell-Based Screening Using High-Throughput Flow Cytometry [Review]. *Assay and Drug Development Technologies*, *9*(1), 13-20. https://doi.org/10.1089/adt.2010.0308

Boune, S., Hu, P., Epstein, A. L., & Khawli, L. A. (2020). Principles of N-Linked Glycosylation Variations of IgG-Based Therapeutics: Pharmacokinetic and Functional Considerations. *Antibodies (Basel)*, *9*(2), 22. https://doi.org/10.3390/antib9020022

Brown, A. (2014). *Tools for Next-Generation Transcriptional Control in Chinese Hamster Ovary Cell Factories* University of Sheffield]. https://etheses.whiterose.ac.uk/7335/

Brown, A. J., Gibson, S. J., Hatton, D., & James, D. C. (2017). In silico design of context-responsive mammalian promoters with user-defined functionality [Article]. *Nucleic Acids Research*, *45*(18), 10906-10919. https://doi.org/10.1093/nar/gkx768

Brown, A. J., Sweeney, B., Mainwaring, D. O., & James, D. C. (2014). Synthetic Promoters for CHO Cell Engineering [Article]. *Biotechnology and Bioengineering*, *111*(8), 1638-1647. https://doi.org/10.1002/bit.25227

Carl, M., Soheila, B., Yao, Y., Jeanne, B., Gopalan, R., Daniela, T., . . . Laurence, F.-D. (2020). Crystal Structure and Characterization of Human Heavy-Chain Only Antibodies Reveals a Novel, Stable Dimeric Structure Similar to Monoclonal Antibodies. *Antibodies (Basel)*, *9*(4), 66. https://doi.org/10.3390/antib9040066

Cartwright, J. F., Arnall, C. L., Patel, Y. D., Barber, N. O. W., Lovelady, C. S., Rosignoli, G., . . . James, D. C. (2020). A platform for context-specific genetic engineering of recombinant protein production by CHO cells. *J Biotechnol*, *312*, 11-22. https://doi.org/10.1016/j.jbiotec.2020.02.012

Chahar, D. S., Ravindran, S., & Pisal, S. S. (2020). Monoclonal antibody purification and its progression to commercial scale [Review]. *Biologicals*, *63*, 1-13. https://doi.org/10.1016/j.biologicals.2019.09.007

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., E, P. R. I., . . . Church, G. M. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*, *12*(4), 326-328. https://doi.org/10.1038/nmeth.3312

Chen, F., Kou, T. C., Fan, L., Zhou, Y., Ye, Z. Y., 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 [Article]. *Biotechnology and Bioprocess Engineering*, *16*(6), 1157-1165. https://doi.org/10.1007/s12257-011-0069-8

Chiba, Y., Suzuki, M., Yoshida, S., Yoshida, A., Ikenaga, H., Takeuchi, M., . . . Ichishima, E. (1998). Production of human compatible high mannose-type (Man5GlcNAc2) sugar chains in Saccharomyces cerevisiae. *J Biol Chem*, *273*(41), 26298-26304. https://doi.org/10.1074/jbc.273.41.26298

Chin, C. L., Chin, H. K., Chin, C. S. H., Lai, E. T., & Ng, S. K. (2015). Engineering selection stringency on expression vector for the production of recombinant human alpha1-antitrypsin using Chinese Hamster ovary cells [Article]. *Bmc Biotechnology*, *15*, 15, Article 44. https://doi.org/10.1186/s12896-015-0145-9

Chotteau, V., Zhang, Y., & Clincke, M. F. (2015). Very High Cell Density in Perfusion of CHO Cells by ATF, TFF, Wave Bioreactor, and/or CellTank Technologies - Impact of Cell Density and Applications. In G. Subramanian (Ed.), *Continuous Processing in Pharmaceutical Manufacturing* (pp. 339-356). Wiley-V C H Verlag Gmbh.

Cockett, M. I., Bebbington, C. R., & Yarranton, G. T. (1990). HIGH-LEVEL EXPRESSION OF TISSUE INHIBITOR OF METALLOPROTEINASES IN CHINESE HAMSTER OVARY CELLS USING GLUTAMINE-SYNTHETASE GENE AMPLIFICATION [Article]. *Bio-Technology*, *8*(7), 662-667. https://doi.org/10.1038/nbt0790-662

Coelho, J. P. L., & Feige, M. J. In case of stress, hold tight: phosphorylation switches PDI from an oxidoreductase to a holdase, tuning ER proteostasis [Editorial Material; Early Access]. *Embo Journal*, 3, Article e104880. https://doi.org/10.15252/embj.2020104880

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 [Article]. *Biotechnology Progress*, *30*(1), 132-141. https://doi.org/10.1002/btpr.1809

Deer, J. R., & Allison, D. S. (2004). High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1 alpha gene [Article]. *Biotechnology Progress*, *20*(3), 880-889. https://doi.org/10.1021/bp034383r

