A thesis submitted to


The University Of Sheffield.

Department of Chemical and Biological Engineering.

For the degree of
Doctor of Philosophy (PhD)

# Multigene engineering of secretion in mammalian cell factories

**Nicholas O. W. Barber**

August 2018. Corrections submitted July 2019.

# Declaration.

I, Nicholas Owen Waterman Barber, declare that I am the sole author of this thesis and that the research presented within is a result of my own efforts and achievements. Where this is not the case, this has clearly been stated. I confirm that this work has not been submitted for any other degrees.

## Abstract

Chinese hamster ovary cells are widely used in the production of biopharmaceuticals such as monoclonal antibodies. Since their first use for this application in the 1980s much engineering of manufacturing processes and the CHO cell itself has been undertaken, leading to increased cell productivity, product specificity and overall product titres and yields. Much CHO cell engineering has thus far focused on increasing transcription and translation levels, with the engineering of protein folding and modification also being targeted. Whilst this has increased the protein production capacity of the CHO cell it has also had the effect of increasing pressure on the CHO biosynthetic and secretory pathways. The secretory pathway of the CHO cell is not adapted to high levels of secretion as is seen in plasma cells, the main function of which is to produce and secret high levels of antibodies. As such an increase in secretory load upon the CHO cell has introduced bottlenecks that limit the amount of recombinant product that can be transported out of the cell. Engineering strategies to ease these bottlenecks have so far focused on expression of single genes or two genes in combination. Two engineering strategies were utilised to engineer the CHO cell secretory pathway and enhance CHO cell productivity. Firstly a literature- and 'omics-driven approach was used to select genes with which to engineer the CHO secretory pathway. Engineering of the secretory and biosynthetic pathways, using single genes and multiple genes in combination, showed that this approach can increase CHO productivity levels through enhancing specific productivity and overall titre. Secondly, directed evolution was used as a method to bring about myriad fine global changes within the CHO biosynthetic and secretory pathways. From these results we hypothesise that a more holistic, global approach to biosynthetic and secretory pathway engineering proves more successful in increasing CHO cell productivity when compared to more directed techniques.

## Acknowledgements

After four years of study I now feel confident that I could take on a research project to a good level. Whilst it may have been favourable for this feeling to occur previously, in this respect I suppose the doctoral training programme has had its desired effect. The following acknowledgements have formed over the last four years. Many times I have thought of the day I could write this small section. Needless to say it will probably not match the hype.

Sincerest thanks go to my academic supervisors, Prof. David James and Dr. Andrew Peden, for the opportunity to work with them, their project outlook and guidance. A great debt of thanks is owed to my industrial supervisor, Dr. Clare Lovelady, for her support, guidance, organisation and thesis writing help. Thanks go to my sponsors at MedImmune for project support and context, and the BBSRC (through the White Rose DTP).

My heartfelt thanks to those who I have interacted with daily in the lab. The entire DCJ group have been of a great help throughout and are too numerous to list lest I forget one of many pivotal figures, but special thanks are due to those I have worked most closely alongside over the past four years: Clare Arnall, Joe Cartwright, Yash Patel and Devika Kalsi. Thanks also to the Peden lab group who put up with me showing up relatively unannounced to get under their feet.

The support outside of my studies has been pivotal in me reaching this point. Sheffield University Orienteering and Fell Running Club (ShUOC) and the Sheffield University Singers' Society (SingSoc) have been channels of relief, friendship and (in)sanity during my two stints in Sheffield. A list of helpful friends could extend this thesis beyond its already interminable length, but lunch time pool sessions with Dan Hartmann were always anticipated; two-way shrink session runs with Craig Fishwick in the last year were enjoyable; and "how are the hamsters doing?" messages from Shaun Davies a reminder of a simpler time.

My utmost love and dedication are due to Hannah Bradley for surfing the rollercoaster of the last 3.5 years at my side, celebrating the highs and supporting the lows.

Finally, those who have offered me their unyielding love and support in everything I have applied myself to for the past three decades – Jo, Julian and Catherine Barber.

Nicholas O. W. Barber, Sheffield, August 2018. Corrections submitted July 2019.

# Contents

## Abbreviations

| | | |
|---|---|---|
| °C | - | degrees centigrade |
| ‘omic | - | combination of transcriptomic and proteomic data |
| 1° | - | primary |
| 2° | - | secondary |
| aa | - | amino acids |
| AAT | - | Amino acid transporters |
| Amp | - | Ampicillin |
| ANOVA | - | Analysis of variance |
| AP | - | Adaptor protein |
| APS | - | Ammonium persulphate |
| ATF | - | Activating transcription factor |
| ATP | - | Adenosine triphosphate |
| BFA | - | Brefeldin A |
| BHK | - | Baby hamster kidney |
| BiP | - | Immunoglobulin binding protein |
| BL1 | - | emission detection at 530 ± 30 nm |
| BLAST | - | Basic local alignment search tool |
| bp | - | base pair(s) |
| BSA | - | Bovine Serum Albumin |
| CATCHR | - | Complexes associated with tethering containing helical rods |
| CD | - | Chemically defined |
| cDNA | - | Complementary DNA |
| CerS2 | - | Ceramide synthase 2 |
| CERT | - | Ceramide transfer (protein) |
| CFPS | - | Cell free protein synthesis |
| CHO | - | Chinese hamster ovary |
| CMV | - | Cytomegalovirus |
| $CO_2$ | - | Carbon Dioxide |
| COG | - | Conserved oligomeric Golgi (complex) |
| COP | - | Coat protein complex |
| CRISPR | - | Clustered Regularly Interspaced Short Palindromic Repeats |
| Ct | - | Cycle time |
| CypB | - | Cyclophilin B |
| DAG | - | Di-acyl glycerol |
| DCJ | - | David James Laboratory group |
| di$H_2O$ | - | De-ionised water |
| DMSO | - | Di-Methyl Sulphoxide |
| DNA | - | Deoxyribonucleic Acid |
| DoE | - | Design of experiment |

| | | |
|---|---|---|
| DT | - | Doubling time |
| DTE | - | Difficult to Express |
| e125 | - | Erlenmeyer 125 mL flask |
| EB | - | Elution Buffer (10 mM Tris-Cl, pH 8.5) |
| eIF2α | - | Eukaryotic initiation factor 2α |
| ER | - | Endoplasmic Reticulum |
| ERAD | - | ER-associated protein degradation |
| ETE | - | Easy to express |
| ERES | - | ER exit sites |
| ERO1 | - | ER oxidoreductase |
| ERSE | - | ER stress element |
| EtOH | - | Ethanol |
| $F_{ab}$ | - | Antigen binding fragment (of antibody) |
| FBS | - | Fetal Bovine Serum |
| $F_c$ | - | Constant fragment (of antibody) |
| FDA | - | Food and drug administration |
| FKPM | - | Fragments per kb of transcript per million of mapped reads |
| FP | - | Fluorescence Polarisation |
| FSC | - | Forward scatter |
| G30 | - | 30 Generations (±2) growth |
| GASE | - | Golgi apparatus stress response element |
| GBF1 | - | Golgi-specific Brefeldin A resistance 1 |
| GCA | - | Golgicide A |
| GCLM | - | Glutamate-cysteine ligase modifier subunit |
| gDNA | - | Genomic DNA |
| Gen | - | Generation number |
| GFP | - | Green fluorescent protein |
| GGAs | - | Golgi-localising, γ-ear-containing, Arf-binding proteins |
| GO | - | Gene ontology |
| GRAS | - | Generally regarded as safe |
| GRP | - | Glucose-regulating protein |
| GS | - | Glutamine synthetase |
| GSH | - | Glutathione |
| HC | - | (Mab) Heavy chain |
| HCP | - | Host cell protein |
| HF | - | High Fidelity |
| HS | - | High sensitivity |
| HSP | - | Heat shock protein |
| HTS | - | High throughput screening |
| IF | - | Immunofluorescence (microscopy) |
| IgG | - | Immunoglobulin G monoclonal antibody |
| IRE1 | - | inositol-requiring enzyme 1 |
| IVCD | - | Integral of viable cell density |

| | | |
|---|---|---|
| Kan | - | Kanamycin |
| kDa | - | kilodalton |
| KEGG | - | Kyoto Encyclopedia of Genes and Genomes |
| LB | - | Luria Bertani (broth) |
| LC | - | (Mab) Light chain |
| LD | - | Lethal Dose |
| Mab | - | Monoclonal Antibody |
| MaxiPrep | - | Maxi DNA preparation kit (Plasmid *Plus*) |
| MCB | - | Master cell bank |
| MeOH | - | Methanol |
| MGEV | - | Multi-gene expression vector |
| MHRA | - | Medicines and Healthcare products Regulatory Agency |
| min(s) | - | minute(s) |
| mg | - | milligram |
| mM | - | milimolar |
| MMPs | - | Matrix metaloproteases |
| MSX | - | Methionine Sulphoxamine |
| MTC | - | Multisubunit tethering complex |
| Onco-KIT | - | Tyrosine receptor kinase |
| PAGE | - | Polyacrylamide gel electrophoresis |
| PB | - | PrestoBlue cell viability reagent |
| PBS | - | Phosphate buffered Saline |
| PC | - | Phosphatidylcholine |
| PCR | - | Polymerase Chain Reaction |
| PDI(A) | - | Protein disulphide isomerases |
| PEI | - | Polyethylenimine |
| PERK | - | PKR-like ER kinase |
| PFA | - | Paraformaldehyde |
| pg | - | picograms |
| PI | - | Phosphatidylinositol |
| PLG | - | Phase Lock Gel |
| PLL | - | Poly-L-Lysine (solution) |
| PM | - | Plasma membrane |
| PRKD1 | - | Protein Kinase D |
| pSILAC | - | pulsed stable isotope labeling of amino acids in cell culture |
| PTM | - | Post-translational modification |
| Qp | - | Specific productivity |
| RIPA | - | Radioimmunoprecipitation assay buffer |
| RCF | - | Relative Centrifugal Field |
| RFU | - | Relative Fluorescence Unit(s) |
| RL1 | - | emission detection at 660 ± 20 nm |
| RNA | - | Ribonucleic Acid |
| RNA-seq | - | RNA sequencing (transcriptomic data) |

| | | |
|---|---|---|
| rpm | - | Revolutions per minute |
| RT-PCR | - | Reverse transcription PCR |
| SAMY | - | Secreted a-amylase |
| SD | - | Standard deviation |
| SDS | - | sodium dodecyl sulphate |
| SEAP | - | Secreted alkaline phosphatase |
| SEM | - | Standard error of the mean. |
| SGMS1 | - | Sphingomyelin synthetase |
| SM | - | Sec1/Munc18 family of proteins |
| S/MAR | - | Scaffold/Matrix attachment region |
| SNAP | - | Soluble NSF attachment proteins |
| SRP | - | Signal recognition particle |
| SSC | - | Side scatter |
| STAR | - | Stabilising anti repressor element |
| Stx5 | - | Syntaxin 5 |
| SV40e | - | Simian vacuolating virus 40 (promoter) |
| $t$ | - | time (normally days) |
| T25 | - | 25 $cm^2$ tissue culture flasks |
| TAE | - | Tris, acetate, EDTA (buffer) |
| TALEN | - | Transcription activator-like effector nucleases |
| TCD | - | Total cell density |
| TEE | - | Transcriptional enhancing element |
| TEMED | - | Tetramethylethylenediamine |
| TFRE | - | Transcription factor regulatory/response elements |
| TGN | - | Trans-Golgi network |
| TI | - | Targeted integration |
| $T_m$ | - | Melting temperature |
| tPA | - | Tissue plasminogen activator |
| TRAPP | - | Transport protein particle |
| t-SNARE | - | target (membrane localised) SNARE |
| µ (mu) | - | Specific growth rate |
| UCOE | - | Ubiquitous chromatin opening element |
| µL | - | microlitre |
| µm | - | micron (micro meter) |
| µM | - | micromolar |
| UPR | - | Unfolded Protein Response |
| UTR | - | Untranslated region |
| V | - | Volts |
| VEGF | - | Vascular endothelial growth factor |
| VCD | - | Viable cell density |
| v-SNARE | - | Vesicle (membrane localised) SNARE |
| VSV-G | - | Vesicular-stomatis virus-G |
| WB | - | Western Blot(ting) |

| | | |
|---|---|---|
| WCB | - | Working cell bank |
| WCP | - | Whole Cell Protein |
| XBP1 | - | X-box binding protein |

## List of figures

## List of tables

# 1) Literature review and introduction: Biologic manufacture and the mammalian cell secretory pathway.

*This chapter presents a detailed overview of the role of the CHO cell as a mammalian cell factory in biotechnology, alongside an introduction to the secretory pathway within mammalian cells. It starts with an introduction to the biopharmaceutical industry and how biopharmaceuticals are produced. The CHO cell and the processes by which it is used to produce biologics is introduced, alongside an overview of previous engineering approaches that have been taken to enhance CHO productivity. Finally a review of the mammalian cell secretory pathways carried out to better inform engineering strategies used within the rest of the project.*

## 1.1. Introduction to the biopharmaceutical (biologic) industry

A biopharmaceutical (also known as a biological medical product and henceforth referred to as a biologic) is a form of therapeutic pharmaceutical drug manufactured or part-synthesised in a living organism. They are produced by biotechnological techniques such as recombinant DNA or hybridoma technology, with the ultimate aim of alleviating the symptoms of, or curing, a certain disease or condition (Rader 2008; Walsh 2002). In most cases biologics are proteins used to treat a wide array of diseases and conditions, especially cancers, inflammation-related diseases, metabolic and endocrine disorders. Recently gene therapies requiring high-quality nucleic acid have become more prevalent, with the DNA or RNA used also classed as a biologic (Ginn et al. 2018). To gain approval for human use, protein biologics need to be specific to their intended target, well characterised and of high purity. As such their manufacturing processes are highly regulated by bodies such as the US Food and Drug Administration (FDA) and the UK's Medicines and Healthcare products Regulatory Agency (MHRA).

### 1.1.1. The Biopharmaceutical market

Recent data shows that sales of biologics approved for human use totaled $140 billion in 2013, with 54 biologics approved for human use between 2010 and 2014. Biologic sales increased 18.2% in 2012 compared to 2011, an increase seven-times greater than the 2.5% increase seen across the entire pharmaceutical sector. The nine major protein biologic families comprise of monoclonal antibodies (Mabs); hormones; growth factors; fusion proteins; cytokines; anti-coagulants; recombinant vaccines and therapeutic enzymes (table 1-1). The most

prevalent protein biologic family are Mabs, with nine approved by the FDA in 2017 (de la Torre and Albericio 2018). Mabs accounted for 31% of worldwide biologic approvals in the four years to 2014 and made up 38% ($24.6 billion) of total US biologic sales in 2012 (Aggarwal 2014; Walsh 2014). By 2014, 47 Mab products had been approved by US and European regulators. By 2016/2017 this had increased to 74 Mab-based products making up 65.6% of total recombinant protein biologic sales which amounted to $106.9 billion, with combined worldwide sales expecting to reach $125 billion by 2020 (Ecker et al. 2015; Kesik-Brodacka 2018; Strohl 2018).

As the technology to discover disease markers improves there is greater need to find and design biologics to target them. The process by which new biologics can be tested and manufactured has improved greatly, expanding the pipeline of drug companies and increasing the number of biologics coming to market. However, it is now likely that the continued increase in biologic sales is reaching a peak, with the 2008 global recession and on-going austerity and financial pressures increasing pressure on the pricing of drugs in general (Aggarwal 2014). The patents of many biologics, including many 'blockbusters' (a drug garnering annual sales exceeding $1 billion) have recently - or will soon - expire in both the EU and US (Kesik-Brodacka 2018; Walsh 2014). To exploit this companies are developing biosimilars (effectively a molecularly similar copy of an already licensed biologic) so they are prepared with competing products for market as patents expire (Kirchhoff et al. 2017; Ledford 2018).

## 1.1.2. Monoclonal antibodies

Monoclonal antibodies (Mabs; also known as Immunoglobulins, and including Immunoglobulin G; IgG) are the highest selling class of biologic (Aggarwal 2014). This is due to their specificity to disease targets, their ability to be completely humanised (both resulting in reduced side effects and immunogenicity) and their ability to be conjugated to another therapeutic agent or radioisotope such that it can be better targeted or used for diagnostic purposes (Wang et al. 2007). Antibodies naturally produced within the body are proteins that bind to specific antigens (via molecular markers known as epitopes) and then neutralise them. Mabs are made by immune system lymphocytes that alter deepening upon the Mab produced. As such isolation of specific lymphocytes to produce specific antibodies is not feasible (Tansey and Catterall 1994).

### *1.1.2.1. Monoclonal antibody structure*

Due to the large number of antigens that can be present within a body, myriad antibodies are required to target them. Functional diversity within the antibody family is, however, based upon a conserved molecular backbone. A roughly Y-shaped molecule comprises of a constant region ($F_c$) forming the base of the Y,

with two variable domains that bind antigen forming the top of the Y (antigen binding fragment; $F_{ab}$). The $F_c$ region comprises of two heavy chain (HC) molecules whilst each of the two arms of the $F_{ab}$ domain are made up of one light chain (LC) molecule along with the upper section of one of the HC molecules that make up the constant region. The specificity of epitope binding is conferred by the structure of the top of the $F_{ab}$. The tips of the HC and LC combine to interact with a specific epitope depending upon its structure (figure 1-1). Whilst the $F_{ab}$ interacts with non-normal epitopes found on foreign bodies (e.g. bacteria) and malfunctioning cells (e.g. cancer cells) through the variable domains, the $F_c$ region interacts with effector cells and molecules (including other antibodies) to mediate an immune response.

The Mab LC and HC are each encoded by separate genes, the gene products forming an antibody molecule containing two HCs and two LCs due to glycosylation, disulphide bond formation and their heterogeneity. However some humanised HCs can fold and dimerise without the presence of LC, subverting the ER's quality control system. This process is dependent upon the HC variable domain (Stoyle et al. 2017). It is because of these complexities in Mab formation and modification that their synthesis outside of the body can not be performed in a chemical reaction, with a mammalian cell required to complete their complex folding processes. Despite their conserved structure antibody behaviour still varies greatly and has resulted in the development of functional and commercially viable Mab pharmaceuticals has not been straight forward (Wang et al. 2007).

Table 1-1. Overview of different biologic families (next page.)