Derouazi, M., Martinet, D., Besuchet Schmutz, N., Flaction, R., Wicht, M., Bertschinger, M., . . . Wurm, F. M. (2006). Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. *Biochem Biophys Res Commun*, *340*(4), 1069-1077. https://doi.org/10.1016/j.bbrc.2005.12.111

Donovan, K. W., & Bretscher, A. (2015). Tracking individual secretory vesicles during exocytosis reveals an ordered and regulated process. *J Cell Biol*, *210*(2), 181-189. https://doi.org/10.1083/jcb.201501118

Eletto, D., Dersh, D., Gidalevitz, T., & Argon, Y. (2014). Protein Disulfide Isomerase A6 Controls the Decay of IRE1 alpha Signaling via Disulfide-Dependent Association [Article]. *Molecular Cell*, *53*(4), 562-576. https://doi.org/10.1016/j.molcel.2014.01.004

Ellgaard, L., & Ruddock, L. W. (2005). The human protein disulphide isomerase family: substrate interactions and functional properties [Review]. *Embo Reports*, *6*(1), 28-32. https://doi.org/10.1038/sj.embor.7400311

Fan, L. C., Kadura, I., Krebs, L. E., Hatfield, C. C., Shaw, M. M., & Frye, C. C. (2012). Improving the efficiency of CHO cell line generation using glutamine synthetase gene knockout cells [Article]. *Biotechnology and Bioengineering*, *109*(4), 1007-1015. https://doi.org/10.1002/bit.24365

Fan, L. C., Kadura, I., Krebs, L. E., Larson, J. L., Bowden, D. M., & Frye, C. C. (2013). Development of a highly-efficient CHO cell line generation system with engineered SV40E promoter [Article]. *Journal of Biotechnology*, *168*(4), 652-658. https://doi.org/10.1016/j.jbiotec.2013.08.021

Feige, M. J., Hendershot, L. M., & Buchner, J. (2010). How antibodies fold [Review]. *Trends in Biochemical Sciences*, *35*(4), 189-198. https://doi.org/10.1016/j.tibs.2009.11.005

Ferreira, J. P., Peacock, R. W. S., Lawhorn, I. E. B., & Wang, C. L. (2011). Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters. *Syst Synth Biol*, *5*(3-4), 131-138. https://doi.org/10.1007/s11693-011-9089-0

Flintoff, W. F., & Essani, K. (1980). METHOTREXATE-RESISTANT CHINESE-HAMSTER OVARY CELLS CONTAIN A DIHYDROFOLATE-REDUCTASE WITH AN ALTERED AFFINITY FOR METHOTREXATE [Article]. *Biochemistry*, *19*(18), 4321-4327. https://doi.org/10.1021/bi00559a027

Forster, M. L., Sivick, K., Park, Y. N., Arvan, P., Lencer, W. I., & Tsai, B. (2006). Protein disulfide isomerase-like proteins play opposing roles during retrotranslocation [Article]. *Journal of Cell Biology*, *173*(6), 853-859. https://doi.org/10.1083/jcb.200602046

Fox, J. L. (2012). First plant-made biologic approved. *Nature biotechnology*, *30*(6), 472-472. https://doi.org/10.1038/nbt0612-472

Gagnon, M., Hiller, G., Luan, Y. T., Kittredge, A., DeFelice, J., & Drapeau, D. (2011). High-End pH-Controlled Delivery of Glucose Effectively Suppresses Lactate Accumulation in CHO Fed-Batch Cultures [Article]. *Biotechnology and Bioengineering*, *108*(6), 1328-1337. https://doi.org/10.1002/bit.23072

Geisler, C., & Jarvis, D. L. (2018). Adventitious viruses in insect cell lines used for recombinant protein expression [Review]. *Protein Expression and Purification*, *144*, 25-32. https://doi.org/10.1016/j.pep.2017.11.002

Gibson, B., Wilson, D. J., Feil, E., & Eyre-Walker, A. (2018). The distribution of bacterial doubling times in the wild [Article]. *Proceedings of the Royal Society B-Biological Sciences*, *285*(1880), 9, Article 20180789. https://doi.org/10.1098/rspb.2018.0789

Goey, C. H., Alhuthali, S., & Kontoravdi, C. (2018). Host cell protein removal from biopharmaceutical preparations: Towards the implementation of quality by design. *Biotechnology Advances*, *36*(4), 1223-1237. https://doi.org/10.1016/j.biotechadv.2018.03.021

Gomord, V., Chamberlain, P., Jefferis, R., & Faye, L. (2005). Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol*, *23*(11), 559-565. https://doi.org/10.1016/j.tibtech.2005.09.003

Gonzalez, R., Andrews, B. A., & Asenjo, J. A. (2002). Kinetic model of BiP- and PDI-mediated protein folding and assembly [Article]. *Journal of Theoretical Biology*, *214*(4), 529-537. https://doi.org/10.1006/jtbi.2001.2478