*The nine main classes of biologics are shown with a description of representative diseases they are used to treat. One specific example of each class is highlighted and the total sales from 2012 for each class is shown to highlight the scale of the biopharmaceutical industry. Table is modified from Aggarwal 2014.*

**Table 1-1. Overview of different biologic families**

*(Modified from Aggarwal 2014).*

| Biologic category | Description | Market example; disease target | 2012 sales ($ billion) |
|---|---|---|---|
| **Monoclonal Antibodies (Mab)** | Antibodies normally made by immune cells. Binding to an epitope highlights a problem to the immune system, resulting in a response. | Herceptin® (Trastuzumab); Her2-positive breast cancers. | 24.6 |
| **Hormones** | Signaling molecules produced by glands and transported around the body by the circulatory system. | Lantus® (Insulin glargine); Diabetes management. | 16.1 |
| **Growth Factors** | Protein that stimulates cellular growth, healing, cell differentiation and cellular process regulation. | Neulasta® (Pegfilgrastim); White blood cell booster. | 8.1 |
| **Fusion Proteins** | A single protein made from two separate genes fused together – e.g. an effector molecule is targeted by fusion to a Mab Fc region. | Enbrel® (Etanercept); autoimmune disease. | 5.8 |
| **Cytokines** | Small proteins with role in cell signaling and importance in the immune system. | Avonex® (Interferon β1a); Multiple sclerosis. | 4.9 |
| **Therapeutic enzymes** | Bio-catalysts synthesised by living cells. Used to replace enzymes that are mutated or lacking in the patient. | Pulmozyme® (Dornase α); Cystic fibrosis. | 1.4 |
| **Blood factors** | Clotting factor proteins present in blood to prevent bleeding and promote healing. | Novoseven® (Factor VIIa); haemophilia | 1.2 |
| **Recombinant Vaccines** | Attenuated versions of disease-causing proteins used to 'train' the immune system, providing immunity. | Gardasil® (recombinant HPV vaccine); HPV. | 1.1 |
| **Anti-coagulants** | Opposite of Blood Factors – stop blood clotting and blocking organs/ blood vessels. | Coumadin® (Warfarin); Deep vein thrombosis. | 0.4 |

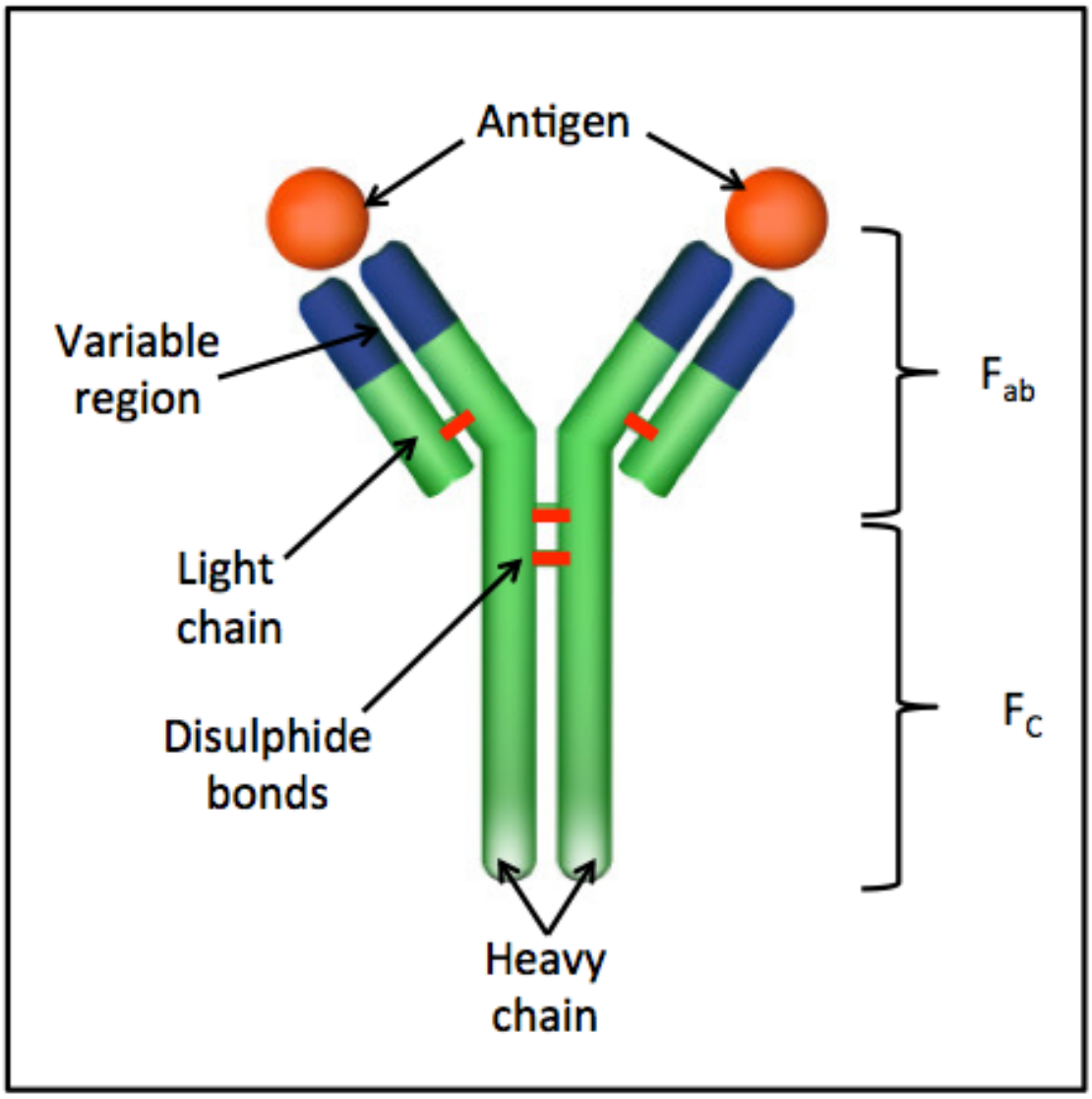


**Figure 1-1. Linear representation of IgG antibody**

*The Y-shaped linear structure of a Mab is made up of two HC joined together by two disulphide bonds. A LC joins to each HC by a single disulphide bond. The top of the molecule, comprising the N terminals of the HC and LCs, varies between antibodies and makes up the antigen binding domain. The bottom of the antibody remains more constant and activates immune response through interactions with immune system molecules and cells (pre-annotated image taken from http://www.dxdiscovery.com/monoclonal-antibodies.html).*

## 1.2 Biologic Manufacture

The development of a biologic follows a workflow of target identification; discovery of a molecule that acts upon it (e.g. by phage display, hybridoma discovery techniques); enhancement of the molecule to improve its efficacy before manufacturing and finally testing to see whether the molecule can be mass-produced at a satisfactory scale, cost and purity. With the complexities of biologics exceeding those that can be achieved by chemical synthesis, the biopharmaceutical industry utilises cellular organisms as factories with which to manufacture biologics. The gene encoding a biologic is inserted into the chosen cell by transformation (prokaryotes) or transfection (eukaryotes) and the cell machinery co-opted to manufacture the recombinant protein. Highly producing cells are selected and expanded to make cell banks that when grown in a large-scale industrial process produce high levels of recombinant product at consistent quality.

### 1.2.1. Non-mammalian manufacturing platforms

#### *1.2.1.1. Bacterial systems.*

Simple prokaryotes, particularly *Escherichia coli*, were initially used for production of recombinant proteins. They are well characterised; easy to genetically manipulate; easy to culture; fast growing (entire process lasting up to 48 h); able to produce high protein titres (up to 20 g/L) and therefore relatively in-expensive and quick as an expression system (Rosano and Ceccarelli 2014). However, prokaryotic systems have their limitations. They are unable to produce the complex folding and post-translational modifications (PTMs) required by many highly-specific biologics (Kesik-Brodacka 2018). Internal aggregation of recombinant protein, which bacteria can not fully secrete, can lead to the biologic product being caught in difficult-to-solubilise inclusion bodies, making downstream recovery of product labour intensive and expensive. Breaking the cell to retrieve the product produces contaminating exotoxins from the cell wall, requiring especially careful purification procedures (Mamat et al. 2015).

#### *1.2.1.2. Yeast systems.*

Utilisation of basic eukaryotes can overcome some of these issues. Yeast platforms such as *Saccharomyces cerevisiae* or *Pichia pastoris* (which are both recognised as 'Generally regarded as safe' [GRAS] by the FDA) produce recombinant proteins that are more accurate and functional than those from bacteria. Yeast contain more of the cellular machinery required to perform the basic folding and PTMs that are required for protein stability, function and efficacy. At 6-8 days in duration yeast processes are longer than bacterial processes, but recombinant protein production can be inducible, enhancing production levels up to 30 g/L, with secretion of recombinant proteins simplifying downstream processing (Celik and Calik 2012). Improved vector design; work that has allowed the secretion of normally non-secreted proteins; as well as host cell engineering and improvement

of glyco-engineering to better humanise glycosylation patterns, have led to yeast-produced biologics gaining regulatory approval for human and animal use (Ahmad et al. 2014).

#### *1.2.1.3. Plant-based systems.*

The improvement of production of transgenic plants has led to increased numbers of plant-produced biopharmaceuticals. A clinically-tested and FDA-approved anti-Ebola antibody and another anti-viral have both been produced in tobacco leaves. Human growth hormone was initially produced in tobacco and sunflower plants, the product having similar physical and antigenic properties to yeast- and human-derived products. The folding of complex proteins by plants is accurate and they also have the ability to introduce some PTMs. There is no danger of contamination with animal pathogens and initial set-up and plant husbandry costs are inexpensive. However large-scale production requires long production times and large amounts of land, whilst transgene expression cannot be consistently controlled (Lomonossoff and D'Aoust 2016; Rasala and Mayfield 2015; Whaley et al. 2011; Yao et al. 2015). Green microalgae have also been utilised more for recombinant protein production due to them being safe, scalable, easy to genetically manipulate and inexpensive to grow. Use of insect-based expression systems can produce products with many accurate PTMs, but not the desired complex glycosylation patterns required (Kesik-Brodacka 2018).

For some purposes and molecular targets, the lack of precise specificity of a biologic is not problematic and so production in microbial systems - both bacteria and yeast – are widely used for research and industry. However, as drugs have been developed to target more complex conditions, biologics are required to be more specific. This has resulted in recombinant proteins that are more difficult for the cell to produce to the required specificity and regulatory standards, especially with regards to protein folding and PTM. Microbial and basic eukaryote cells are unable to produce the high level of specificity required for most use within humans, with only mammalian cells being able to perform the precise and complex folding, PTMs and processing this new breed of biologic requires (Dalton and Barton 2014).

### 1.2.2. Mammalian manufacturing platforms

Production of recombinant proteins in mammalian cells has increased rapidly, contributing 60% of all biologics produced by 2014 (Walsh 2014; Wurm 2004). Biologic production in mammalian cells has been shown to be effective, though some engineering of the product is required for it to function correctly (Baik et al. 2015). Culturing of a mammalian cell line engineered to express a product allows production of highly specific and functional recombinant proteins which are secreted from the cell, making downstream processing relatively simple and similar to that required for yeast processes. However, mammalian cell culture is time consuming and labour intensive, with a doubling time of around 24 hours

and continuous sub-culturing or feeding strategies required to maximise culture growth. This results in processes 2-4 weeks in duration (De Jesus and Wurm 2011). *In vivo* mammalian cells normally exist as part of a tissue structure but to attain the cell densities required to make protein production worthwhile suspension growth is required. Cultured mammalian cells are much more susceptible to contamination and small changes in growth conditions than microbial cells. Changes in media composition and process control have increased mammalian cell growth, survival time and productivity. Genetic engineering of cell lines has enhanced the ability of cells to produce biologics to the required specificity, produce higher titres, quicker growth and better survival in suspension culture (Almo and Love 2014; Hacker et al. 2009). Various mammalian cell lines can be used for biologic manufacture, including mouse myeloma lines (NS0, Sp2/0), murine hybridomas, human cell lines (HEK293, HeLa, PER.C6) and baby hamster kidney (BHK) cells (Birch and Racher 2006). However, the Chinese Hamster Ovary cell has become the most prevalent mammalian cell factory utilised by industry for the large-scale production of complex biologics.

## 1.3. The Chinese Hamster Ovary cell: A mammalian cell factory for biologic manufacture

Chinese Hamster Ovary (CHO) cells are the preferred expression system for recombinant biologics. The CHO expression platform accounts for 63% of all mammalian production systems and 35.5% of total biologic production (Walsh 2014). Due to its role in industry CHO is well characterised, its genome having been sequenced and published (Lewis et al. 2013; Xu et al. 2011). However it is genetically plastic enough to be easily manipulated and transformed with DNA encoding recombinant proteins such as biologics. Effector genes can also be used to aid the amplification of the product gene, further enhancing biologic production levels. In optimised fed-batch manufacturing processes cell densities can reach 10-15 million cells $mL^{-1}$, with specific productivities of 50-60 pg/cell/day resulting in product yields of 1-5 g/L for Mabs and similar molecules (Wurm 2004). More recently antibody titres have reached 10-13 g/L (Huang et al. 2010). Since approval of the first mammalian-expressed biopharmaceutical – tissue plasminogen activator (tPA) in 1987 (Collen and Lijnen 2004; Collen et al. 1984), CHO cells have become the most commonly used mammalian expression system in biologic production.

The CHO cell line is an imortalised cell line first isolated in 1957 as part of a study to develop human and animal cell lines “hardy and reliable” enough to survive long-term cultivation, but with a low enough chromosome number (2n = 22) such that they could be used for genetic studies and be easily genetically manipulated

(Puck et al. 1958). There are four main CHO lines deriving from this original isolate: CHO-DXB, CHO-S, CHO-K1 and CHO DG44 (Wurm 2013). Initially grown as an adherent cell line under static conditions, suspension variants (SV) have been developed, allowing high cell densities to be attained (Wurm 2004).

## 1.3.1. Development of high-producing monoclonal CHO cell lines

For industrial biologic production to be economically viable recombinant protein needs to be produced in the order of g/L, whereas many laboratory-based systems can only reach mg/L titres. Many biological, molecular and process engineering changes have been made in recombinant protein production to enhance titre levels.

### *1.3.1.1 Biologic design and enhancement*

The development of a biologic progresses from identification of a problem and the molecular process that instigates it, through to discovery and design of a molecule that can target the problem. Careful design of the molecule, using both generic and more molecule-specific strategies, can improve production levels by reducing difficult-to-express regions or motifs that result in aggregation (Hussain et al. 2017). The optimised gene encoding the biologic is inserted into a plasmid DNA vector and placed under a strong promoter from which its transcription is activated. The highly transcriptionally active Human Cytomegalovirus (CMV) promoter is often used to drive transcription of the GOI, with the less active Simian vacuolating virus 40 early promoter (SV40e) often used to run the chosen selective marker gene on the plasmid. Changing the stoichiometry of Mab heavy and light chain genes can improve Mab folding kinetics, resulting in increased Mab production levels (Pybus et al. 2014). The inclusion of an accessory gene(s) on the same or separate vector can allow tweaking or enhancement of a specific cellular pathway important to product production (e.g. ER stress, folding, glycosylation, secretion), increasing product quality and yield from CHO cells (Florin et al. 2009; Johari et al. 2015; Le Fourn et al. 2014; Pybus et al. 2014; Tigges and Fussenegger 2006). Vector engineering has improved transcription rates of recombinant genes (Cacciatore et al. 2010). Recently, synthetic promoter design has improved with them becoming more prevalent in cell engineering, allowing a greater and wider control over transcription rates whilst decreasing the length of the promoter region compared to those from natural sources (Brown and James 2016; Brown et al. 2014). Utilisation of natural and synthetic regulatory elements for RNA processing (e.g. poly-A tail; 5' and 3' untranslated regions; introns and splice sites to encourage nuclear export) further enhance rates of gene transcription (Birch and Racher 2006).

### *1.3.1.2. Recombinant gene expression*

Plasmid vectors (containing the GOI, any required accessory genes, selectable marker(s) and any requisite untranslated control/regulatory elements) are introduced in to the cell by transfection. Transfection of CHO cells with plasmid

DNA can be either virally, chemically of physically mediated. Multiple genes can be transfected at once using multi-gene expression vectors (MGEVs) containing, for example, Mab HC, LC and selection marker. Transcription of each separate gene can be independently controlled by separate promoters.

The machinery by which viruses insert DNA into cells to replicate can be co-opted to insert and express recombinant genes by viral transduction (Muzyczka 1992). Use of polymers or lipids to facilitate introduction of DNA into the cell (using e.g. Polyethylenimine, PEI; Lipofectamine™) requires initial complexing of the DNA with the polymer. The complex is then taken up by the cell by endocytosis, the DNA eventually being delivered to the nucleus where it can be transcribed. Polymers (especially PEI) are generally inexpensive and have little impact upon cell growth, though transfection rates can be moderate (Mozley et al. 2014; Thompson et al. 2012). Electroporation physically introduces holes in to the CHO cell wall by passing an electric shock through a mixture containing the cells and DNA. The DNA is taken up into the cell cytoplasm through the holes before delivery to the nucleus where it is transcribed. Transfection rates are generally quite high (above 90%) with cell recovery improving as culture time extends, depending on the strength and duration of the charge (Teissie et al. 2005). Nucleofection is an improved electroporation method where DNA is better delivered to the cell nucleus (Maurisse et al. 2010).

CHO cells can be either transiently transfected (where plasmid DNA is transcribed whilst it persists within the cell population for 4-6 days), or stably transfected (where a plasmid integrates into the genome and is selected for). Transient transfection produces small amounts of protein so is ideal for the initial testing and screening of genes (both product and accessory) or characterisation studies. It is also scalable between different culture volumes (Abbott et al. 2015; Baldi et al. 2007). The development of long-term transient expression techniques has resulted in large-scale transient cultures producing up to 2 g/L antibody titres (Daramola et al. 2014; Derouazi et al. 2004). Stable transfection is used for long-term high-level product production. The stable integration of the recombinant DNA into the genome is maintained by the application of selective pressure to ensure all surviving cells contain the plasmid (Kraemer et al. 2010).

#### *1.3.1.3 Gene expression and amplification systems*

Producing CHO cell lines with high expression levels of good quality product is required to make them viable for the manufacture of biologics. This primarily requires a cell line with a high copy number of the gene of interest (GOI) stably integrated into the genome at a point with high activity. Amplification of the gene produces cell lines with high GOI activity levels, with two platforms - dihydrofolate reductase (DHFR) knockout and Glutamine synthetase (GS) knockout/reduction – predominantly used (Cacciatore et al. 2010; Durocher and Butler 2009).

#### *1.3.1.3.1. The DHFR system.*

DHFR is required for the synthesis of nucleic acid precursors. DHFR-negative CHO mutants such as CHO-DG44 require supplementation of growth media with glycine, hypoxanthine and thymidine to survive (Kraemer et al. 2010; Urlaub and Chasin 1980; Urlaub et al. 1983). Inclusion of a DHFR gene on to a vector encoding the GOI provides a section marker for cells expressing the GOI, as well as allowing for amplification of the gene to enhance product productivity levels. Treatment of transfected cells with the DHFR inhibitor methotrexate (MTX) will kill cells not expressing DHFR. With both the DHFR gene and GOI present on the same vector this ensures cells expressing the GOI persevering within the cell population. Gradually increasing the concentration of MTX will ensure only cells with high active copy numbers of DHFR and the GOI survive, resulting in an enriched population of cells with high productivity levels (Cacciatore et al. 2010; Kraemer et al. 2010). Attenuation of the DHFR selection marker can further enhance the selection and amplification power of this system (Chin et al. 2015).

#### *1.3.1.3.2. The GS system.*

The GS enzyme is essential in CHO cells for synthesis of the essential amino acid glutamine. The GS expression system was developed by Lonza and can utilise cells both deficient for and containing (albeit at low endogenous levels) GS (Bebbington et al. 1992; Cockett et al. 1990). The GOI is contained on a vector also containing an exogenous GS gene. Growth of transfected cells in glutamine-free media, coupled with treatment with the GS inhibitor methionine sulfoxamine (MSX), results in the death of cell containing only low levels of endogenous or exogenous GS. Only cells containing higher levels of active GS survive these conditions, resulting in all cells within the population expressing the GOI.

Gene amplification can be performed by increasing MSX levels, with high-producing cells then selected (Kraemer et al. 2010). The GS system has several advantages over the DHFR system, as it does not require a mutant host cell line, although the efficiency of cell line generation can be enhanced by knocking out endogenous GS (Fan et al. 2012). GS also facilitates a more rapid selection process, shortening the time required to develop a high producing cell line (Nakamura and Omasa 2015; Noh et al. 2013). Using a weak promoter (e.g. SV40e) to drive transcription of the selection gene (GS or DHFR) aids the amplification process as to survive cells will need to be more productive (Ng et al. 2007). Utilisation of a strong promoter (e.g. CMV) to drive GOI transcription helps maximize its expression level (Fan et al. 2013).

### *1.3.1.4. Cell line development and use of monoclonal cell lines.*

For higher and more consistent levels of production, including manufacture, stable transfection is required. After transfection cells are selected to ensure genes are being expressed, and amplification can be performed to enhance productivity.

Once recovered from the selection process the transfected pool is expanded. For reproducibility and regulatory purposes, all manufacturing cell lines are required to be monoclonal, that is derived from a single progenitor cell. Recombinant molecule production from monoclonal cells produces more reproducible manufacturing process and biologic product (ICH-Q5D 1998). Fluorescence-activated cell sorting (FACS), cell printing, microfluidic systems and limited dilution techniques can all be used to separate heterogeneous pool populations into single cell colonies. As well as fulfilling regulatory requirements, the single cell selection process allows for favourable phenotypes to be selected for and poor producers to be discarded, streamlining and accelerating the cell line development process as well as the overall production pipeline (Lindgren et al. 2009).

Single cell colonies are progressed and expanded static culture, high throughput (often automated) systems and continuously monitored for productivity and growth levels to ensure the best-performing cell line is taken forward for use in manufacturing processes (Browne and Al-Rubeai 2007; Lindgren et al. 2009; Rouiller et al. 2016). A pool of transfectants has a wide range of phenotypes determined, in part, by the site at which stable integration into the host cell genome has occurred.

#### *1.3.1.4.1. Improving recombinant gene activity in mammalian cell factories.*

Integration sites may be transcriptionally repressed due to DNA being packaged as heterochromatin. Insertion of a recombinant gene sequence could also interrupt essential genes, affecting cellular growth or processes. To reduce this heterogeneity, increase transcriptional activity and therefore product yields, sight directed/targeted integration (TI) can be used to introduce DNA at a specific genomic region with well defined expression properties (Baumann et al. 2017). Genetic elements that modify genomic chromatin can also be used to ensure the GOI is in a transcriptionally active conformation. This can be achieved by the addition to the transfected DNA of a Ubiquitous Chromatin Opening Element (UCOE); Scaffold/Matrix Attachment Region (S/MAR); Stabilising Anti Repressor (STAR) elements; or insulators that are linked to the GOI. These elements reduce or negate epigenetic processes that can impinge upon transgene expression by ensuring the region of genomic DNA the plasmid resides remains in an open and active confirmation and also limiting gene silencing by methylation. Use of these elements ensure tighter control over transcription levels and reduce the variation seen from random integration. This increases the number of high-producing cells in a transfected cell pool, reducing cell line development timelines. UCOEs have been shown to increase Mab productivity in CHO cells by more than 6-fold (Benton et al. 2002; Betts and Dickson 2015; Betts and Dickson 2016; Neville et al. 2017; Saunders et al. 2015).

Even with targeted integration, transfectant pools have a wide range of cellular phenotypes. It is desirable to identify a single-cell derived colony with desirable

growth and productivity characteristics as quickly as possible. This accelerates the development of a high-producing cell line, allowing product analysis, medical trials and eventually large-scale manufacture to occur in a shortened timeframe. However, behavior of a cell line early in the selection process does not necessarily indicate its behavior in the final production process (Porter et al. 2010). Improving high throughput screening (HTS) techniques at larger scales is starting to improve the cell selection process with regards to comparable performance across multiple production scales (Abbott et al. 2015; Rouiller et al. 2016). Once a high-performing clone has been selected CHO cells are still susceptible to phenotypic drift that can impact upon phenotype and productivity, potentially limiting the duration for which a cell line can be reliably used (Jadhav et al. 2013).

#### *1.3.1.5. Optimising CHO growth and manufacturing processes*

Improvement of CHO growth levels by media and process development was one of the initial methods used to enhance titres. Bioprocess optimisation has seen a 100-fold increase in CHO cell yields of recombinant proteins to between 1-5 g/L, with a specific cellular productivity (Qp) of 50 - 90 pg/cell/day (Hacker et al. 2009; Wurm 2004). A decrease in culture temperature during the stationary phase of cell growth can result in a coordinated cellular response in mammalian cells. A combined modulation of the cell cycle, translation, transcription and metabolism can reduce cell growth, allowing more cellular resources to be directed towards protein production. This then leads to a higher specific productivity and, if performed correctly, an increased overall titre. This is also seen in transiently-expressing cultures, with mild hypothermia also improving Mab glycosylation (Al-Fageeh et al. 2006; Sou et al. 2018; Torres et al. 2018; Yee et al. 2009).

For downstream purification and regulatory purposes, animal-derived products such as serum can not be used in manufacturing cell culture, despite their enhancing effects upon growth and productivity levels in a research environment. The development of CHO cell lines able to grow in serum-free shaken suspension culture, coupled with the optimisation of chemically defined (CD) media conducive to this, has allowed the removal of serum from the manufacturing process. This alleviates regulatory hurdles and boosts production levels when compared to basal medium (Almo and Love 2014; Kim and Lee 2009; Zhang et al. 2013).

The effect of process scale-up needs to be taken in to account during the cell line and process development process to ensure the final cell line chosen can perform during large scale production. Compared to small research-scale cell culture volumes, the large volumes of liquid culture required for manufacturing processes suffer more from poor mixing. This can lead to microenvironments within the bioreactor. Increasing mixing speeds to combat these results in higher shear forces which further impact upon culture growth, as well as increasing host cell

protein (HCP) levels in the growth medium which can pose a purification issue (Hogwood et al. 2014). Concentration gradients of gas, feed substrates and cell metabolites can develop, resulting in a heterogeneous mixture and exposure of cells to inhibitory environments which can result in a reduction in cell growth, productivity and overall yields (Wang et al. 2015a). Alteration of bioprocess parameters and scales can reduce product aggregation and increase CHO biologic production levels (Baik et al. 2015; Paul et al. 2018).

Addition of chemicals to cell culture can reduce the impact of harmful metabolites, with antioxidants able to mitigate the effect of reactive oxygen species, hydrogen peroxide and lactate to reduce oxidative stress on the culture and enhance productivity (Ha et al. 2018). Addition of cell cycle inhibitors or histone modifiers can alter the transcriptomic and proteomic profile of the cell, resulting in enhanced productivity (Du et al. 2015; Gatti et al. 2007; Yee et al. 2009). The production of HCPs during cell culture provides challenges for the downstream processing of a product. HCP contamination can occur due to constitutive secretion of host proteins, cell lysis upon apoptosis or cell bursting upon contact with bioreactor surfaces. Due to the number of host proteins a cell produces it is likely that a sub-section of HCPs will often purify alongside the product (Hogwood et al. 2014; Valente et al. 2015; Wang et al. 2009a; Yuk et al. 2015).

#### *1.3.1.6. The future of CHO as the major player in large-scale biologic manufacture*

Despite its wide use in biologic manufacture there remain some questions around the suitability of CHO as a widely-used manufacturing platform in the future. The requirement to produce more complex and difficult to express (DTE) biologics will require more stringent demands upon the CHO cell.

Expansion of the drug development pipeline from basic Mabs and globular proteins has resulted in the necessity to produce bispecifics, antibody-drug conjugates, radioimmunoglobulins, fusion proteins, antibody fragments and immunocytokines. Many of these will require complex and specific folding and glycosylation that CHO cells may struggle to perform (Spiess et al. 2015; Strohl 2018).

##### *1.3.1.6.1. Meeting future biologic demand: Production platform considerations.*

Mammalian systems have the capacity to produce the drug quantities currently required to target diseases such as cancers and immune deficiency in north American and western European markets. However as the demand for biologics increases both within and without of these areas, so will the pressure upon the current manufacturing system. The drug requirements of the developing world are similar to those in more developed countries, but for biologics to become more readily available in less wealthy countries the costs of manufacturing would need to be greatly reduced from current levels ($50 - $300/g) to $10/g or lower (Kelley 2009; Matthews et al. 2017). A combination of biologics being developed for

treatment of more chronic conditions (e.g. Alzheimer's, heart disease); the size of affected populations; and the continual treatment required, puts vast strain upon the worldwide CHO manufacturing process. For example, a 1 g/month dose of biologic per patient for a total patient population of one million would require 12 metric tons per annum, a value double the worldwide CHO manufacturing capacity in 2016 (Matthews et al. 2017).

Without rapid increases in mammalian cell productivity levels and capacity, (which may yet come about through ongoing optimisation and improvement of long-term perfusion manufacturing technologies) an alternative manufacturing platform capable of mimicking human post-translational modifications whilst producing the required level of materials would have to be found. Eukaryotic microorganisms such as yeast are used in industrial scale protein production but are not currently suitable for the high specificity required for many biologics, especially with regards protein folding and PTMs.

It is in the engineering out of non-mammalian PTMs, and engineering in of mammalian PTMs, that the recent advances in applicability of powerful gene-editing tools such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR), coupled with the simplicity and plasticity of yeast and bacterial genomes, could be used to produce yeast strains with the ability to produce high-quality biologics at a scale thus far only achieved by mammalian cells (Love et al. 2017; Mougiakos et al. 2016; Ryan and Cate 2014). Use of plant-based systems for biologic production is inexpensive compared to both mammalian and eukaryotic systems (0.1% and 2-10% of production costs respectively), though there are issues with regard to downstream processing and replacement of essential food crops (Yao et al. 2015).

#### *1.3.1.6.2. Meeting future biologic demand: Technology and equipment considerations.*

The widening of drug development pipelines will require increased throughput and speed with which to both optimise production process as well as produce small amounts of drugs for testing and trials. HTS processes are becoming increasingly prevalent to optimise cell line productivity (Arnall et al., manuscript in preparation; Mora et al. 2018). The development of Cell Free Protein Synthesis (CFPS) systems have thus far focused on prokaryotic cell lines, but recent advances in mammalian CFPS systems (in both CHO and HeLa cells) have allowed rapid, reproducible production of complex proteins at a scale of 500 - 600 mg/L in a 2-5 hour time frame (Martin et al. 2017; Tran et al. 2017). However further work is required to improve glycosylation and disulphide bond formation in CFPS systems and expanding these small-scale processes to a manufacturing-relevant scale seems somewhat distant.

With diminishing returns after more than 25 years of process optimisation, genetic engineering of the CHO cell has further increased biologic yields by improving productivity levels of the cell itself, improving yields and allowing the production of more complex biopharmaceuticals (Fischer et al. 2015). Engineering targets to improve yields have so far focused on enhancing transcriptional and translational control of the product gene (Brown and James 2016). Metabolic engineering and cell-cycle control has improved cell growth within culture (Fussenegger et al. 1998). Glyco-engineering has enhanced PTMs such that they closer match patterns found in humans (Chung et al. 2017). Aggregation of products has been reduced (Le Fourn et al. 2014), whilst engineering of the Endoplasmic Reticulum (ER) and the Unfolded Protein Response (UPR) has improved protein processing, reduced protein aggregation within the cell and reduced apoptosis whilst also increasing organelle size and therefore the productivity capacity of the cell (Rahimpour et al. 2013; Tigges and Fussenegger 2006). Engineering of the biosynthetic and secretory pathways has also enhanced cellular productivity (Becker et al. 2008; Hansen et al. 2017; Zhou et al. 2018).

## 1.4. The mammalian cell secretory pathway

The mammalian secretory pathway is made up of many discrete membrane-bound compartments comprising organelles, vesicles, secretory tubules, endosomes and lysosomes. Biological membranes are fluid and easily undergo topological changes due to their fluidity and quasi-2-dimensional structure, resulting in membrane scission and fusion events essential for intra- and extra-cellular trafficking as well as general cell maintenance (Knorr et al. 2017). Processing of both host and recombinant proteins takes place within the structures. Organelles and vesicles contain both resident proteins, which remain in the compartment and confer its biochemical function upon it, and transient 'cargo' proteins passing through the organelle *en route* to their final destination. There are two main mammalian transport pathways controlling transport in different directions within the cell. Retrograde transport moves cargos from their current location - be it the cell surface or a peripheral organelle - further inside the cell. Endocytosis initiates this pathway at the cell surface. Anterograde transport shuttles cargo proteins from within the ER towards the extremities of the cell via vesicles and the Golgi apparatus. In the case of proteins transported out of the cell this is called secretion or exocytosis.

There are two types of cellular secretion. **Constitutive secretion** comprises of the continuous delivery of cellular components to the plasma membrane and the exocytosis of extracellular factors. Meanwhile, specific cells such as neurons and endocrine/immune cells practice **regulated secretion** of specific molecular

modulators and effectors (e.g. neurotransmitter release into the synapse; hormone/antibody release in to the blood stream). The secretion of these molecules is not continuous, with cargo being stored at a high concentration in secretory granules with their release being activated and controlled by the presence and level of a signal such as a calcium pulse down an axon or change in blood sugar levels (Burgess and Kelly 1987).

The secretory pathway begins with transport of the nascent polypeptide chain from the ribosome into the ER lumen via the translocon, a process mediated by interaction between the protein's signal peptide and the ER's signal recognition particles which mediates docking of the ribosome at the translocon (Halic et al. 2006). Within the ER protein folding and initial PTMs takes place with the help of chaperones, foldases, bond-forming enzymes and glycosylation enzymes. In the Golgi complex further PTMs occur, including glycan modification and proteolytic cleavage. Transport through the Golgi stack is believed to be through maturation of the cisternae, though some vesicle transport does occur (Glick and Luini 2011; Glick and Nakano 2009), with transport to the trans-Golgi network (TGN) being vesicle mediated. Finally at the TGN proteins are sorted to their required destination, in the case of secreted proteins this being the plasma membrane (PM) for eventual release from the cell (Lee et al. 2004; Papanikou and Glick 2014).

Whilst intracellular protein transport and secretion occur in all cells to some level, most cell types are not secretory specific. In the case of secretory-specific cell types naïve cells mature into protein processing and secreting factories. For example, antibody-secreting plasma cells mature from naïve B-cells. B-cells are activated by the immune system, resulting in their terminal differentiation into plasma cells, the antibody producing and secreting factories of the immune system. This change in cell phenotype is brought about by a large increase in secretory pathway components resulting in organelle biogenesis, increasing ER and Golgi volumes (figure 1-2; Ribourtout and Zandecki 2015; Shaffer et al. 2004; Shapiro-Shelef and Calame 2005). Plasma cell protein production capacity has been calculated to reach approximately 100 pg of IgG monomers per cell per day, from cells a similar size (10-15 μm diameter) to CHO cells (Dinnis and James 2005).

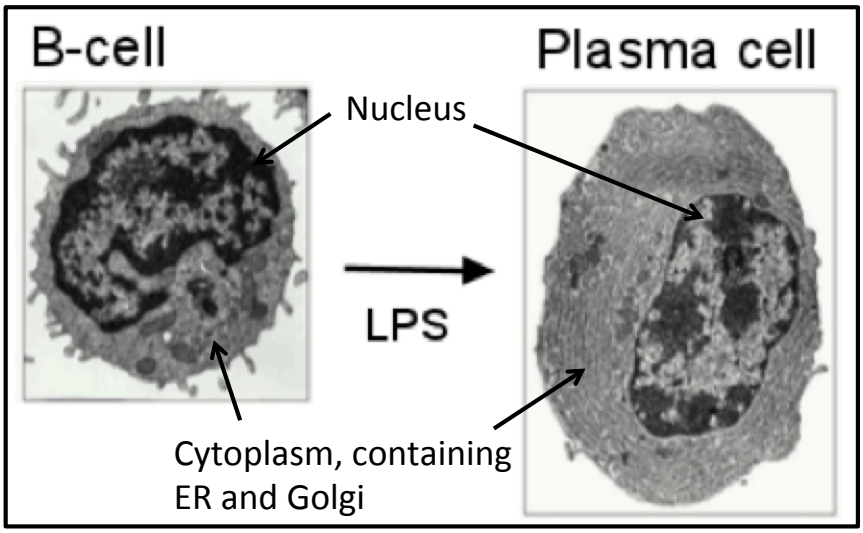


**Figure 1-2. Image of plasma cell morphology.**

*Activation of a B-cell by exposure to Lipopolysaccharide (LPS) results in terminal differentiation to a plasma cell that is specialised in the secretion of antibodies. The cell size expands slightly (to approximately 15 µm in diameter), but the main morphological difference is seen in the cytoplasm. Whereas the B-cell contains some ER and one normal-sized Golgi, in the plasma cell the cytoplasm is filled with long strands of rough ER and one large Golgi body. Image provided by A. Peden*

.

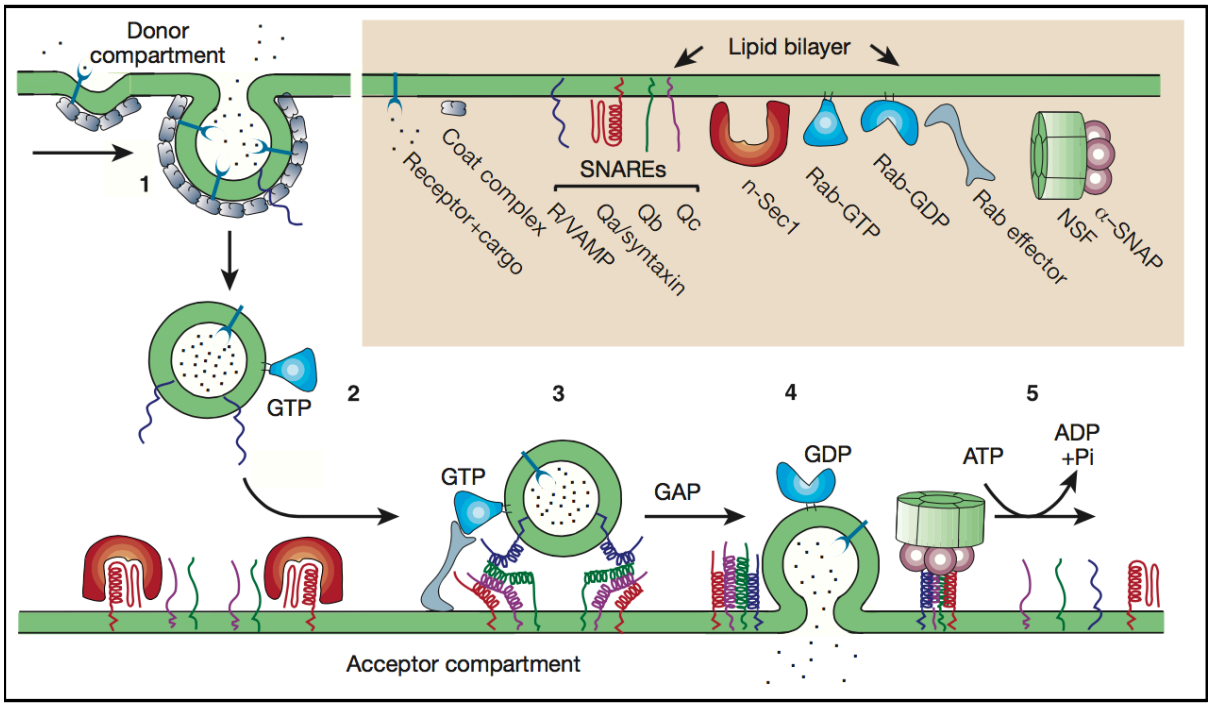


**Figure 1-3. An overview of the life-cycle of a transport vesicle.**

*[1] Interaction of cargo proteins with cargo receptors leads to coat protein recruitment, membrane bending and vesicle budding. Upon budding many coat proteins are shed (see section 1.4.1.1). [2] Interaction between vesicle and cytoskeleton components mediate transport towards target membrane (see section 1.4.1.3). [3] Tethering and docking of vesicles to target membranes provides docking and fusion specificity (see section 1.4.1.4). [4] Zippering and collapse of SNARE proteins instigates fusion of the vesicle to the target membrane and the release of cargo proteins (see section 1.4.1.5). [5] SNARE disassembly initiates recycling of vesicle components back to the donor membrane (see section 1.4.2). Figure taken from Bock et al. 2001*.