Grabherr, M. G., Pontiller, J., Mauceli, E., Ernst, W., Baumann, M., Biagi, T., . . . Grabherr, R. M. (2011). Exploiting nucleotide composition to engineer promoters. *PLoS One*, *6*(5), e20136-e20136. https://doi.org/10.1371/journal.pone.0020136

Grossmann, M., Wong, R., Teh, N. G., Tropea, J. E., EastPalmer, J., Weintraub, B. D., & Szkudlinski, M. W. (1997). Expression of biologically active human thyrotropin (hTSH) in a baculovirus system: Effect of insect cell glycosylation on hTSH activity in vitro and

in vivo [Article; Proceedings Paper]. *Endocrinology*, *138*(1), 92-100. https://doi.org/10.1210/en.138.1.92

Gstraunthaler, G. (2003). Alternatives to the use of fetal bovine serum: Serum-free cell culture [Article]. *Altex-Alternativen Zu Tierexperimenten*, *20*(4), 275-281.

Gupta, K., Tolzer, C., Sari-Ak, D., Fitzgerald, D. J., Schaffitzel, C., & Berger, I. (2019). MultiBac: Baculovirus-Mediated Multigene DNA Cargo Delivery in Insect and Mammalian Cells [Review]. *Viruses-Basel*, *11*(3), 14, Article 198. https://doi.org/10.3390/v11030198

Gupta, S. K., & Shukla, P. (2017). Microbial platform technology for recombinant antibody fragment production: A review [Review]. *Critical Reviews in Microbiology*, *43*(1), 31-42. https://doi.org/10.3109/1040841x.2016.1150959

Haberle, V., & Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol*, *19*(10), 621-637.

Harfst, E., & Johnstone, A. P. (1992). CHARACTERIZATION OF THE GLUTAMINE-SYNTHETASE AMPLIFIABLE EUKARYOTIC EXPRESSION SYSTEM APPLIED TO AN INTEGRAL MEMBRANE-PROTEIN - THE HUMAN THYROTROPIN RECEPTOR [Article]. *Analytical Biochemistry*, *207*(1), 80-84. https://doi.org/10.1016/0003-2697(92)90504-z

Hu, T., Yeh, J. E., Pinello, L., Jacob, J., Chakravarthy, S., Yuan, G. C., . . . Frank, D. A. (2015). Impact of the N-Terminal Domain of STAT3 in STAT3-Dependent Transcriptional Activity. *Molecular and Cellular Biology*, *35*(19), 3284-3300. https://doi.org/10.1128/mcb.00060-15

Islam, M. A., Park, T. E., Singh, B., Maharjan, S., Firdous, J., Cho, M. H., . . . Cho, C. S. (2014). Major degradable polycations as carriers for DNA and siRNA [Review]. *Journal of Controlled Release*, *193*, 74-89. https://doi.org/10.1016/j.jconrel.2014.05.055

Jacobs, P. P., Geysens, S., Vervecken, W., Contreras, R., & Callewaert, N. (2009). Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology. *Nat Protoc*, *4*(1), 58-70. https://doi.org/10.1038/nprot.2008.213

Jansen, G., Maattanen, P., Denisov, A. Y., Scarffe, L., Schade, B., Balghi, H., . . . Thomas, D. Y. (2012). An Interaction Map of Endoplasmic Reticulum Chaperones and Foldases [Article]. *Molecular & Cellular Proteomics*, *11*(9), 710-723. https://doi.org/10.1074/mcp.M111.016550

Jayapal, K. R., Wlaschin, K. F., Hu, W. S., & Yap, M. G. S. (2007). Recombinant protein therapeutics from CHO cells - 20 years and counting [Article]. *Chemical Engineering Progress*, *103*(10), 40-47.

Jeong, J., Cho, N., Jung, D., & Bang, D. (2013). Genome-scale genetic engineering in Escherichia coli [Review]. *Biotechnology Advances*, *31*(6), 804-810. https://doi.org/10.1016/j.biotechadv.2013.04.003

Jessop, C. E., Watkins, R. H., Simmons, J. J., Tasab, M., & Bulleid, N. J. (2009). Protein disulphide isomerase family members show distinct substrate specificity: P5 is targeted to BiP client proteins [Article]. *Journal of Cell Science*, *122*(23), 4287-4295. https://doi.org/10.1242/jcs.059154

Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y. Z., Wright, C. M., Estes, S. D., . . . James, D. C. (2019). CHO genome mining for synthetic promoter design [Article]. *Journal of Biotechnology*, *294*, 1-13. https://doi.org/10.1016/j.jbiotec.2019.01.015

Kaplon, H., Muralidharan, M., Schneider, Z., & Reichert, J. M. (2020). Antibodies to watch in 2020 [Article]. *Mabs*, *12*(1), 24, Article Unsp 1703531. https://doi.org/10.1080/19420862.2019.1703531

Kaplon, H., & Reichert, J. M. (2018). Antibodies to watch in 2018 [Article]. *Mabs*, *10*(2), 183-203. https://doi.org/10.1080/19420862.2018.1415671