### 1.4.1. Overview of the mammalian secretory pathway: The life cycle of a vesicle

Organelles are bound by fluid lipid membranes, with alterations in environmental conditions resulting in topological changes modifying the geometry of the surface. This property allows the formation and scission of vesicular and tubular structures which are utilised in intracellular transport and exocytosis (Knorr et al. 2017). Despite the differences between different organelles and the proteins that mediate secretion to and from them, vesicular-mediated transport between two separate membrane-bound entities generally follows a set pattern.

The life cycle of a secretory vesicle starts with the recruitment and loading of the correct cargo and effector proteins, resulting in the formation of a vesicle which then buds from the parent organelle (Faini et al. 2013). A mixture of directed transport (utilising the cell's cytoskeleton) and undirected transport ferries a vesicle towards its target membrane (Brownhill et al. 2009). Here tethering and docking ensure correct vesicle targeting and bring the vesicle and target membranes into close enough proximity that fusion of the two occurs and the vesicle's contents are delivered (Bao et al. 2018; D'Agostino et al. 2017). An overview of a vesicle's life-cycle is shown in figure 1-3. Five main protein families regulate eukaryotic vesicular transport: Vesicle coat proteins; Rab GTPases; tethers; SNAREs and SNARE-regulators. All are conserved across phylogeny from yeast to man and across cell organelles involved in vesicle trafficking - ER, Golgi apparatus (containing *cis*- *medial*- and *trans*-Golgi regions), endosomes, lysosomes and the PM - albeit in slightly different forms (Bock et al. 2001). Despite cargo sorting at the donor organelle, bulk flow of donor-resident or incorrectly folded proteins can occur, though normally at low levels (Bethune and Wieland 2018).

#### *1.4.1.1. Cargo selection, vesicle formation and fission at the donor organelle*

Three distinct sets of cytosolic coat proteins complexes (COP) – COPI, COPII and Clathrin – function to form transport vesicles as well as selecting specific cargos via sorting signals on the cargo proteins to enrich them within these vesicles (Lee et al. 2004). COPII proteins are involved in ER-Golgi transport and COPI in intra-Golgi and Golgi-ER retrograde transport, whilst Clathrin is involved in post-Golgi and endocytic transport. All can transport proteins, lipids and small molecules, have a common ancestry and function in similar ways (Faini et al. 2013; Jensen and Schekman 2011; Owen et al. 2004). However the processes of COPI- and COPII-mediated vesicle transport are most relevant to protein secretion from mammalian cells. Rab GTPases (in their active GTP-bound state) are recruited to transport vesicles and interact with effector molecules to regulate a vesicle's life cycle (Fukuda 2008). Rab activity is controlled by Guanine Exchange Factors (GEFs) and GTPase activating proteins (GAPs) specific to each Rab-GTPase's function. A network of GTPases, their GEFs, GAPs, effectors, cascades and feedback loops results in the ability to change and control the activity of secretory pathway

components (Mizuno-Yamasaki et al. 2012). Vesicle formation commences with the activation of an adaptor protein (AP) by a GTPase. Activated APs recruit coat proteins (which themselves can act as an extra AP), cargo or cargo adaptors (which capture soluble cargoes from the ER lumen) and vesicle effector proteins (e.g. tethers, SNAREs) to the vesicle budding site (Cai et al. 2007a). Build-up of coat complexes leads to membrane bending which eventually results in the physical process of vesicle fission from the donor membrane (Bethune and Wieland 2018).

#### *1.4.1.1.1. Cargo selection and vesicle formation at the ER*

The COPII coat complex mediates vesicle budding from the ER membrane at ER exit sites (ERES,) with vesicle formation undergoing a stepwise assembly. The small GTPase Sar1 is activated by the GEF Sec12 (Barlowe and Schekman 1993), with Sar1-GTP interacting with the ER membrane and inducing membrane curvature (Long et al. 2010). Increased membrane curvature has been shown to increase Sar1-GTP's affinity for ER-membrane binding and GTPase activity, creating a positive-feedback loop linking Sar1 to membrane fission (Hanna et al. 2016). Membrane-bound Sar1 recruits the COPII inner coat proteins Sec23 and Sec24, which form a heterodimer. Whilst Sec23 interacts with Sar1-GTP, Sec24 recruits cargo proteins, cargo adaptors and vesicle effectors (e.g. SNAREs and tethers) to the ERES, forming a pre-budding complex that excludes ER resident proteins (Aridor 2018; Miller et al. 2002; Sato and Nakano 2005). The Sec16 protein, believed to have a scaffold role, has also been shown to mark ERES and is required for COPII vesicle formation, being recruited by Sar1 independently of Sec23/24 (Sprangers and Rabouille 2015; Watson et al. 2006).

There are many different cargo adaptor proteins (APs) that can be involved in vesicle formation, either travelling with their cargo or merely facilitating packaging into vesicles, with APs only interacting with fully-mature cargo proteins (Barlowe and Helenius 2016). Some specific APs can expand the COPII vesicle beyond its normal size (50 - 100 nm) so that large proteins such as collagen can be accommodated and transported (Gorur et al. 2017; Saito et al. 2017; Saito et al. 2014).

The COPII complex of outer coat proteins, the Sec13-Sec31 dimer, binds the pre-budding complex, further bending the ER membrane, promoting vesicle fission and releasing it from the membrane and into the cytoplasm (Barlowe et al. 1994). Other proteins such as TANGO1 interact with COPII machinery around the vesicle neck, with larger vesicles formed by recruitment of extra membrane regions and not extra coat proteins (Raote et al. 2017). The presence of higher levels of Sec13-31 increases the hydrolysis of Sar1-GTP to inactive Sar1-GDP, leading to the shedding of COPII coat proteins, leaving a membrane-bound vesicle in the cytoplasm (Bi et al. 2007; Bonifacino and Glick 2004; Faini et al. 2013; Jensen and Schekman 2011; Saito et al. 2017). Not all coat proteins are necessarily shed as

they have been shown to have a role in vesicle targeting and tethering (Cai et al. 2007b). Overexpression of secretory cargo induces the UPR, upregulating COPII and leading to the formation of more ERES, showing that ER stress response has an impact upon vesicle formation events. (Farhan et al. 2008).

#### *1.4.1.1.2. Cargo selection and vesicle formation within the Golgi*

The COPI coat complex mediates vesicle formation and budding at the Golgi, instigating both intra-Golgi and Golgi-ER retrograde transportation. The latter process is involved in the recycling of COPII vesicle components to the ER (Nelson et al. 1998). COPI vesicle formation for retrograde transport is instigated by recruitment of activated AP Arf1 GTPase to the Golgi membrane. Activation of Arf1 is carried out by a range of Arf-GEFs, which one depending upon the position within the Golgi stack. The Arf-GEF Golgi-specific Brefeldin A resistance 1 (GBF1) is responsible for regulation of Arf1 activation at the ER-Golgi interface (Garcia-Mata et al. 2003). Activated Arf1-GTP recruits a pre-formed COPI coatomer complex containing seven subunits to the Golgi membrane (Harakuge et al. 1994; Yu et al. 2012). Sorting signals on cargo proteins mediate their packaging in to COPI vesicles through interaction with COP subunits and APs (Jackson et al. 2010; Yu et al. 2012). Coatomer attachment to the membrane leads to coatomer polymerisation, increased membrane curvature and eventually scission and vesicle release, which may also be mediated by Arf1 dimerisation upon membrane bending (Arakel and Schwappach 2018; Beck et al. 2009). Hydrolysis of Arf1-GTP leads to incomplete shedding of coatomer complexes from the vesicle, with those remaining aiding vesicle targeting and tethering (Arakel and Schwappach 2018; Bethune and Wieland 2018; Faini et al. 2013).

#### *1.4.1.1.2. Cargo loading and vesicle formation at the TGN*

The Golgi is polarised, allowing transport through the stacks that make up the organelle to be directed (Allan et al. 2002; Boncompain and Perez 2013). Cargo eventually reaches the trans-Golgi network (TGN) where the endocytic and exocytic pathways converge in the sorting of lipids and proteins for delivery to their divergent final destinations, including endosomes, lysosomes, the PM, or back through the Golgi stack. For accurate delivery, cargo has to be packaged into transport carriers associated with docking and transport machinery specific to their destination (De Matteis and Luini 2008). Cargo proteins contain specific signals (e.g. tyrosine based; di-Leucine motifs) that are recognised at the TGN by APs that also recruit requisite vesicle components (Bethune and Wieland 2018; Owen et al. 2004; Traub 2005). Cargo proteins have been shown to exit the TGN in large tubules (between 0.3 - 1.7 μm in length, much larger than 50 - 100 nm diameter of coated vesicles) studded with kinesin motor proteins that pull specific regions of TGN membrane to form the transport carrier (Hirschberg et al. 1998; Polishchuk et al. 2003). Protein kinase D (PRKD1) regulates fission of PM-targeted transport carriers and is recruited to the TGN membrane by binding diacylglycerol (DAG), a conical fatty acid that is present in subdomains on the TGN

membrane and is produced from ceramide (Bard and Malhotra 2006; Baron and Malhotra 2002; Shemesh et al. 2003). At these DAG-concentrated sites the PRKD1-catalysed phosphorylation of ceramide transfer protein (CERT) and lipid kinases mediates the lipid make-up of the membrane region, with high DAG concentrations leading to bending of the membrane away from the lumen, facilitating transport carrier formation and budding (Bard and Malhotra 2006; Fugmann et al. 2007; Graham and Burd 2011; Hausser et al. 2005). However some transport carrier formation may be induced by the presence of a specific cargo (Yeaman et al. 2004). Transport from the TGN to the PM can also take other forms, for example via the endosome and its recycling pathway to the PM; Clathrin-related APs; transport of ubiquitylated proteins by Golgi-localising, γ-ear-containing, Arf-binding proteins (GGAs); or production of secretory granules in signal-activated secretion that mature during transport to the PM (Spang 2015).

#### *1.4.1.2. The ER-Golgi intermediate compartment (ERGIC)*

Upon budding from the ER, homotypic fusion of COPII vesicle contents leads to the formation of an ERGIC (Appenzeller et al. 1999; Lee et al. 2004). The ERGIC is a stable compartment close to but distinct from the ER to which cargo from ERES is initially shuttled by COPII vesicles for further sorting (by both the signal sequences within the cargo, and the difference in pH between the ERGIC and ER). At the ERGIC concentration and quality control of cargo is performed before subsequent transport to the Golgi via a second vesicular transport step involving COPI (Appenzeller-Herzog and Hauri 2006; Baines and Zhang 2007; Ben-Tekaya et al. 2005). COPI components localise to the ERGIC and are required for its formation (Lavoie et al. 1999; Oprins et al. 1993). Efficient transport of several globular cargo proteins from ERES to the ERGIC requires a cargo-receptor complex formed by the ERGIC53 (aka LMAN1) and MCFD2 proteins that assist ER export and delivery to the ERGIC before LMAN1 is recycled back to the ER (Appenzeller et al. 1999; Nyfeler et al. 2008; Nyfeler et al. 2006; Zhang et al. 2005). Vesicle fusion leading to ERGIC formation is also mediated by vesicle-to-vesicle interactions (Hay et al. 1998), with a specific Stx5 isoform discerning between vesicle fusion at the ERGIC and the *cis*-Golgi before being recycled to the ER (Hui et al. 1997).

#### *1.4.1.3. Inter-organelle vesicle transport*

Vesicle transit between organelles relies upon the cell's cytoskeleton which comprises of actin and microtubule filaments and their associated molecular motors. Vesicles interact with motor proteins such as the microtubule-associated kinesin and dynein (which are directed to microtubules' + and – ends respectively) and actin-associated myosin. Attachment of vesicles to motor proteins is likely mediated by specific adaptor proteins (e.g. kinectin with kinesin; dynactin with dynein), with Rab GTPases also orchestrating motor protein complex recruitment to vesicle membranes (Fokin et al. 2014; Horgan and McCaffrey 2011; Matanis et al. 2002; Short et al. 2002). In a stimulated exocytosis

model in human cells, GFP-tagged cargos displayed both long-rage bi-directional movement dependent on microtubules and kinesin, as well as short-range diffusive-like movements mediated by actin and myosin, suggesting rapid cargo delivery to the correct area of a cell is directed along microtubule tracks, before actin-mediated transport to the vesicle's final destination (Manneville et al. 2003). Vesicle transport can also be directed in a manner independent of motor proteins, with actin polymerisation able to direct vesicles upon their interaction with transport complex surface proteins (Khaitlina 2014). Cytoskeletal components also help maintain organelle structure, lumenal conditions and cellular position (Lazaro-Dieguez et al. 2006), whilst the Golgi is known to be a microtubule nucleation site (Brownhill et al. 2009).

Cargo leaving ERES has been shown to concentrate in pre-Golgi intermediates before moving towards the Golgi complex along microtubules using the microtubule motor complex of dynein/dynactin, disruption of which arrests transport of a model cargo protein (Presley et al. 1997). The COPII Sec23-24 subunit has been shown to directly interact with a dynactin subunit, suggesting this may be a functional mode of tethering cargo to a motor protein (Watson et al. 2005). This is however dependent upon how much of the COPII coat is shed from vesicles upon their budding (Brownhill et al. 2009; Cai et al. 2007b). Tethering of vesicles to motor proteins allows the transport of vesicles along the polarised elements of the cell's cytoskeleton, with motor-protein adaptors also believed to have a role in motor protein activation (Allan et al. 2002; Brownhill et al. 2009; Kamal and Goldstein 2002).

Due to the short distances involved, transport of COPI vesicles between Golgi cisternae is likely diffusion mediated. The actin cytoskeleton is likely to mediate retrograde traffic of COPI vesicles from the Golgi to the ER, with many actin-binding proteins localised to the Golgi and activated in response to the Arf-GTPase activation that initiates vesicle formation and budding. Meanwhile Rho GTPases are believed to modulate actin dynamics to control Golgi-ER transport (Fucini et al. 2000; Wu et al. 2000), although this actin assembly could be involved in vesicle scission and not transport (Fucini et al. 2002). In yeast, transport vehicles budding at the TGN interact with Myosin via a Rab GTPase and exocyst subunit to direct transport towards the PM (Jin et al. 2011). Lipid, ion and small molecule transport between organelles can be mediated by inter-organelle membrane contact sites where close tethering (but not full fusion) between the ER and PM, endosomes and mitochondria occurs (Henne et al. 2015; Phillips and Voeltz 2016).

#### *1.4.1.4. Vesicle tethering at the target organelle*

Once transported to its target destination a transport vesicle interacts with the target membrane, eventually releasing its cargo. Membrane-associated tethering factors present on both the vesicle and target membranes initiate the process that

eventually leads to vesicle fusion and cargo release, mediating a loose interaction between the vesicle and its target membrane, bringing them in to close proximity and conferring targeting specificity (Sztul and Lupashin 2006; Sztul and Lupashin 2009). This allows transmembrane soluble *N*-ethylmaleide-sensitive fusion protein attachment protein receptors (SNAREs) present on both the vesicle and target membranes to interact. SNARE fusion is controlled by regulatory proteins (the Sec1/Munc18 [SM] family). The fusion process lowers the energy barrier that must be overcome for membranes to fuse, resulting in fusion of the vesicle to the target membrane and delivery of vesicle cargo via a pore that opens up between the two membranes (D'Agostino et al. 2017; Gerst 2003). Much of the work on vesicle tethering mechanisms has been carried out in yeast model systems, though many components are conserved in mammalian cells. The role of tethers has recently been shown to go beyond formation of a lose membrane interaction, with them shown to play an essential mechanical role in the later stages of vesicle fusion as well as Golgi structure maintenance (Brown et al. 2011; D'Agostino et al. 2017; Spang 2017).

Vesicle tethering confers some level of specificity in their docking at the target membrane. For example, the exocyst tether complex is located only at areas where vesicle fusion occurs (Kee et al. 1997). Different Golgin tethers, with varying roles across both retrograde and anterograde transports, are located to different parts of the Golgi. They have also been shown to convey specificity towards vesicles originating from different locations (Munro 2011; Wong and Munro 2014). The two types of tethering complex (coiled-coil Golgins; large multi-subunit tethering complexes [MTC]) both interact with SNAREs, membrane lipids, coat proteins and small GTPases in a mechanism for vesicle tethering that is conserved between family members (Sztul and Lupashin 2006). An overview of Golgin and MTC tethering mechanisms is shown in figure 1-4 a. Tether recruitment to secretory bodies is mediated by Rabs, with at least 15 different Rab GTPases present on the Golgi (Cheung and Pfeffer 2016). For example, Rab6 recruits p230/Golgin245 to Tumour Necrosis Factor transport bodies during transport to the PM in HeLa and macrophage cells (Lieu et al. 2008; Micaroni et al. 2013).

Coiled-coil tethers (also referred to as Golgins) are long, rod-like molecules containing α-helices that often form dimers. They extend a significant distance (theoretically between 100 - 600 nm) from the Golgi, allowing vesicle capture at range and forming fibrous bridges between inbound vesicles, Golgi membranes and cisternae (Cheung and Pfeffer 2016; Munro 2011; Sztul and Lupashin 2006; Witkos and Lowe 2015). In general, a tether's N terminus "catches" a vesicle through interaction with specific motifs, coat subunits or vesicle-located tethers whilst being anchored – either through peripheral interactions with an adaptor protein or via a membrane-integrated tail anchor - to the target membrane at its C-terminus (Brown et al. 2011; Cheung and Pfeffer 2016; Gillingham and Munro 2016; Guo et al. 2008). The Golgin contains several Rab-GTPase binding sites

which function in drawing the vesicle closer to, and its stabilisation on, the target membrane, allowing for interaction with fusion machinery (Witkos and Lowe 2017).

Multisubunit tethering complexes are compartment-specific tether complexes that promote specific vesicle fusion at the correct target destination, although they function over shorter distances (up to 30 nm) than coiled-coil tethers. MTCs contain between three and ten subunits which can be classed as 'complexes associated with tethering containing helical rods" (CATCHR; containing long rods of stacked helical bundles), or non-CATCHR complexes which have a more diverse subunit composition. Despite this structural difference, all MTCs share similar protein interactions and therefore a common functional mechanism (Dubuke and Munson 2016; Hong and Lev 2014). The transport protein particle (TRAPP) complexes are multi-subunit complexes with the least similarity to other MTCs. Some subunits having Rab-GEF activity and all TRAPP genes are essential for human cell survival. (Cai et al. 2008; Kim et al. 2016).

#### *1.4.1.4.1. Vesicle tethering at the ER-Golgi interface.*

The majority of vesicle tethering in ER-to-Golgi and intra-Golgi transport pathways is mediated by *cis*-Golgi Golgins that interact with components on in-coming COPII vesicles from the ER. For example the *cis*-Golgi localised GM130 interacts with a C-terminal acidic patch on the COPII vesicle-localised coiled-coil tether P115. P115 is well defined (especially in yeast as its homologue Uso1) as essential for ER-Golgi transport and Golgi structure (Alvarez et al. 2001; Barr et al. 1998; Grabski et al. 2012; Puthenveedu and Linstedt 2001; Sztul and Lupashin 2006). The tail of P115 contains four coiled-coils separated by proline-rich hinges, allowing flexibility for an accordion-like collapse towards the target membrane (Sapperstein et al. 1995), whilst interaction with SNAREs is mediated by the two proteins' similar coiled-coil motifs (Wang et al. 2015b).