Kaplon, H., & Reichert, J. M. (2019). Antibodies to watch in 2019 [Article]. *Mabs*, *11*(2), 219-238. https://doi.org/10.1080/19420862.2018.1556465

Karmali, P. P., & Chaudhuri, A. (2007). Cationic Liposomes as non-viral carriers of gene medicines: Resolved issues, open questions, and future promises [Review]. *Medicinal Research Reviews*, *27*(5), 696-722. https://doi.org/10.1002/med.20090

Karst, D. J., Steinebach, F., Soos, M., & Morbidelli, M. (2017). Process performance and product quality in an integrated continuous antibody production process [Article]. *Biotechnology and Bioengineering*, *114*(2), 298-307. https://doi.org/10.1002/bit.26069

Kaufman, R. J., & Sharp, P. A. (1982). AMPLIFICATION AND EXPRESSION OF SEQUENCES COTRANSFECTED WITH A MODULAR DIHYDROFOLATE-REDUCTASE COMPLEMENTARY-DNA GENE [Article]. *Journal of Molecular Biology*, *159*(4), 601-621. https://doi.org/10.1016/0022-2836(82)90103-6

Kelley, B. (2020). Developing therapeutic monoclonal antibodies at pandemic pace. *Nat Biotechnol*, *38*(5), 540-545. https://doi.org/10.1038/s41587-020-0512-5

Kim, M., O'Callaghan, P. M., Droms, K. A., & James, D. C. (2011). A Mechanistic Understanding of Production Instability in CHO Cell Lines Expressing Recombinant Monoclonal Antibodies [Article]. *Biotechnology and Bioengineering*, *108*(10), 2434-2446. https://doi.org/10.1002/bit.23189

Kim, N. S., Byun, T. H., & Lee, G. M. (2001). Key determinants in the occurrence of clonal variation in humanized antibody expression of CHO cells during dihydrofolate reductase mediated gene amplification [Article]. *Biotechnology Progress*, *17*(1), 69-75. https://doi.org/10.1021/bp000144h

Kunert, R., & Reinhart, D. (2016). Advances in recombinant antibody manufacturing. *Appl Microbiol Biotechnol*, *100*(8), 3451-3461. https://doi.org/10.1007/s00253-016-7388-9

Lai, H., He, J., Hurtado, J., Stahnke, J., Fuchs, A., Mehlhop, E., . . . Chen, Q. (2014). Structural and functional characterization of an anti-West Nile virus monoclonal antibody and its single-chain variant produced in glycoengineered plants. *Plant Biotechnol J*, *12*(8), 1098-1107. https://doi.org/10.1111/pbi.12217

Le, H., Vishwanathan, N., Jacob, N. M., Gadgil, M., & Hu, W. S. (2015). Cell line development for biomanufacturing processes: recent advances and an outlook [Review]. *Biotechnology Letters*, *37*(8), 1553-1564. https://doi.org/10.1007/s10529-015-1843-z

Lee, S. M., Hickey, J. M., Miura, K., Joshi, S. B., Volkin, D. B., King, C. R., & Plieskatt, J. L. (2020). A C-terminal Pfs48/45 malaria transmission-blocking vaccine candidate produced in the baculovirus expression system [Article]. *Scientific Reports*, *10*(1), 14, Article 395. https://doi.org/10.1038/s41598-019-57384-w

Lee, S. M., Wu, C. K., Plieskatt, J. L., Miura, K., Hickey, J. M., & King, C. R. (2017). N-Terminal Pfs230 Domain Produced in Baculovirus as a Biological Active Transmission-Blocking Vaccine Candidate [Article]. *Clinical and Vaccine Immunology*, *24*(10), 14, Article e00140-17. https://doi.org/10.1128/cvi.00140-17

Lewis, N. E., Liu, X., Li, Y. X., Nagarajan, H., Yerganian, G., O'Brien, E., . . . Palsson, B. O. (2013). Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome [Article]. *Nature Biotechnology*, *31*(8), 759-+. https://doi.org/10.1038/nbt.2624

Li, J. D., Zhang, C. C., Jostock, T., & Dubel, S. (2007). Analysis of IgG heavy chain to light chain ratio with mutant Encephalomyocarditis virus internal ribosome entry site [Article]. *Protein Engineering Design & Selection*, *20*(10), 491-496. https://doi.org/10.1093/protein/gzm038

Lin, P. C., Chan, K. F., Kiess, I. A., Tan, J., Shahreel, W., Wong, S. Y., & Song, Z. W. (2019). Attenuated glutamine synthetase as a selection marker in CHO cells to efficiently isolate highly productive stable cells for the production of antibodies and other biologics. *Mabs*, *11*(5), 965-976. https://doi.org/10.1080/19420862.2019.1612690

Liu, H., & May, K. (2012). Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs*, *4*(1), 17-23. https://doi.org/10.4161/mabs.4.1.18347