Other than GM130-P115 interactions, CASP is a Golgi-localised transmembrane protein related to the Golgin giantin that interacts with SNAREs involved in ER-Golgi transport (Gillingham et al. 2002); GCP60, GMAP210 and Golgin-45 are *cis*-Golgi located Golgins involved in ER-Golgi transport; and Golgin-84 likely plays a role in tethering and Golgi structure (Sztul and Lupashin 2006). ER-Golgi transport is also mediated by the pseudo-MTC 7-subunit TRAPPI complex. The Bet3 subunit of TRAPPI specifically binds the Sec23 COPII coat subunit, ensuring specificity of vesicle capture and marking vesicles for fusion to each other or with the *cis*-Golgi (Cai et al. 2007b). Meanwhile, the three-subunit DSL1 CATCHR MTC is ER-localised and required for COPI-vesicle recognition and tethering at the ER in Golgi-ER retrograde transport. DSL1 interacts with the COPI coat and target membrane SNAREs but not GTPases, unlike other MTCs (Dubuke and Munson 2016). An overview of the mechanism by which coiled-coil tethers interact with vesicles is shown in figure 1-4 b.

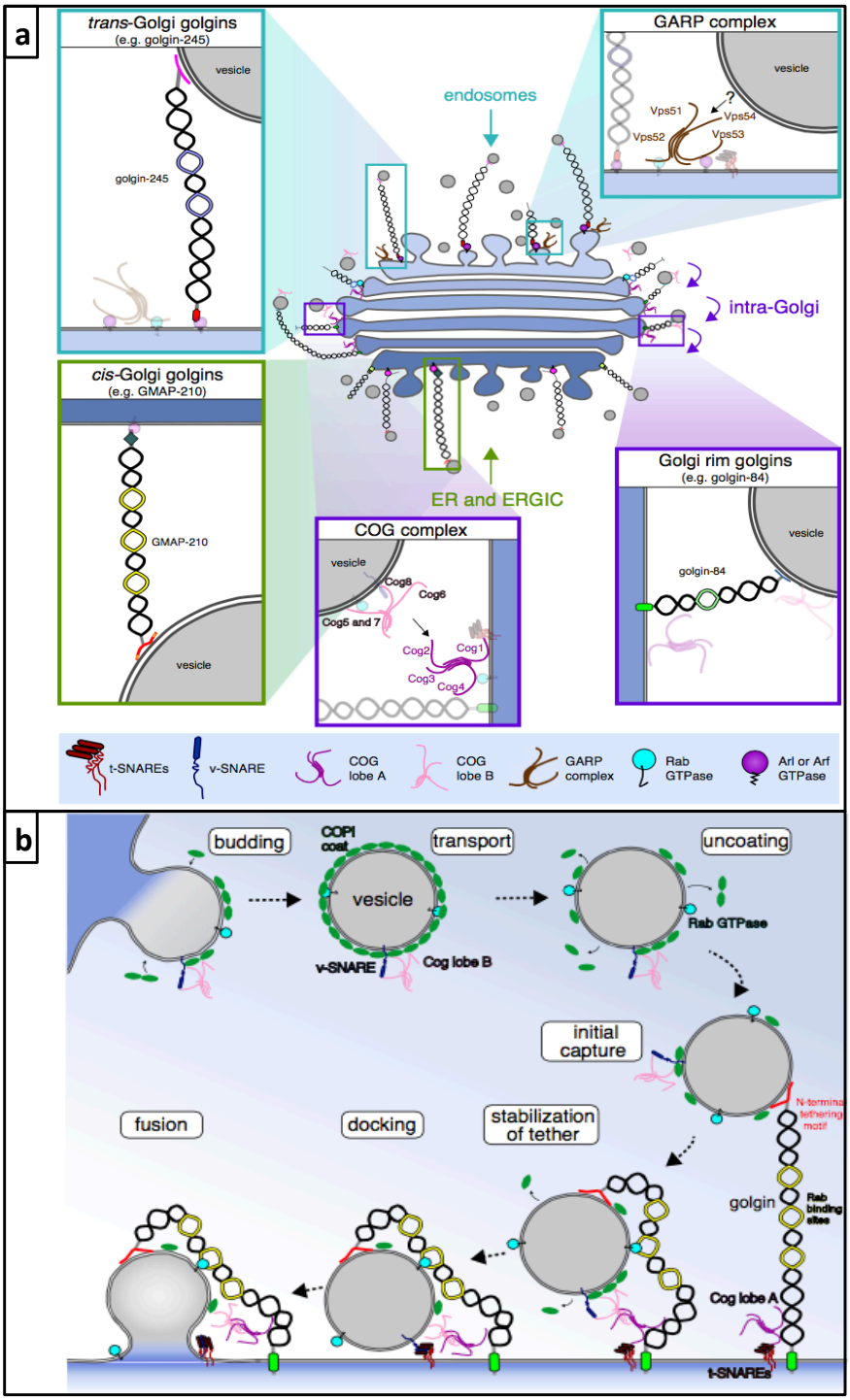


**Figure 1-4. Mode of action of Golgi tether complexes in vesicle recognition.**

*[a] Different coiled-coil and MTC tethers combine to provide specific tethering sites for Golgi-targeted transport vesicles. Coiled-coil Golgins (e.g. golgin-84, -245 and GMAP-210) project from the Golgi membrane, interacting with incoming transport vesicles (from ER/ERGIC, endosomes and, at exposed cisternae rims, within the Golgi itself) via N-terminal tethering motifs. The MTC COG complex mediates intra-Golgi transport alongside some golgins. One COG subdomain is present on Golgi membranes and the other on incoming vesicles. Interaction between the two forms a complete tethering COG complex, closely tethering the vesicle to the target membrane. [b] Mechanism of Golgin-mediated vesicle tethering. Tether N-terminal interacts with vesicle membrane. Collapse of tether mediated by proline-rich 'hinge' regions brings vesicle into close proximity with target membrane, allowing SNARE interaction and vesicle fusion. Both figures from Witkos and Lowe, 2017.*

#### *1.4.1.4.2. Vesicle tethering in intra-Golgi transport*

Transport through the Golgi is mainly by cisternal maturation, in which *cis*-cisternae mature into medial then eventually *trans*-cisternae. During this process COPI-vesicle mediated anterograde transport maintains correct enzymatic compartment contents (Spang 2015). Whilst COPI-mediated anterograde transport of smaller cargo does occur evidence for it is less solid than that for COPI-mediated retrograde transport (Glick and Luini 2011; Glick and Nakano 2009; Papanikou and Glick 2014). There is also evidence of a 'kiss-and-run' model of transport through the Golgi, where temporary direct fusion between cisternae provides a bridge across which transport of molecules can occur, without complete absorption of the vesicle membrane into the donor membrane (Mironov and Beznoussenko 2012).

As intra-Golgi transport vesicles bud in close proximity to their target membrane (and closer than the 100 - 600 nm reach of a Golgin), long-range vesicle tethering is not necessarily required (Witkos and Lowe 2015), though the closer tethering mediated by MTCs is utilised. A subunit of the TRAPPII complex (made up of the seven TRAPPI subunits augmented with three extra subunits) has been shown to bind a COPI coat adaptor protein, mediating the tethering of COPI vesicles in the early Golgi. However this may suggest a further role in ER-Golgi transport (Yamasaki et al. 2009). The TRAPP complexes are able to discern between COPI and COPII coats, providing tethering specificity (Barrowman et al. 2010).

The Conserved Oligomeric Golgi (COG) complex is an eight subunit CATCHR complex, separated into two modules that interacts with small GTPases on vesicle and target membranes. It is necessary for retrograde transport through the Golgi compartments (Dubuke and Munson 2016). Some Golgins are present at the peripheral buds of the intra-Golgi cisternae, but most present within the Golgi stack function in the maintenance of Golgi structure, with the Golgins GRASP55 and GRASP65 specifying in the *cis* - *medial* Golgi stacking (Sztul and Lupashin 2006). Most Golgins present on the TGN (e.g. Golgin-245, -97, GCC185 and GCC88) function in capture of endosome- derived and endocytic vesicles and so are not involved in secretion (Cheung and Pfeffer 2016; Sztul and Lupashin 2006).

#### *1.4.1.4.3. Vesicle tethering at the plasma membrane*

The exocyst is a CATCHR MTC that recognises and tethers secretory vesicles at the PM, being present at specific vesicle fusion and exocytosis sites (Kee et al. 1997). However direct evidence for tethering is still not available despite the exocyst being the most extensively studied tether system (Wu and Guo 2015). The exocyst's Sec8 subunit is transported to the PM on vesicles, residing there until fusion occurs, showing a separation between the tethering and SNARE-mediated steps of vesicle fusion (Rivera-Molina and Toomre 2013). Like the COG complex, the exocyst is made up of eight subunits in two separate modules that interacts with lipids and small GTPases (Rab, Rho and Ral families) on vesicle and plasma

membranes, as well as with SM proteins (Dubuke and Munson 2016; Hong and Lev 2014; Moore et al. 2007). The Sec3 and Exo70 subunits interact with the PM's inner leaflet whilst Sec15 interacts with the secretory vesicle via the Sec4 Rab-GTPase. Remaining subunits are able to form myriad interactions with each other and build up around these markers to create the tether complex, a process initiated by kinase-mediated phosphorylation of the membrane-marking subunit (Liu and Guo 2012; Sztul and Lupashin 2006; Wang et al. 2004; Wu and Guo 2015).

#### *1.4.1.5. Vesicle fusion at the target organelle*

The SNAREs are a protein family containing a conserved SNARE motif between 60-70 residues in length. A C-terminal transmembrane domain associates with the membrane, whilst a protruding N-terminus interacts with vesicle-mounted SNAREs (Wang et al. 2017). The tethering process allows SNARE proteins on vesicle and target membranes to come into close proximity (Dubuke and Munson 2016). Three target (t)-SNARES (also known as Q-SNARES as they contain a specific glutamine residue) are present on the target membrane, with 1 vesicle (v)-SNARE (or R-SNARE, containing a specific arginine residue) present on the vesicle membrane. Both tethers and v-SNARES are sorted to vesicles alongside cargo via interactions with coat adaptor proteins (Arakel and Schwappach 2018; Peden et al. 2001). Singular SNAREs exist in a *cis* conformation. Upon both SNARE types coming into close proximity they assemble into a complete *trans*-SNARE complex containing two/three t-SNAREs and one v-SNARE (Sollner et al. 1993). The helical regions of the complex 'zip' together from the distal N-terminal ends to the membrane-bound C-terminus, bringing the two opposing membranes together (Fasshauer 2003; Pobbati et al. 2006). This overcomes repulsive forces between the two membranes and allows the merging of the apposing bilayers, eventually leading to the release of vesicle contents (Dubuke and Munson 2016; Wang et al. 2017). An overview of SNARE-mediated vesicle fusion is shown in figure 1-5.

SNARE interactions are mediated by cytosolic SM proteins, with Munc18-2 and -3 widely distributed across many tissue types (Tellam et al. 1995), whilst Munc18-1 is only found in neuronal cells (Latham and Meunier 2007). Yeast studies have also shown Sec1 positively regulates formation of SNARE complexes in Golgi to PM transport (Gerst 2003). Ectopic overexpression of SM proteins in CHO cells improves SNARE-mediated vesicle fusion at both the Golgi and PM, increasing cell secretory capacity (Peng and Fussenegger 2009b). Tethering complexes are also able to regulate SNARE complex formation (Bock et al. 2001; D'Agostino et al. 2017; Gerst 2003; Hong and Lev 2014; Malsam and Soellner 2011). Once membrane fusion and cargo delivery has occurred SNARE complexes are broken down, a process mediated by proteins such as NSF, Tomosyn and Amisyn. These bind SNAREs, displace the SM proteins and interrupt the *trans*-SNARE complex, allowing components to be recycled (Baker and Hughson 2016; Gerst 2003).

Expansion of the fusion pore can also induce conversion of *trans*-SNARE complexes to a cis formation, aiding SNARE disassembly (Suedhof 2013).

Whilst exocyst tethering components are restricted to vesicle-docking areas of the PM, SNAREs are broadly distributed across target membranes (Kee et al. 1997). The transportation of syntaxins to the PM is mediated by their transmembrane domain, with SM proteins holding them in an inactive conformation during transport (Salaun et al. 2004). The transmembrane domains of a complete SNARE quadruplex induce mechanical stress upon the vesicle and plasma membranes, aiding the merging of the two and the formation of a fusion pore (Vardjan et al. 2009). It is SNARE complex formation (and not vesicle fusion) that is the rate-limiting step in the vesicle fusion process (Smith and Weisshaar 2011). An increase in SNARE complexes increases the rate of cargo release from a single pore, enabling the transfer of larger cargo molecules as both pore size and stability are affected by SNARE copy number (Bao et al. 2018; Mohrmann and Sorensen 2012). Depletion of both SNAP29 or Stx19 SNAREs decreases vesicle fusion events at the PM (Gordon et al. 2010). Actin and Myosin II have roles in regulating the duration of fusion pore opening and therefore how much cargo can leave the vesicle during kiss-and-run exocytosis events (Aoki et al. 2010; Trouillon and Ewing 2014). In PC12 cells SNAP-25 is expressed more than 10-fold higher than Stx1, ensuring there is less chance of the formation of an inert complex between a SNAP-25 and two Stx1 molecules, highlighting how stoichiometric levels of SNAREs can better control accurate fusion events (Pobbati et al. 2006). Whilst as little as two SNARE complexes can mediate membrane fusion it is likely that rapid capturing and docking of vesicles requires the formation of more than one SNARE complex (Mohrmann and Sorensen 2012; Sinha et al. 2011).

### 1.4.2. Recycling of transport machinery

The absence of an efficient recycling system from target back to donor organelles would cause net cargo transport to halt, resulting in collapse of donor organelles (Malsam and Soellner 2011). As mentioned in section 1.4.1.4, COPI vesicle transport is involved in retrograde transport through the Golgi as well as the repatriation of ER-Golgi transport components such as VAMP to the ER (with most COPII coat components being shed shortly after budding). Clathrin-coated vesicles are involved in endocytic transport from the PM to the TGN as well as TGN to lysosome transport (Bittner et al. 2013; Traub 2005). A reduction in the SNARE SNAP29 not only decreases the frequency of fusion events at the PM but also increases the number of docked vesicles there, suggesting some SNAREs may have a role in the initiation of vesicle recycling (Gordon et al. 2010).

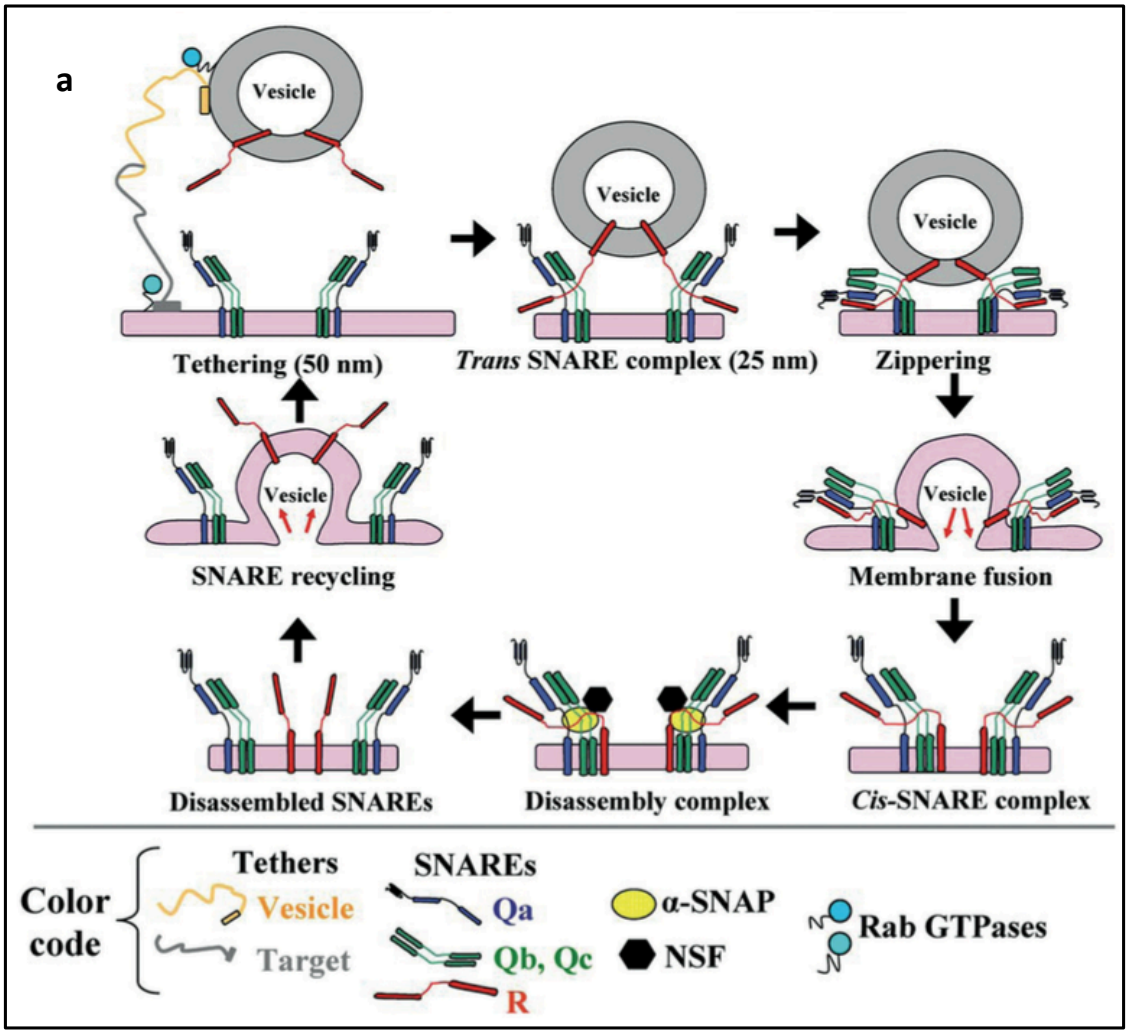


**Figure 1-5. Function of SNAREs in vesicle fusion at target membrane.**

*Starting from [a], tethering at a large distance from the target membrane initiates vesicle capture, bringing v- and t-SNAREs into close proximity and allowing the formation of a quaternary trans-SNARE complex. Zippering of the SNAREs, mediated by SM proteins, overcomes the replant forces between the opposing vesicle and target membranes, bringing them in to close proximity and allowing membrane fusion to occur, resulting in a cis-SNARE complex. NSF proteins mediate SNARE disassembly, allowing v-SNARE recycling via loading into retrograde transport vesicles (Image from Sehgal and Lee 2011).*

### 1.4.3. Secretory organelle stress and the response of the mammalian cell

When the ER's capability to correctly fold, process and package nascent polypeptides (both cargo and resident) is perturbed the build-up of unfolded polypeptide chains leads to ER stress, which eventually induces the unfolded protein response. The UPR aims to regain cellular homeostasis and maintain ER function by balancing the load of incoming protein (by slowing transcription and translation) with the ER's folding capacity (by increasing ER chaperone levels). Due to the increase of protein levels in secretory-specific cells the UPR has a central role in the normal development and function of secreting cells (Moore and Hollien 2012; Walter and Ron 2011). If the UPR is unsuccessful in resolving ER stress apoptotic response pathways are activated, leading to cell death (Jaeger et al. 2012).