Maasho, K., Marusina, A., Reynolds, N. M., Coligan, J. E., & Borrego, F. (2004). Efficient gene transfer into the human natural killer cell line, NKL, using the Amaxa nucleofection system (TM) [Article]. *Journal of Immunological Methods*, *284*(1-2), 133-140. https://doi.org/10.1016/j.jim.2003.10.010

Maksimenko, O. G., Deykin, A. V., Khodarovich, Y. M., & Georgiev, P. G. (2013). Use of Transgenic Animals in Biotechnology: Prospects and Problems [Review]. *Acta Naturae*, *5*(1), 33-46. https://doi.org/10.32607/20758251-2013-5-1-33-46

Malekian, R., Sima, S., Jahanian-Najafabadi, A., Moazen, F., & Akbari, V. (2019). Improvement of soluble expression of GM-CSF in the cytoplasm of Escherichia coli using chemical and molecular chaperones [Article]. *Protein Expression and Purification*, *160*, 66-72. https://doi.org/10.1016/j.pep.2019.04.002

Mamat, U., Wilke, K., Bramhill, D., Schromm, A. B., Lindner, B., Kohl, T. A., . . . Woodard, R. W. (2015). Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins [Article]. *Microbial Cell Factories*, *14*, 15, Article 57. https://doi.org/10.1186/s12934-015-0241-5

Mann, C. J. (2007). Rapid isolation of antigen-specific clones from hybridoma fusions [Editorial Material]. *Nature Methods*, 8-9.

Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy Asthma Clin Immunol*, *14*(Suppl 2), 49-49. https://doi.org/10.1186/s13223-018-0278-1

Moscou, M. J., & Bogdanove, A. J. (2009). Simple Cipher Governs DNA Recognition by TAL Effectors. *Science*, *326*(5959), 1501-1501. https://doi.org/10.1126/science.1178817

Mulukutla, B. C., Kale, J., Kalomeris, T., Jacobs, M., & Hiller, G. W. (2017). Identification and control of novel growth inhibitors in fed-batch cultures of Chinese hamster ovary cells [Article]. *Biotechnology and Bioengineering*, *114*(8), 1779-1790. https://doi.org/10.1002/bit.26313

Nakamura, H., & Funahashi, J. (2013). Electroporation: Past, present and future [Review]. *Development Growth & Differentiation*, *55*(1), 15-19. https://doi.org/10.1111/dgd.12012

Nematpour, F., Mahboudi, F., Khalaj, V., Vaziri, B., Ahmadi, S., Ahmadi, M., . . . Davami, F. (2017). Optimization of monoclonal antibody expression in CHO cells by employing epigenetic gene regulation tools [Article]. *Turkish Journal of Biology*, *41*(4), 622-628. https://doi.org/10.3906/biy-1702-18

Nett, J. H., Stadheim, T. A., Li, H., Bobrowicz, P., Hamilton, S. R., Davidson, R. C., . . . Gerngross, T. U. (2011). A combinatorial genetic library approach to target heterologous glycosylation enzymes to the endoplasmic reticulum or the Golgi apparatus of Pichia pastoris. *Yeast*, *28*(3), 237-252. https://doi.org/10.1002/yea.1835

Ng, S. K., Wang, D. I. C., & Yap, M. G. S. (2007). Application of destabilizing sequences on selection marker for improved recombinant protein productivity in CHO-DG44 [Article]. *Metabolic Engineering*, *9*(3), 304-316. https://doi.org/10.1016/j.ymben.2007.01.001

Noh, S. M., Shin, S., & Lee, G. M. (2018). Comprehensive characterization of glutamine synthetase-mediated selection for the establishment of recombinant CHO cells producing monoclonal antibodies [Article]. *Scientific Reports*, *8*, 11, Article 5361. https://doi.org/10.1038/s41598-018-23720-9

O'Flaherty, R., Bergin, A., Flampouri, E., Mota, L. M., Obaidi, I., Quigley, A., . . . Butler, M. (2020). Mammalian cell culture for production of recombinant proteins: A review of

the critical steps in their biomanufacturing [; Review]. *Biotechnology advances*, 107552. https://doi.org/10.1016/j.biotechadv.2020.107552

Ogawa, R., Kagiya, G., Kodaki, T., Fukuda, S., & Yamamoto, K. (2007). Construction of strong mammalian promoters by random cis-acting element elongation. *Biotechniques*, *42*(5), 628-633. https://doi.org/10.2144/000112436

Oka, O. B. V., Yeoh, H. Y., & Bulleid, N. J. (2015). Thiol-disulfide exchange between the PDI family of oxidoreductases negates the requirement for an oxidase or reductase for each enzyme [Article]. *Biochemical Journal*, *469*, 279-288. https://doi.org/10.1042/bj20141423

Owczarek, B., Gerszberg, A., & Hnatuszko-Konka, K. (2019). A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals. *Biomed Res Int*, *2019*, 4216060. https://doi.org/10.1155/2019/4216060

Palmberger, D., Wilson, I. B. H., Berger, I., Grabherr, R., & Rendic, D. (2012). SweetBac: A New Approach for the Production of Mammalianised Glycoproteins in Insect Cells [Article]. *Plos One*, *7*(4), 8, Article e34226. https://doi.org/10.1371/journal.pone.0034226

Perez-Pinera, P., Ousterout, D. G., Brunger, J. M., Farin, A. M., Glass, K. A., Guilak, F., . . . Gersbach, C. A. (2013). Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Methods*, *10*(3), 239-242. https://doi.org/10.1038/nmeth.2361

Pilbrough, W., Munro, T. P., & Gray, P. (2009). Intraclonal Protein Expression Heterogeneity in Recombinant CHO Cells [Article]. *Plos One*, *4*(12), 11, Article e8432. https://doi.org/10.1371/journal.pone.0008432

Pobre, K. F. R., Poet, G. J., & Hendershot, L. M. (2019). The endoplasmic reticulum (ER) chaperone BiP is a master regulator of ER functions: Getting by with a little help from ERdj friends. *J Biol Chem*, *294*(6), 2098-2108. https://doi.org/10.1074/jbc.REV118.002804

Popovic, B., Gibson, S., Senussi, T., Carmen, S., Kidd, S., Slidel, T., . . . Lowe, D. (2017). Engineering the expression of an anti-interleukin-13 antibody through rational design and mutagenesis [Article]. *Protein Engineering Design & Selection*, *30*(4), 303-311. https://doi.org/10.1093/protein/gzx001

Povey, J. F., O'Malley, C. J., Root, T., Martin, E. B., Montague, G. A., Feary, M., . . . Smales, C. M. (2014). Rapid high-throughput characterisation, classification and selection of recombinant mammalian cell line phenotypes using intact cell MALDI-ToF mass spectrometry fingerprinting and PLS-DA modelling. *J Biotechnol*, *184*, 84-93. https://doi.org/10.1016/j.jbiotec.2014.04.028

Preissler, S., Rato, C., Perera, L. A., Saudek, V., & Ron, D. (2017). FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP. *Nat Struct Mol Biol*, *24*(1), 23-29. https://doi.org/10.1038/nsmb.3337

Priola, J. J., Calzadilla, N., Baumann, M., Borth, N., Tate, C. G., & Betenbaugh, M. J. (2016). High-throughput screening and selection of mammalian cells for enhanced protein production [Review]. *Biotechnology Journal*, *11*(7), 853-865. https://doi.org/10.1002/biot.201500579

Pybus, L. P., Dean, G., West, N. R., Smith, A., Daramola, O., Field, R., . . . James, D. C. (2014). Model-Directed Engineering of ''Difficult-to-Express'' Monoclonal Antibody Production by Chinese Hamster Ovary Cells [Article]. *Biotechnology and Bioengineering*, *111*(2), 372-385. https://doi.org/10.1002/bit.25116

Rahbarizadeh, F., Rasaee, M. J., Forouzandeh, M., & Allameh, A. A. (2006). Over expression of anti-MUC1 single-domain antibody fragments in the yeast Pichia pastoris. *Mol Immunol*, *43*(5), 426-435. https://doi.org/10.1016/j.molimm.2005.03.003

Raju, T. S., Briggs, J. B., Borge, S. M., & Jones, A. J. S. (2000). Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics [Article]. *Glycobiology*, *10*(5), 477-486. https://doi.org/10.1093/glycob/10.5.477

Rosenbaum, M., Andreani, V., Kapoor, T., Herp, S., Flach, H., Duchniewicz, M., & Grosschedl, R. (2014). MZB1 is a GRP94 cochaperone that enables proper immunoglobulin heavy chain biosynthesis upon ER stress [Article]. *Genes & Development*, *28*(11), 1165-1178. https://doi.org/10.1101/gad.240762.114

Rössger, K., Charpin-El-Hamri, G., & Fussenegger, M. (2014). Bile acid-controlled transgene expression in mammalian cells and mice. *Metab Eng*, *21*, 81-90. https://doi.org/10.1016/j.ymben.2013.11.003

Sautter, K., & Enenkel, B. (2005). Selection of high-producing CHO cells using NPT selection marker with reduced enzyme activity [Article]. *Biotechnology and Bioengineering*, *89*(5), 530-538. https://doi.org/10.1002/bit.20374

Schildknegt, D., Lodder, N., Pandey, A., Egmond, M., Pena, F., Braakman, I., & van der Sluijs, P. (2019). Characterization of CNPY5 and its family members [Article]. *Protein Science*, *28*(7), 1276-1289. https://doi.org/10.1002/pro.3635

Schillberg, S., Raven, N., Spiegel, H., Rasche, S., & Buntru, M. (2019). Critical analysis of the commercial potential of plants for the production of recombinant proteins. *Front Plant Sci*, *10*, 720-720. https://doi.org/10.3389/fpls.2019.00720