The UPR is controlled by an array of signaling cascades initially mediated by three ER-transmembrane receptors: PKR-like ER kinase (PERK), inositol-requiring enzyme 1 (IRE1) and ATF6. All three can mediate both ER recovery and apoptotic responses to ER stress, albeit at varying levels (Hussain et al. 2014; Jaeger et al. 2012). Under normal ER conditions the ER lumenal domains of these proteins interact with Ig binding protein (BiP or Glucose-regulating protein [GRP] 78). BiP binds ATP tightly, with a depletion in cellular ATP levels decreasing protein folding activity and therefore BiP's ability to interact with the unfolded proteins. This lack of interaction with folded proteins and ATP leads to BiP instead binding the transmembrane regulators (Morris et al. 1997). Upon an increase in lumenal nascent polypeptide chains, BiP dissociates from the transmembrane receptors to preferentially bind unfolded proteins, freeing ATF6, IRE1 and PERK to be activated. This initiates ER-to-nucleus signaling cascades and activates the UPR (Bertolotti et al. 2000; Gardner et al. 2013; Samali et al. 2010).

Free from BiP, ATF6 translocates (through constitutive secretion) to the Golgi where it is cleaved, freeing its cytosolic transcriptionally active region so it can enter the nucleus and initiate transcription of ER molecular chaperones, protein disulphide isomerases, CHOP and XBP1 (Shen et al. 2002). Both PERK and IRE, upon being freed from BiP, oligomerise and are phosphorylated to an active state (Chalmers et al. 2017; Gardner et al. 2013). Active PERK phosphorylates eukaryotic initiation factor 2α (eIF2α) which in turn attenuates general translation levels, reducing protein flux to the ER (Harding et al. 2000). Despite this general response, activated eIF2α does initiate translation of ATF4, a transcriptional activator for pro-apoptotic factors (Gardner et al. 2013). Active IRE1 splices XBP1 mRNA, allowing production of the transcriptional activator XBP1s that, upon entering the nucleus, activates transcription of ER chaperones (He et al. 2010). An overview of the UPR to ER stress is shown in figure 1-6.

The UPR to ER stress ensures that only correctly-folded proteins can travel on from the ER, ensuring fidelity in protein structure and folding (Walter and Ron 2011). Development of a non-invasive fluorescence-based reporter that is activated alongside UPR activation allows the UPR status of transgene-producing CHO cells to be quantified and monitored, showing that UPR activation is regulated throughout production culture, that different producing CHO clones vary in the pattern and timing of their UPR-related transcriptional activity and that external factors (media composition and osmolarity) significantly impact upon UPR activation (Du et al. 2013). The destiny of the ER is intrinsically linked to that of the Golgi. Control of Golgi capacity is regulated by wider cellular demands, including the ER stress response. ER stress decreases protein transport to the Golgi, altering its glycolipid composition (Renna et al. 2006).

The ER chaperone HSP47 has been shown to protect cells from Golgi stress-induced apoptosis, although how remains unclear (Miyata et al. 2013). An excessive increase in protein transport to the Golgi can reduce its ability to function correctly with regards to glycosylation and vesicular transport leading to Golgi stress (Sasaki and Yoshida 2015). Like ER stress and the UPR, Golgi stress is abrogated by the Golgi stress response (Yoshida 2009). Golgi stress results in an increase in transcription induction from a transcription enhancing region called the Golgi apparatus stress response element (GASE), which drives transcription of Golgi-related genes including glycosylation enzymes, Golgi structural proteins and post-Golgi vesicle transport factors (Oku et al. 2011). The transcription factor TFE3 has been identified as binding the GASE. Normally residing in its phosphorylated form within the cytoplasm, Golgi stress leads to dephosphorylation of TFE3, allowing it to enter the nucleus and bind the GASE (Beckmann et al. 1990; Taniguchi et al. 2015). A transcription factor similar to TFE3, MLX, also binds the GASE but results in the decrease of its transcriptional activity (Taniguchi et al. 2016).

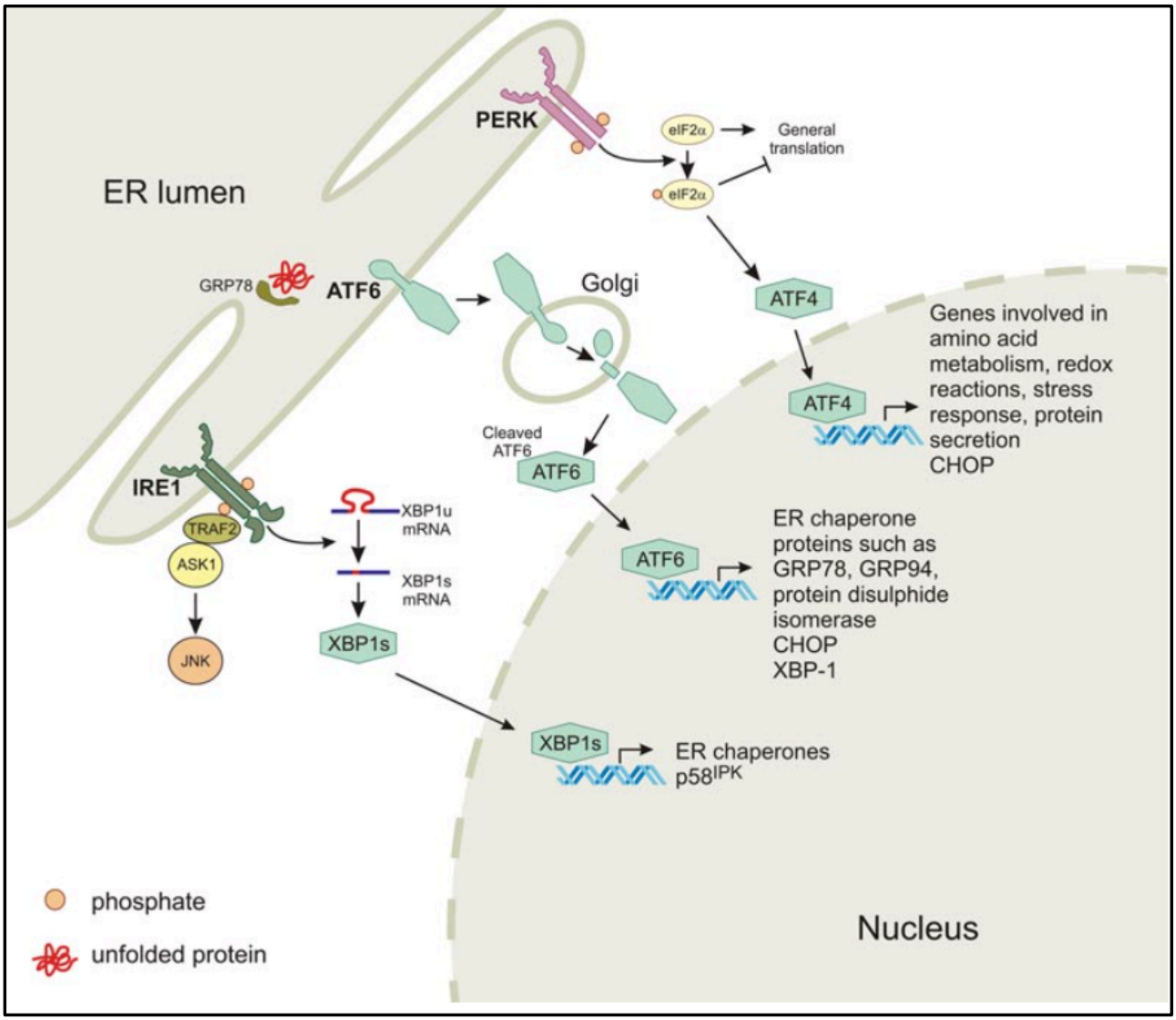


**Figure 1-6. Activation of the UPR in response to ER stress.**

*An increase in unfolded protein leads to it preferentially binding BiP/GRP78, disassociating from IRE1, ATF6 and PERK. Once not bound, PERK monomers dimerise and phosphorylate each other, leading to phosphorylation and activation of eIF2α. This knocks down general translation whilst activating ATF4 which in turn activates transcription of genes involved in amino acid metabolism, redox reactions, protein secretion and stress response. Over-activation of ATF4 can lead to initiation of the apoptotic CHOP pathway. Free IRE1 molecules dimerise and are phosphorylated. Active IRE1 splices XBP1 mRNA, producing the active XBP1s variant which up-regulates transcription of ER chaperones. BiP-free ATF is free to translocate (via secretion) to the Golgi where it is cleaved, allowing its transcriptionally active domain to enter the nucleus and initiate transcription of ER chaperones, PDIs and XBP1. Figure taken from Jaeger et al. 2012.*

## 1.5. Engineering of the CHO cell

The engineering and optimisation of CHO growth conditions, manufacturing processes and media formulation appears to have reached a state of diminishing returns with regards to maximising cell growth and yields from standard CHO cells (Wurm 2004). To further improve CHO cell capacity and productivity it is required to investigate and work within the chassis of the CHO cell to enhance its biosynthetic capacity. With CHO cells not being a specialist protein producing or secreting cell line, there is plenty of scope for synthetic biology and genetic engineering strategies to improve the productivity of easy to express (ETE) and DTE recombinant proteins. The biosynthetic pathway incorporates the production, from a DNA sequence via mRNA and tRNA to a nascent protein that is translocated into the ER where it undergoes folding and PTMs, before entering the secretory pathway which transports the product via the Golgi apparatus to the plasma membrane so it can be secreted from the cell. This section will summarise previous attempts to engineer the CHO cell to enhance its productivity.

### 1.5.1. Previous engineering of CHO cells to enhance productivity

Engineering of CHO cells has thus far focused upon both synthetic biology and genetic engineering strategies. These have been used to optimise transcription and translation, the early biosynthetic pathway (protein processing, folding and processing modifications) and secretory pathway enhancement. Whilst molecular mapping of cell pathways provides a good knowledge of interactions occurring within a cell, the increased availability of transcriptomic and proteomic ('omic) datasets, of many different cell lines, has provided data to inform the selection of genes with which to genetically engineer the cell in an attempt to enhance cell performance (Kumar et al. 2015; Lewis et al. 2016). However, it is worth noting that many synthetic biology and genetic engineering of CHO cells has so far only been tested in small scale shaken culture and whilst this is a good indicator of larger-scale cell culture, these strategies may not be entirely relevant to industrial scale production.

#### *1.5.1.1 Engineering of CHO Transcription and Translation*

The fist steps of the biosynthetic pathway – the central biological tenants of DNA transcription to RNA, followed by RNA translation to a nascent polypeptide sequence – are a natural point at which to start engineering of the CHO cell with a view to enhancing productivity.

##### *1.5.1.1.1. Transcription*

A major factor in expression of a recombinant gene is its copy number and transcriptional activity. Use of strong promoters, targeting integration of recombinant genes to a highly active part of the genome and use of elements that decrease the risk of epigenetic silencing can all achieve increased transcription

and productivity levels (Benton et al. 2002; Deer and Allison 2004; Koduri et al. 2001; O'Callaghan and James 2008; Saunders et al. 2015). All recombinant gene transcription must start from a promoter element. In mammalian cell expression these are normally either endogenous or virus-derived. Despite the availability of the CHO genome sequence, it has so far proved difficult to determine the regulatory sequences that control expression of CHO genes, so endogenous CHO promoters are rare (Brown and James 2016). However, the increase in CHO transcriptomic data sets will likely lead to an increase in the identification of CHO cell gene promoters at varying activity levels (Charaniya et al. 2009; Doolan et al. 2008; Harreither et al. 2015; Nissom et al. 2006). Of the virus-derived promoters, CMV and SV40e are the most commonly used due to their ability to drive high rates of transcription (Brown and James 2016; Chatellard et al. 2007). However maximal transcription levels are not always necessary, either for a selectivity marker, GOI or effector gene.

Synthetic biology approaches have been used to produce promoter elements with which expression levels of recombinant genes within CHO cells can be tightly controlled at myriad levels, allowing precision control of relative stoichiometry of numerous functional components (Brown and James 2016). This allows tighter control of gene stoichiometry which has an effect upon Mab production from HC and LC (Pybus et al. 2014). It also opens up the possibility of varying expression levels of recombinant genes, selection markers and accessory genes. Viral transcription factor regulatory element (TFREs; also known as transcription factor binding sites) used in CHO cells have been screened to determine their relative abundance. A large promoter library, with varying levels of transcriptional efficiency, has been produced by placing blocks of TFREs upstream of a CMV core promoter. Testing of the promoter blocks effect upon secreted alkaline phosphatase (SEAP) expression levels showed that in some cases synthetic promoters can exceed the levels seen from a standard CMV promoter (Brown et al. 2014). Overexpression of the chromatin regulating transcription factor YY1 resulted in an up-to 6-fold increase in antibody titre across various diverse cell lines, both hamster and human derived (Tastanova et al. 2016).

#### *1.5.1.1.2. Translation*

In many cases productivity levels do not increase proportionally with mRNA levels and transgene copy numbers, suggesting translational and post-translational steps are limiting in the production of recombinant proteins (Hussain et al. 2014; Johari et al. 2015). Basic tenants of translation include the use of an AUG start codon downstream of a Kozak sequence that aids ribosome binding at the correct point of the mRNA sequence (Kozak 1987). With the vast majority of biologics being humanised their codon sequence may not be optimised to work within the CHO cell, resulting in sub-optimal translation levels and thus

reduced protein expression. Codon optimisation of the gene sequence of the recombinant protein so that it better matches the codon distribution of the CHO cell can result in enhanced translation levels and an eventual 10-13-fold increase in productivity (Chung et al. 2013; Quax et al. 2015). Use of synthetic non-coding SINEUP RNAs in CHO cells can promote and enhance translation, increasing the production of recombinant protein (Patrucco et al. 2015).

#### *1.5.1.2. Engineering the CHO cell biosynthetic and secretory pathways*

Engineering of improvement of transcriptional, translational and post-translational processes within the CHO cells, whilst enhancing overall productivity levels, can shift synthetic bottlenecks from the nucleus to the ER and further along the biosynthetic pathway. This can result in the overloading the secretory pathway which transports cargo from the ER to the plasma membrane (PM) via the Golgi apparatus and numerous secretory vesicles (Hansen et al. 2017; Zhou et al. 2018).

##### *1.5.1.2.1. ER and UPR engineering*

At the start of a secreted protein's life-cycle, signal recognition particles (SRPs) mediate docking of active ribosomes to the ER, directing nascent polypeptide chains through the translocon so they can be transported into the ER lumen where the SRP is removed. The protein folding process is initiated whilst translocation is in progress, although PTMs such as disulphide bond formation occurs once the entire protein has entered the ER lumen (Robinson et al. 2017). In CHO cells producing a DTE Mab with a malfunctioning SRP complex, engineering of the CHO cell by transient overexpression of SRPs in CHO cells rescued secretion, resulting in a 20 - 40% increase in titre (Le Fourn et al. 2014).

A combination of UPR activity and high cargo protein levels in secretory-specific cells (e.g. plasma cells) confirms that the UPR has a central role in the normal development and function of highly-secreting cells (Moore and Hollien 2012; Walter and Ron 2011). The UPR has also been shown to impact upon vesicle formation and ER maturation (Farhan et al. 2008). UPR engineering by transient overexpression of the active spliced form of the regulating transcription factor X-box binding protein 1 (XBP1s; see section 1.4.3) saw a 5- and 3-fold increase in the expression levels of the model products SEAP and secreted a-amylase (SAMY) respectively. XBP1 also appeared to propagated an increase in ER volume (Tigges and Fussenegger 2006), a 15-85% increase in productivity of multiple therapeutic proteins (Rajendra et al. 2015) and between a 1.4- to 2-fold increase in the titre of various Mabs (Becker et al. 2008; Johari et al. 2015; Pybus et al. 2014). Transient co-expression alongside the ER oxidoreductase ERO1-Lα saw a 5.3- to 6.2-fold increase in Mab yield (Cain et al. 2013). The increase in productivity brought about by XBP1s expression is due to an increase in overall CHO biosynthetic

capacity, with expression having no effect upon transgene mRNA levels (Rajendra et al. 2015).

Of other UPR-related genes used to engineer the CHO cell, stable overexpression of activating transcription factor (ATF) 4 resulted in a 2-fold increase in Antithrombin II production levels (Ohya et al. 2008), though ATF4, CHOP and HSPA5 transient overexpression in CHO cells saw a decrease in Mab expression levels (Rajendra et al. 2015). Overexpression of ATF6 saw an up-to 2.5-fold increase in Mab productivity, though this was dependent upon the specific Mab. Expression of the immunoglobulin binding protein (BiP), which is a UPR marker, saw a 1.5-fold increase in Mab titre (Johari et al. 2015; Pybus et al. 2014). Meanwhile miRNA-mediated inhibition of ATF6's antagonist, ATF6β, resulted in an increase in Qp and overall titre of a stably-expressing IgG whilst also enhancing viable cell density (Pieper et al. 2017b).

Overexpression of ER chaperones and enzymes catalysing processes that aid protein folding is a more specific way to enhance the ER's protein folding capacity. Stable overexpression of members of the HSP70 family in baby hamster kidney (BHK) cells resulted in a 50% increase in expression levels of blood factor VIII (Ishaque et al. 2007). Transient overexpression of Tor1a increased secreted luciferase and IgG4 levels in CHO by 2.5- and 1.4-fold respectively (Josse et al. 2010). Expression of protein disulphide isomerases (PDI[A]s) that catalyse the formation of disulphide bonds increased Mab production by up to 37%, depending upon the Mab expressed (Borth et al. 2005; Johari et al. 2015; Pybus et al. 2014). The chaperone cyclophilin B (CypB) functions in folding of PM-directed proteins including Mabs (Price et al. 1994). Its transient overexpression saw an up-to 1.4-fold increase in Mab titre, again dependent upon the specific Mab being expressed (Johari et al. 2015; Pybus et al. 2014). Stable expression of ERO1-Lα, which transfers disulphide bonds to PDI saw an increase in CHO productivity levels (Cain et al. 2013; Hussain et al. 2014). Stable co-overexpression of the ER chaperones Calnexin and Calreticulin saw an 1.9-fold overall increase in recombinant thrombopoietin levels within the cell, though this did not result in increased secretion (Chung et al. 2004).

#### *1.5.1.2.2. Secretory pathway engineering*

Direct expression of CHO secretory pathway components appears to have been limited to genes involved in vesicle docking and fusion at the target membrane (see sections 1.4.1.4 and 1.4.1.5 for a more in-depth explanation). Overexpression in CHO of the Soluble *N*-ethylmaleide-sensitive fusion protein attachment protein receptors (SNAREs) SNAP-23 and VAMP8, which are both involved in vesicle docking at the Golgi, increased CHO IgG1 production by between 2 - 2.5 fold (Peng et al. 2011). Likewise overexpression of the SNARE modulating proteins Sly1 and

Munc18c saw an increase in the levels of SEAP, SAMY and another model protein by up to 5-fold (Peng and Fussenegger 2009a; Peng and Fussenegger 2009b). Engineering strategies that have targeted the CHO secretory pathway resulting in increased productivity levels are summarised in figure 1-7.

#### *1.5.1.2.3. Wider engineering targets*

Modulation of ceramide transport from the ER to the Golgi by stable overexpression in CHO cells of an active version of the ceramide transport protein (CERT) saw a 26 - 45% increase in IgG and tPA production levels (Florin et al. 2009; Rahimpour et al. 2013). Stable expression of an oncogenic mutant of tyrosine receptor kinase (onco-KIT) increased global protein synthesis and improved cell growth and stress resistance. This resulted in a highly-proliferating cell line with a greater than 2-fold increase in productivity of a GFP-Mab Fc fusion protein (Mahameed and Tirosh 2017). Stable overexpression of Glutamate-cysteine ligase modifier subunit (GCLM) resulted in increased production of the antioxidant glutathione (GSH). Whilst its overexpression in CHO cells resulted in a 71% increase in IgG1 titre, expression of a related protein which resulted in even higher GSH levels did not enhance productivity (Orellana et al. 2017). This shows that pushing the level of intracellular metabolites and signaling molecules can have a deleterious effect when compared to smaller changes. Upregulation of amino acid transporters (AATs) can enhance the cellular uptake of specific amino acids that are more prevalent in Mabs than normal cell processes, as well as enhancing levels of amino acids required for general CHO housekeeping procedures such as GSH synthesis and protein synthesis (Geoghegan et al. 2018).