Schlabach, M. R., Hu, J. K., Li, M., & Elledge, S. J. (2010). Synthetic design of strong promoters. *Proc Natl Acad Sci U S A*, *107*(6), 2538-2543. https://doi.org/10.1073/pnas.0914803107

Schlatter, S., Stansfield, S. H., Dinnis, D. M., Racher, A. J., Birch, J. R., & James, D. C. (2005). On the optimal ratio of heavy to light chain genes for efficient recombinant antibody production by CHO cells [Article; Proceedings Paper]. *Biotechnology Progress*, *21*(1), 122-133. https://doi.org/10.1021/bp049780w

Sgrignani, J., Garofalo, M., Matkovic, M., Merulla, J., Catapano, C. V., & Cavalli, A. (2018). Structural Biology of STAT3 and Its Implications for Anticancer Therapies Development. *International Journal of Molecular Sciences*, *19*(6), Article 1591. https://doi.org/10.3390/ijms19061591

Shepelev, M. V., Kalinichenko, S. V., Deykin, A. V., & Korobko, I. V. (2018). Production of Recombinant Proteins in the Milk of Transgenic Animals: Current State and Prospects [Review]. *Acta Naturae*, *10*(3), 40-47. https://doi.org/10.32607/20758251-2018-10-3-40-47

Shepherd, C., Oka, O. B. V., & Bulleid, N. J. (2014). Inactivation of mammalian Ero1 alpha is catalysed by specific protein disulfide-isomerases [Article]. *Biochemical Journal*, *461*, 107-113. https://doi.org/10.1042/bj20140234

Shimizu, Y., Meunier, L., & Hendershot, L. M. (2009). pERp1 is significantly up-regulated during plasma cell differentiation and contributes to the oxidative folding of immunoglobulin [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(40), 17013-17018. https://doi.org/10.1073/pnas.0811591106

Sinzger, C., Digel, M., & Jahn, G. (2008). Cytomegalovirus cell tropism. *Human Cytomegalovirus*, *325*, 63-83.

Sola, I., Castilla, J., Pintado, B., Sanchez-Morgado, J. M., Whitelaw, C. B. A., Clark, A. J., & Enjuanes, L. (1998). Transgenic mice secreting coronavirus neutralizing antibodies into the milk [Article]. *Journal of Virology*, *72*(5), 3762-3772. https://doi.org/10.1128/jvi.72.5.3762-3772.1998

Spadiut, O., Capone, S., Krainer, F., Glieder, A., & Herwig, C. (2014). Microbials for the production of monoclonal antibodies and antibody fragments. *Trends Biotechnol*, *32*(1), 54-60.

Strasser, R., Stadlmann, J., Schähs, M., Stiegler, G., Quendler, H., Mach, L., . . . Steinkellner, H. (2008). Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J*, *6*(4), 392-402. https://doi.org/10.1111/j.1467-7652.2008.00330.x

Swindley, O. (2021). *Manipulation of Endoplasmic Reticulum Folding and Assembly Machinery for Improved Production of Complex Biotherapeutics* University of Sheffield]. https://etheses.whiterose.ac.uk/29353/

Takagi, Y., Kikuchi, T., Wada, R., & Omasa, T. (2017). The enhancement of antibody concentration and achievement of high cell density CHO cell cultivation by adding nucleoside [Article]. *Cytotechnology*, *69*(3), 511-521. https://doi.org/10.1007/s10616-017-0066-7

Tange, T. O., Nott, A., & Moore, M. J. (2004). The ever-increasing complexities of the exon junction complex [Review]. *Current Opinion in Cell Biology*, *16*(3), 279-284. https://doi.org/10.1016/j.ceb.2004.03.012

Tian, J., He, Q., Oliveira, C., Qian, Y. M., Egan, S., Xu, J. L., . . . Li, Z. J. (2020). Increased MSX level improves biological productivity and production stability in multiple recombinant GS CHO cell lines. *Engineering in Life Sciences*, *20*(3-4), 112-125. https://doi.org/10.1002/elsc.201900124

Tornøe, J., Kusk, P., Johansen, T. E., & Jensen, P. R. (2002). Generation of a synthetic mammalian promoter library by modification of sequences spacing transcription factor binding sites. *Gene*, *297*(1), 21-32. https://doi.org/10.1016/S0378-1119(02)00878-8

Tripathi, N. K., & Shrivastava, A. (2019). Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development [Review]. *Frontiers in Bioengineering and Biotechnology*, *7*, 35, Article 420. https://doi.org/10.3389/fbioe.2019.00420

Valderrama-Rincon, J. D., Fisher, A. C., Merritt, J. H., Fan, Y. Y., Reading, C. A., Chhiba, K., . . . DeLisa, M. P. (2012). An engineered eukaryotic protein glycosylation pathway in Escherichia coli [Article]. *Nature Chemical Biology*, *8*(5), 434-436. https://doi.org/10.1038/nchembio.921

van Anken, E., Pena, F., Hafkemeijer, N., Christis, C., Romijn, E. P., Grauschopf, U., . . . Braakman, I. (2009). Efficient IgM assembly and secretion require the plasma cell induced endoplasmic reticulum protein pERp1 [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(40), 17019-17024. https://doi.org/10.1073/pnas.0903036106