Whilst ectopic expression of genes is a proven way to alter the synthetic and secretory phenotype of the CHO cell, the process of getting DNA within the cell does have some effect upon cell growth. The transcription and translation of the effector genes adds further to the metabolic burden already placed upon the cell by the recombinant GOI. As such, knockdown of genes that negatively impact upon cell productivity levels, such as miRNA and CRISPR, may provide an accessible method of cell line engineering without further burdening the CHO cell (Fischer et al. 2017; He et al. 2015). Knockdown of the ER-resident ceramide synthase 2 (CerS2) and the Rab1 GAP Tbc1D20, mediated by siRNA, increased the Qp of CHO cells stably expressing IgG. This response was greatly enhanced when both genes were knocked down, alongside an enhancement in cell growth. An altered cellular ceramide composition and an increase in Rab1 activity respectively is believed to have improved vesicular transport at the ER level (Pieper et al. 2017a).

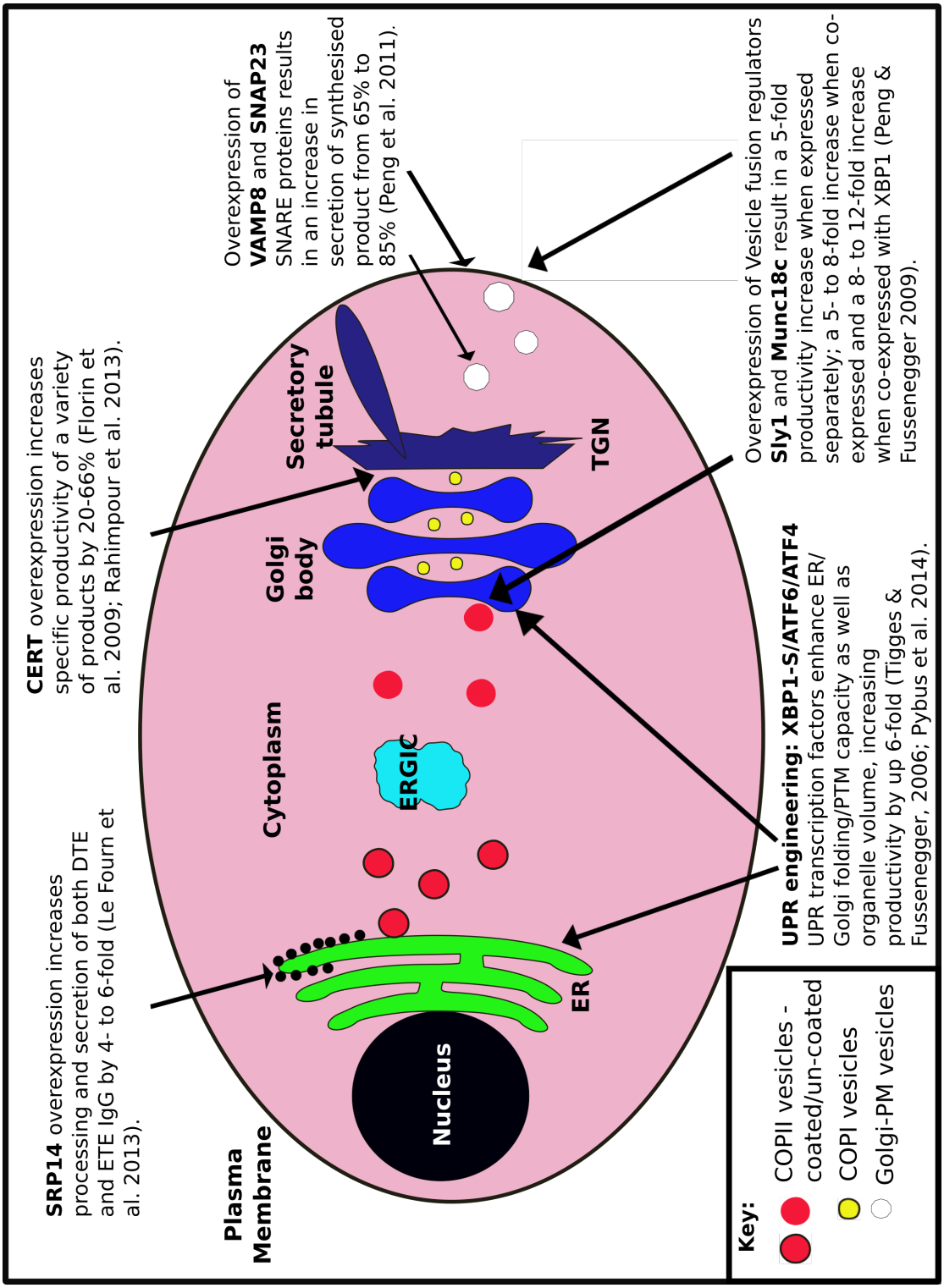


**Figure 1-7. Previous engineering of the CHO secretory pathway.**

*Engineering of the CHO secretory pathway has so far focussed on translocation of nascent proteins into the ER, engineering of ceramide transport to the Golgi, increasing vesicle docking and fusion through SNARE and SNARE effector proteins and engineering of the UPR through expression of active transcription factors that mediate the UPR.*

#### *1.5.1.3. Other methods of CHO cell genetic engineering*

Overexpression of one or more effector gene(s) may help produce an enhanced productivity phenotype, but production of further recombinant genes (alongside that of the recombinant product) can add a further metabolic burden upon the cell. As such, methods to enhance CHO cell productivity without the need for further transcription and translation burden are of increasing importance alongside the introduction of effector genes in to cells (Fischer et al. 2015). MicroRNAs are small, non-protein coding mRNAs that are capable of post-transcriptionally regulating entire mammalian cellular pathways through, enhancing gene expression through translation inhibition or destabilisation (Barron et al. 2011b). Stable expression of microRNAs in CHO cells enhances process steps in cell line development and significantly increases final product yield (Barron et al. 2011a; Fischer et al. 2017).

Use of targeted gene mutation technologies such as CRISPR and transcription activator-like effector nucleases (TALENs) can be used to knock down and/or mutate specific genes, the activity of which of which can impact upon product quality, cell growth and productivity (Cong et al. 2013; Yang et al. 2015). RNA interference techniques such as small interfering RNAs, utilising mRNA-complementary sequences and thus blocking its translation, can also be used to knock down gene expression, being used to improve apoptosis resistance, glycosylation and cell metabolism, increasing specific productivity and titres as well as product quality (Fischer et al. 2015; Mori et al. 2004; Zhou et al. 2011).

That the engineering of some pathways has positive effects upon production levels with some cells and products, but the opposite in others, suggests that there is some level of specificity with regards to engineering strategy and the cell line and model Mab used. However, in the case of global transcriptional activators such as XBP1s, there does seem to be an overwhelming consensus that its co-expression results in an improvement in protein production levels.

### 1.5.2. Why further engineer the CHO secretory pathway?

A non-linear relationship between GOI mRNA and titre levels shows that in CHO cells there is a post-transcriptional bottleneck in the biosynthesis of recombinant proteins. Were this not the case, the level of transgene mRNA would be in direct proportion to productivity level (Hansen et al. 2017; Johari et al. 2015). It stands to reason that, in a cell line not specified for high protein production or secretion levels, whilst the optimisation of transcription, translation can increase the levels of a recombinant protein a cell is capable of producing, the cell may not be able to cope with this increase in cargo in a biosynthetic capacity (protein folding, PTMs and secretion from the cell). Likewise enhancing protein folding and glycosylation

would only increase the burden upon the protein secretory pathway. Increasing mRNA and protein levels merely shifts any productivity bottleneck limiting cellular productivity further down the biosynthetic pathway. Increased protein levels can also result in aggregation and blocking within the secretory pathway which can activate the UPR, reducing cell growth or targeting cells for apoptosis (Reinhart et al. 2014; Zhou et al. 2018).

A biologic may prove DTE due to many reasons. An inability for the complete molecule to assemble from its constituent parts; improper protein chain folding; improper PTM processing; specific structural elements or sequences within the polypeptide sequence leading to the formation of an inactive conformation; or regions allowing easy aggregation with other molecules may occur. Many of these issues result in activation of cellular stress responses, especially ER and Golgi stress, as improperly folded proteins build up with the ER and Golgi (Hussain et al. 2014; Reinhart et al. 2014).

The formation of complete Mabs from constituent heavy and light chains can be affected by the levels of these components within the cell. The stoichiometry of HC:LC genes transcribed and translated by the cell can have an impact upon the formation of complete Mabs, with the expression levels of DTE Mabs being improved by alteration of the transfected HC:LC gene ratio (Pybus et al. 2014; Strutzenberger et al. 1999). Modification of a protein sequence can alleviate structural issues such as incorrect folding and interactions leading to aggregation. This can be streamlined by utilisation of an engineering strategy based upon related proteins that are easily expressed, accelerating the molecular development process (Hussain et al. 2017). As previously discussed, engineering of ER folding and the UPR through co-expression of ER-resident molecular chaperones (e.g. BiP, PDI) or active UPR transactivators (e.g. ATF6, XBP1s) can increase biosynthetic and secretory flux. Furthermore, treatment of cells with chemical chaperones (e.g. DMSO, glycerol) or UPR signal modulators (e.g. PERK inhibitors) can decrease the apoptotic pathways of the UPR and can result in a 6-fold increase in DTE expression levels (Johari et al. 2015; Pybus et al. 2014).

Molecular engineering of problem biologics is probably the most time-efficient method of improving DTE expression, with optimisation of each molecule an important step in the research and development pipeline (Hussain et al. 2017). Whilst specific host-cell engineering could be used to improve productivity, producing cell lines bespoke to a certain molecule would be a time and resource intensive process starting from basic principles for each different molecule. Utilisation of engineering strategies and HTS systems could speed up this process but it would still prove too lengthy to be viable in an industrial environment. However, production of a few general cell lines with enhanced productivity and secretion levels of model recombinant proteins will streamline the process. By

providing a chassis cell line(s) with secretory bottlenecks already removed and thus having an already high secretory capacity, engineering of the protein production pathway upstream of this (from transfection and transcription through to glycosylation) will allow for a more immediate feed-back as to the effect these tweaks will have upon overall titre. It may also be possible to pair a specific DTE protein with a specific cell line depending upon the reason it is DTE.

## 1.6. Conclusion and project introduction

The CHO cell is an important mammalian cell factory that is well used in the production of biopharmaceuticals due to its ability to carry out PTMs, relatively quick growth, genetic plasticity and regulatory status. Engineering of CHO manufacturing processes and more latterly the CHO cell itself, has resulted in increased yields of a wide range of recombinant proteins. However, as drugs companies diversify, trying to make more complex molecules to target more complex and a wider range of disease targets, there is still need to increase the capacity of the overall CHO production chassis.

With the availability and improving annotation of the CHO genome, as well as increasing transcriptomic and proteomic data sets of both various CHO strains and non-related cell lines with high productivity characteristics, there is an increasing knowledge as to how CHO responds on an 'omic level to different stimuli and production processes. There is also a widening base of knowledge – such as transcriptomic and proteomic data - which can be mined to provide directed engineering strategies to try and improve the productivity levels of CHO cells. Whilst much elucidation of the eukaryotic secretory pathway has been carried out in yeast, many yeast secretory proteins have close homologues in mammalian cells. Furthermore increasing levels of secretory pathway dissection is being carried out in mammalian cells.

Whilst improving yields of biologics, single gene engineering of CHO biosynthetic and secretory pathway components will likely only result in the shifting of bottlenecks elsewhere within these pathways, overloading them elsewhere. To produce CHO cells with a truly enhanced biosynthetic and secretory phenotype a more holistic and global approach to CHO cell engineering is required. Facilitating myriad genetic changes throughout the CHO cell, focusing especially upon the biosynthetic and secretory pathways, should allow optimisation of the processing and transport of increasing levels of recombinant protein from the ER through to the PM, increasing general CHO productivity as well as expanding its expression repertoire.

The overall aim of this project is to improve the productivity phenotype of the CHO cell through engineering of its secretory pathway. It is envisaged that

expansion of the CHO secretory pathway (alongside the biosynthetic pathway) will remove bottlenecks that reduce CHO productivity levels. It has already been shown that single-gene engineering of the CHO secretory pathway can increase CHO productivity, through SNARE expression and control, ceramide transport to the Golgi, enhancing polypeptide translocation into the ER and engineering of the UPR. Whilst some genes have been co-expressed to good effect, none of these strategies have taken a multigene approach to the engineering of the CHO secretory pathway. This project aims to identify new genetic targets for secretory pathway engineering. These will be used, alongside previously used targets, in a multigene strategy, to engineer the CHO secretory pathway through direct overexpression of single and combinatorial targets. Directed evolutionary, strategies will also be used to produce myriad small alterations within the CHO cell. All of this will be performed with the aim of enhancing the secretory capacity of the CHO cell to propagate an increase in overall CHO productivity levels.

# 2) Materials and methods

## 2.1. Cell Culture and Maintenance

A suspension-adapted CHO-K1-derived cell line with low endogenous Glutamine Synthetase (GS) levels was provided by MedImmune (Granta Park, Cambridge). This host cell line (henceforth referred to as MedI-CHO) was used for the majority of cell culture experimental work, with all other CHO cell lines derived from it during the course of the project (including evolved and stably-expressing cell lines). Information on cell lines used and developed during the project are summarised in table 2-1.

### 2.1.1. Cell Growth Media

Standard base growth media for all CHO cell cultures was proprietary chemically defined (CD)-CHO growth media (GIBCO, Invitrogen, Carlsbad, CA, USA).

#### *2.1.1.1. CHO cell growth media*

Suspension MedI-CHO was cultured in CD-CHO media supplemented with 6 mM L-Glutamine (GIBCO, Invitrogen). Cell lines stably expressing a recombinant DNA construct were cultured in CD-CHO supplemented with 50 μM Methionine Sulphoxamine (MSX – Sigma, St. Louis, MO, USA) as a selective agent. For long-term transient expression studies a specifically developed cell line (MedI-Tran) was used. MedI-Tran was maintained in CD-CHO supplemented with 25 μM MSX and 100 μg/mL Hygromycin-B. All CHO cell lines, the abbreviations used throughout the project and growth media supplementation they require are shown in table 2-1.

### 2.1.2 Shaken CHO cell culture

Standard cell growth was carried out under orbital shaken conditions in growth media relevant to each cell line as described previously (section 2.1.1; table 2-1). Passages in to fresh growth media were carried out every three or four days with a seeding density of $0.2 \times 10^6$ cells $mL^{-1}$. All shaken cultures were carried out in Multitron HT (Infors, Reigate, UK) orbital shaking incubators at 37 °C, supplemented with 5% $CO_2$. Where possible humidity was maintained between 70-85%.

**Table 2-1. CHO cell lines used during project**

| Cell Line | Description | Growth Media | In-text abbreviation |
|---|---|---|---|
| **A) Cell lines provided by MedImmune** | | | |
| MedImmune host CHO cell line | CHO-K1 derived GS knockout. Host cell line used for majority of expression and evolution work. | CD-CHO, 6 mM L-Glutamine | MedI-CHO |
| Long-term transient CHO line | CHO line developed for long-term transient expression of recombinant genes. | CD-CHO; 25 µM MSX; 100 µg/mL Hygromycin-B | MedI-Tran |
| **B) Cell lines produced by NB during project** | | | |
| Stably-expressing CHO cells | CHO cell line stably expressing recombinant Mab gene(s). | CD-CHO, 50 µM MSX | Initial cell line /recombinant Mab dependent |
| Evolved CHO cell line | MedI-CHO line evolved against various secretory blocking agent. | CD-CHO, 6 mM L-Glutamine, 1 µM relevant secretory blocking agent | MedI-BA (BFA); MedI-FLI (FLI-06) |
| Chemically filtered CHO cell lines | MedI-CHO line filtered with various chemical combinations. | CD-CHO, 6 mM L-Glutamine, (+ filter drug) | MedI-DF (dual filter); MedI-BF (BFA filter) |

*An overview of the cell lines used throughout the project. **A**. Cell lines provided directly provided by MedImmmune and maintained as per their culturing protocol. **B**. Cell lines produced by NB, by stable transfection or directed evolution treatment. These cell lines were derived from MedI-CHO cells.*

#### *2.1.2.1. General CHO cell culture*

General cell maintenance was carried out in 20-30 mL growth media volumes in vented 125/250 mL Erlenmeyer flasks (e125/e250; Corning, New York, NY, USA). Expanded cultures were carried out in larger vented flasks, culture volume not exceeding 24% of the total flask volume. Erlenmeyer cultures were shaken at 140 revolutions per minute (RPM) with a 50 mm throw. Smaller cultures (5-10 mL) for evolution and cell maintenance were grown in 50 mL TubeSpin bioreactor tubes (TPP, Trasadingen, Switzerland; henceforth referred to as 'cultiflasks'), shaken at 170 RPM with a 25 mm throw/240 RPM with a 50 mm throw. Both these shaking protocols result in the same centrifugal force placed upon cell culture, according to Equation 1 from Infors.

**Equation 1 –** converting shaking incubator speed for different throw diameters.

$$(rpm25)^2 \times 25 = (rpm50)^2 \times 50$$

(where rpm25 = rpm with a 25 mm throw; rpm50 = rpm with a 50 mm throw).

#### *2.1.2.2. High-throughput small scale CHO cell culture*

High-throughput deep-well plate culture allowed cell growth and productivity to be assayed under many different conditions, allowing quick execution of DoE gene-expression experiments and kill curves with many permutations. Both batch and fed-batch cell culture was utilised throughout the project.

Deep well plates were used for high-throughput shaken cell culture. Cells were seeded at 0.2-0.4 x $10^6$ cells $mL^{-1}$ per well in 475 µL (96 well) or 2 mL (24 well) starting volume per well in square well V-bottomed deep well plates (Greiner Bio-One, Stonehouse, UK). Plates were covered with relevant 24- or 96-well 'Sandwich cover' lids (Duetz Enzyscreen B.V, Heemstede, Netherlands) to seal the deep well plates to improve gas transfer and reduce evaporation and cross-contamination between wells. Cells were shaken at 320 rpm with a 25 mm throw, with plates and lids held in place by a specific clamp system (Duetz, Enzyscreen B.V) to allow the vigorous orbital shaking required for optimal cell culture performance. Cell culture was carried out in a MultiTron HT incubator at 37 °C, 5% $CO_2$ with humidity maintained at 85% to reduce culture loss by evaporation.

Batch cultures were grown for between 2-4 days before analysis samples were taken. Fed-batch cultures (e.g. long-term transient gene expression; see section 2.5.2) were maintained for 120 h, with a feed carried out at 72 h post-seeding when cells were supplemented with a 1:1 mix of EfficientFeed™ A and B (Gibco, Invitrogen). The 96-well deep-well plate was covered with a sterile plastic lid (Greiner Bio-One) and centrifuged at 300 x g for 5 minutes to pellet cells. Supernatant (10% of initial seeding volume, 47.5 µL) was removed and replaced

with 47.5 µL of efficient feed mix. Cells were resuspended by gentle pipetting before being returned to the incubator. Growth analysis of both batch and fed-batch cultures was carried out by removing 100 µL of cell culture to flat-bottomed shallow 96-well plates (Nunc) for PrestoBlue® analysis (section 2.1.5.1).