Vanhove, M., Usherwood, Y. K., & Hendershot, L. M. (2001). Unassembled Ig heavy chains do not cycle from BiP in vivo but require light chains to trigger their release [Article]. *Immunity*, *15*(1), 105-114. https://doi.org/10.1016/s1074-7613(01)00163-7

Walsh, G. (2018). Biopharmaceutical benchmarks 2018 [Article]. *Nature Biotechnology*, *36*(12), 1136-1145. https://doi.org/10.1038/nbt.4305

Walter, P., & Ron, D. (2011). The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation [Review]. *Science*, *334*(6059), 1081-1086. https://doi.org/10.1126/science.1209038

Wang, T. Y., & Guo, X. (2020). Expression vector cassette engineering for recombinant therapeutic production in mammalian cell systems [Review; Early Access]. *Applied Microbiology and Biotechnology*, 16. https://doi.org/10.1007/s00253-020-10640-w

Wang, Y., Li, X., Chen, X., Nielsen, J., Petranovic, D., & Siewers, V. (2021). Expression of antibody fragments in Saccharomyces cerevisiae strains evolved for enhanced protein secretion. *Microb Cell Fact*, *20*(1), 134. https://doi.org/10.1186/s12934-021-01624-0

West, N. (2014). *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].
https://etheses.whiterose.ac.uk/6701/

Westwood, A. D., Rowe, D. A., & Clarke, H. R. G. (2010). Improved Recombinant Protein Yield Using a Codon Deoptimized DHFR Selectable Marker in a CHEF1 Expression Plasmid [Article]. *Biotechnology Progress*, *26*(6), 1558-1566. https://doi.org/10.1002/btpr.491

Wiedemann, C., Kumar, A., Lang, A., & Ohlenschläger, O. (2020). Cysteines and Disulfide Bonds as Structure-Forming Units: Insights From Different Domains of Life and the Potential for Characterization by NMR. *Front Chem*, *8*, 280-280. https://doi.org/10.3389/fchem.2020.00280

Wurm, F. M., & Wurm, M. J. (2017). Cloning of CHO Cells, Productivity and Genetic Stability—A Discussion. *Processes*, *5*(2), 20.

Yamaguchi, H., & Miyazaki, M. (2014). Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies [Evaluation Studies; ; Research Support, Non-U.S. Gov't; Review]. *Biomolecules*, *4*(1), 235-251. https://doi.org/10.3390/biom4010235

Ye, J., Ly, J., Watts, K., Hsu, A., Walker, A., McLaughlin, K., . . . Potgieter, T. (2011). Optimization of a glycoengineered Pichia pastoris cultivation process for commercial antibody production. *Biotechnol Prog*, *27*(6), 1744-1750. https://doi.org/10.1002/btpr.695

Yee, C. M., Zak, A. J., Hill, B. D., & Wen, F. (2018). The Coming Age of Insect Cells for Manufacturing and Development of Protein Therapeutics [Review]. *Industrial & Engineering Chemistry Research*, *57*(31), 10061-10070. https://doi.org/10.1021/acs.iecr.8b00985

Zanghi, J. A., Renner, W. A., Bailey, J. E., & Fussenegger, M. (2000). The growth factor inhibitor suramin reduces apoptosis and cell aggregation in protein-free CHO cell batch cultures [Article]. *Biotechnology Progress*, *16*(3), 319-325. https://doi.org/10.1021/bp0000353

Zhang, R., Rao, M., Li, C., Cao, J. Y., Meng, Q. L., Zheng, M., . . . Li, N. (2009). Functional recombinant human anti-HAV antibody expressed in milk of transgenic mice [Article]. *Transgenic Research*, *18*(3), 445-453. https://doi.org/10.1007/s11248-008-9241-0

Zhu, J., Ruan, Y., Fu, X., Zhang, L. C., Ge, G. S., Wall, J. G., . . . Hu, X. J. (2020). An Engineered Pathway for Production of Terminally Sialylated N-glycoproteins in the Periplasm of Escherichia coli [Article]. *Frontiers in Bioengineering and Biotechnology*, *8*, 9, Article 313. https://doi.org/10.3389/fbioe.2020.00313

Zimmermann, R., Eyrisch, S., Ahmad, M., & Helms, V. (2011). Protein translocation across the ER membrane. *Biochim Biophys Acta*, *1808*(3), 912-924. https://doi.org/10.1016/j.bbamem.2010.06.015

Zito, E., Melo, E. P., Yang, Y., Wahlander, A., Neubert, T. A., & Ron, D. (2010). Oxidative Protein Folding by an Endoplasmic Reticulum-Localized Peroxiredoxin [Article]. *Molecular Cell*, *40*(5), 787-797. https://doi.org/10.1016/j.molcel.2010.11.010