## 2.1.3. Static CHO cell culture

Static cell culture was performed in well plates or tissue culture flasks. 6-, 12-, 24- and 96-well plates (Nunc, Roskilde Denmark) with maximum working volumes of 3.0, 2.0, 1.5 and 0.2 mL respectively. For larger static culture, including stable pool production, cells were seeded in 25 $cm^2$ (T25) tissue culture flasks (Corning) in 5-6 mL growth media. For all static culture, cells were seeded at 0.2-0.4 x $10^6$ cells $mL^{-1}$ and incubated at 37 °C in a HeraCell 150 static incubator (ThermoFisher scientific, Waltham, MA, USA), supplemented with 5% $CO_2$ and humidified with a water reservoir to reduce culture loss by evaporation.

## 2.1.4. Cell counting

Cell number and viability level of cells was determined by Trypan Blue dye exclusion with a VI-CELL-XR cell viability analyser (Beckman-Coulter, Brea, CA, USA). Total cell density (TCD; total number of cells x $10^6$ $mL^{-1}$), viable cell density (VCD; total number of live cells x $10^6$ $mL^{-1}$), percentage of cells that are viable (% viability) and cell diameter (microns, µm) were the measurements most utilised to track cell growth and culture health. For measurement, approximately 600 µL of culture was transferred a measuring cup and loaded onto the ViCell. Where a cell count was expected to be over 8.0 x $10^6$ cells $mL^{-1}$, sample was diluted in PBS.

### *2.1.4.1. Cell Counting of small volume, high-throughput cultures*

Measurement of cell growth in small-scale, high-throughput cultures (as described in section 2.1.2.2) could not be measured by ViCell due to time and cost constraints. PrestoBlue® cell viability reagent (Invitrogen) was used to give an indication of cell viability and therefore cell growth. Whilst not providing actual cell concentration levels, PrestoBlue® (PB) is a colourimetric assay which changes from blue to pink in the presence of cells, providing a relative fluorescence unit (RFU) measurement relative to the amount of viable cells present within the sample. From each culture well, 100 µL of culture medium was transferred to a fresh 96 shallow-well plate. On each plate analysed, three blank wells, containing 100 µL of plain CD-CHO medium, were included. To each well 22.5 µL of room-temperature PB solution (a pre-mixed 1:1 PB:CD-CHO solution) was added by multichannel pipette to reduce time variability between wells. The cell/PB mix was shaken for 20 seconds using a PHERAstar® microplate reader (BMG labtech, Offenburg, Germany) shake protocol before static incubation at 37 °C for 35 mins. Post-incubation samples were measured using a PHERAstar® microplate reader. PB was excited at 560 nm, with emission measured at 590 nm.

The RFU values produced by the plate reader were normalised to the average value of the three media-blank wells run on each plate. A sample:blank ratio was calculated by dividing each sample value by the average blank value. Use of this ratio, instead of just subtracting the blank value, allowed for variances in plate reader gain settings between runs, allowing plates from separate experiments to be compared directly. This was essential for growth calculations. To determine growth over a period of time, PB samples were taken at both the start and end points of culture, from which IVCD change and growth rate values could be calculated (section 2.2). A PB standard curve was performed to show the linear relationship between read:blank RFU ratio and cell number (figure 2-1). A serial dilution of MedI-CHO cells, from 10 x $10^6$ cells $mL^{-1}$ diluted 1:1 with growth media in a series of 7, was carried out in 2 mL volumes, well mixed by pipetting before each dilution. A ViCell sample of each dilution was taken to ensure accuracy. Each dilution was split in to four 100 µL aliquots for PB analysis as described above.

The standard curve shows that there is some saturation of the PB assay above 5 x $10^6$ cells $mL^{-1}$, resulting in an $r^2$ value of 0.9695 (figure 2-1 a). Removal of the 10 x $10^6$ data point shows the close linear relationship between RFU and VCD values up to 5 x $10^6$, with an $r^2$ value of 0.9963. This standard curve shows that PB analysis is accurate up to a RFU/blank value of 15, representing a VCD up to approximately 5 x $10^6$ (figure 2-1 b). Whilst there is some relationship above this VCD, it is not as linear as at lower VCD values and as such, RFU:blank ratio values far exceeding a value of 15 can not be fully trusted. This standard curve will not be used to directly calculate VCD values which will be represented as RFU:blank values, but is purely an indication of PB's use as a VCD assay.

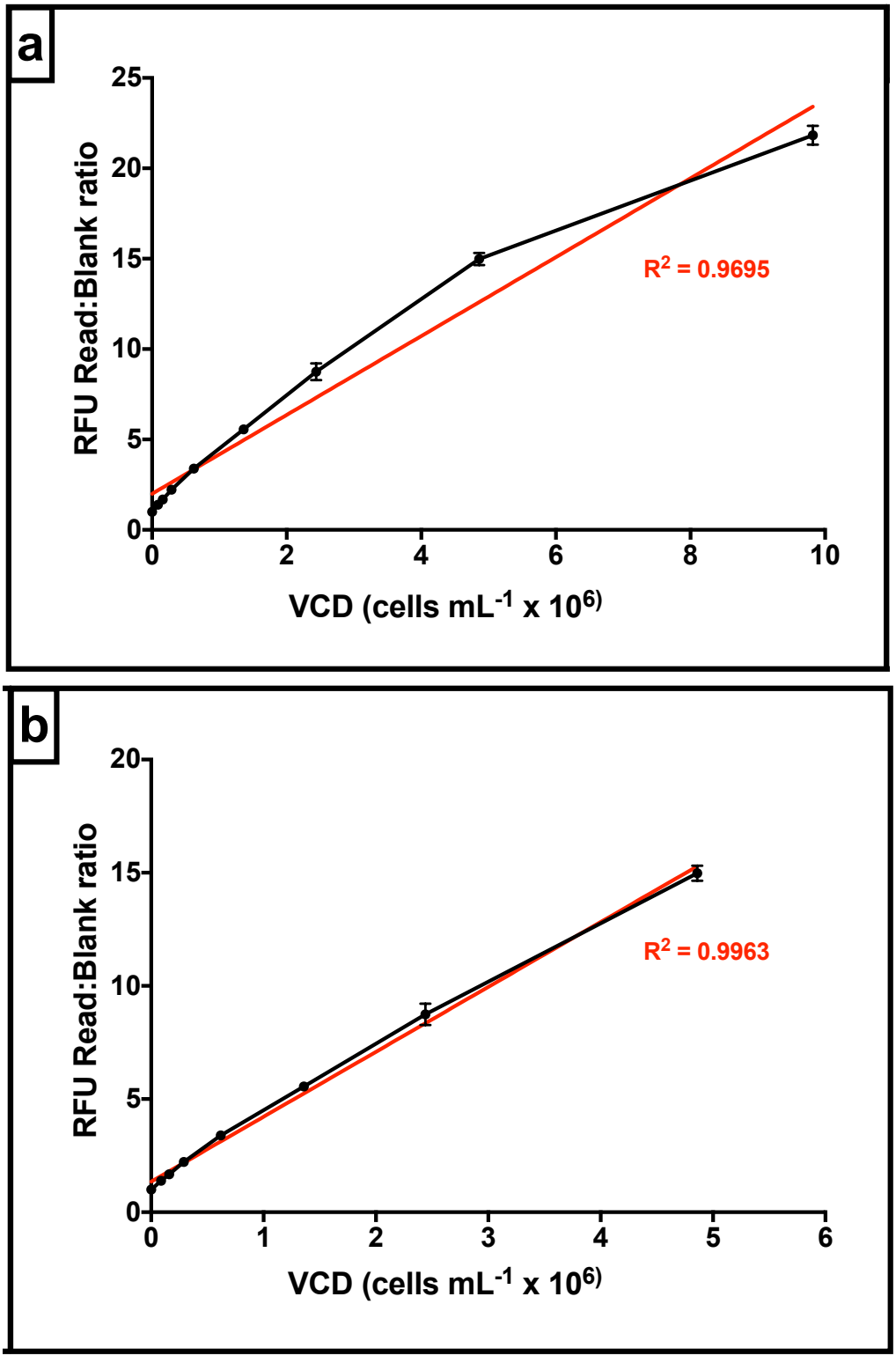


**Figure 2-1. PrestoBlue® standard curve.**

*A serial dilution of MedI-CHO cells was carried out in growth media before analysis by both ViCell and PB. Each PB sample was repeated in triplicate. RFU and ViCell data were plotted against each other to produce a PB standard curve between [a] zero and 10 x $10^6$ cells $mL^{-1}$ and [b] zero and 5 x $10^6$ cells $mL^{-1}$. Linear regression of each curve was performed to provide $r^2$ values. Each data point shows the mean of four technical replicates. Error bars show SEM.*

### 2.1.5. Cell bank production and storage

Cell lines were taken, where possible, from a Master Cell Bank (MCB), from which a larger working cell bank (WCB) was produced as close to the original MCB as possible (minimum four passages of growth from previous vial). Each M/WCB vial contained between 10.0 - 15.0 x $10^6$ cells in 1.5 mL freezing media (chilled CD-CHO, 7.5% Di-Methyl Sulphoxide [DMSO], Sigma). The required volume of culture (to provide 10 - 15 x $10^6$ cells) was centrifuged at 130 relative centrifugal field (RCF) for 5 mins and the supernatant discarded. The cell pellet was gently resuspended in 1.5 mL freezing media per vial required. Cells were transferred into sterile 2 mL cryovials (ThermoFisher) and cooled at 1 °C per minute to -80 °C in a Mr. Frosty (Sigma) containing isopropanol. After overnight freezing vials were transferred to a liquid nitrogen cryostat for long-term storage.

### 2.1.6. Cell bank revival

MCB/WCB vials were thawed at 37 °C in a water bath for four mins. Vial contents (approximately 1.5 mL) were gently transferred to a 50 mL Falcoln tube by 2 mL stripette and washed by slow addition of pre-warmed (37 °C) plain CD-CHO to removed DMSO: CD-CHO (10 mL) was slowly added before the final volume was made up to 45 mL with a further 33.5 mL warm CD-CHO. Cells were pelleted at 130 RCF for 5 mins and the supernatant removed carefully by decanting. Pelleted cells were re-suspended in 10 mL growth media (supplemented CD-CHO) and a cell count carried out. For the first passage cells were seeded at 0.3 x $10^6$ cells $mL^{-1}$ in 25 mL media in a e125 flask and grown for 3 days. All further passages following the usual 3- or 4-day duration at a density of 0.2 x $10^6$ cells $mL^{-1}$. Cells were cultured for at least 4 full passages before being used for experimental purposes to allow cell growth to recover and allow any issues from the freeze/thaw protocol to grow out.

### 2.1.7. Sample harvesting and storage

Were cell growth to be continued after a time point, a sterile sample of the culture was taken and centrifuged in a bucket centrifuge at room temperature (21 °C) at 130 RCF for 5 mins. For end-point samples (after which the cells would not be required for growth), this process was carried out at 1000 RCF.

#### *2.1.7.1. Supernatant sample storage*

Supernatant samples were removed from pelleted cells by pipetting. Samples were transferred to a labeled 1.5 mL microfuge tube or a 96-well plate (up to 200 µL volume) which was covered with a microplate cover (SLS, Nottingham, UK) to minimise evaporation. Supernatant samples were stored at 4 °C for up to a week. Beyond this duration sample storage was at -20 °C. All comparative samples underwent the same freeze-thaw process to allow for direct comparisons to be made. Where possible, Valita™TITER analysis (section 2.6) was carried out within a couple of days of harvesting to produce results as accurate as possible.

#### *2.1.7.2. Cell pellet sample storage*

Cells were centrifuged in 1.5 mL microfuge tubes or 15/50 mL Falcon tubes and as much supernatant as possible removed by decanting, aspiration by pipetting and blotting. Cells pellets were briefly air-dried to allow residual liquid to evaporate. Pellets were stored at -80 °C for up to six months before use.

## 2.2 Cell growth and productivity calculations

A growth curve of the MedI-CHO WCB used for all experimental work was performed so as to define cell growth and better inform sampling and feeding strategies. Cells were seeded in 50 mL total volume of growth media at 0.2 x $10^6$ cells $mL^{-1}$. Cells were grown as described in section 2.1.2, with ViCell samples taken at daily intervals to produce a cell growth curve (figure 2-2). Data collected from ViCell measurements were used to calculate various cell growth metrics to better track cell growth and productivity during the course of cell culture. For these calculations, *t* represents time (normally measured in days), whilst T represents product titre (usually in mg/L). Any subscripts following these symbols represent a specific time-point. E.g. $t_0$ shows culture start time, whilst $T_1$ represents the titre at time-point 1 (normally the end-point of a culture).

### 2.2.1. Growth rate

The growth of cell cultures grown over similar time periods were directly compared by calculating their specific growth rate (μ $day^{-1}$ , equation 2). This utilised a culture's starting VCD (VCD at *t*=0; $X_{v0}$), final VCD (VCD at *t*=1; $X_{v1}$) and change in time (Δt, days) to compare cell growth starting at different VCDs and at different times. The formula was taken from previous work in the DCJ group (Pybus et al. 2014). To ensure μ could be directly compared, culture humidity had to remain constant across experiments so as to not impact upon VCD by diluting or concentrating cultures.

For kill curves and transfections, an initial seeding density was measured by ViCell and used as $X_{v1}$ to provide accurate μ and IVCD levels. For general cell culture and evolution, the calculated seeding density (normally 0.2 x $10^6$) was used as $X_{v1}$.

**Equation 2** – Calculating specific growth rate of mammalian cells**.**

$$\mu = \frac{Ln(X_{v1}) - Ln(X_{v0})}{\Delta t}$$

where Ln = natural log

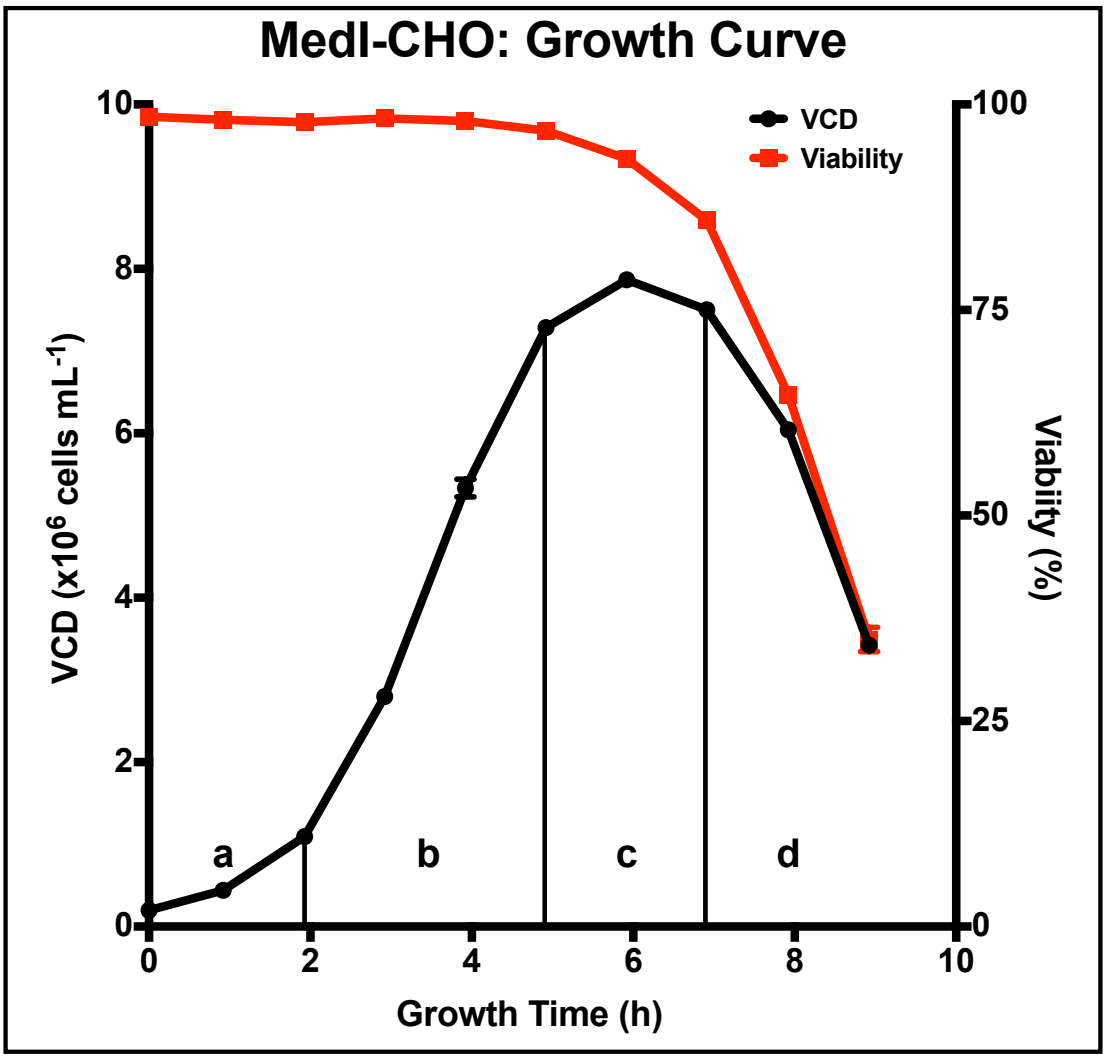


**Figure 2-2. MedI-CHO growth curve.**

*MedI-CHO cells were seeded at 0.2 x $10^6$ cells $mL^{-1}$ in 50 mL cultures and grown for 9 days. ViCell samples for VCD and viability were taken every day. The growth curve shows the different growth phases of a CHO shaken culture, approximately [a] lag phase; [b] log/exponential growth phase; [c] stationary/production phase and [d] death phase.*

### 2.2.2. Doubling time

The time taken for a cell culture to double in size, doubling time (DT, days, Equation 3), was calculated from the specific growth rate of a culture (μ).

**Equation 3 –** calculating the doubling time of a population of cells.

$$DT = \frac{Ln(2)}{\mu}$$

### 2.2.3. Generation number

Generation number (Gen) was used to determine the number of generations (i.e. cell divisions) a culture underwent in one passage (Equation 4). The number of generations a cell line progressed through during a multiple-passage experiment was determined by addition of the generation numbers from each separate passage. Generation number was used to directly compare the relative age of cultures, allowing harvesting, testing and treatments to be carried out at the same relative culture age across different cultures, as well as to assume cell stability.

**Equation 4 –** calculating the generation number of a population of cells**.**

$$Gen = \frac{\Delta t}{DT}$$

### 2.2.4. Integral of viable cell density (IVCD)

The Integral of Viable Cell Density (IVCD, cell day $mL^{-1}$; Equation 5) was calculated to better compare the change in VCD over culture, for cultures grown at different times, using VCD values at the start (t0, $X_{v0}$) and end (t1, $X_{v1}$) of the measurement period. The IVCD of a culture gives the area under a culture's growth curve, defining the total amount of time the cells in a culture have worked, assuming that every cell has the same production capacity (Pybus et al. 2014).

**Equation 5 –** calculating the IVCD of CHO cells.

$$IVCD = \frac{X_{v0} + X_{v1}}{2} \times \Delta t$$

For growth and productivity curves the increase in IVCD across an entire culture (cumulative IVCD) was calculated by addition of consecutive IVCD values to each other.

### 2.2.5. Specific productivity rate (Qp)

The daily specific productivity rate of a culture (Qp, pg $cell^{-1}$ $day^{-1}$; Equation 6) was used to normalise culture productivity levels for varying cell growth rates. It was taken from previous work in the DCJ group (Pybus et al. 2014). Qp was calculated using the change in titre level between time-point 0 ($T_0$) and time point

1 ($T_1$), the change in cell density between the two time points ($X_{v0}$, $X_{v1}$) and the overall change in time (Δt).

**Equation 6 –** calculating the specific productivity of CHO cells**.**

$$Qp = \frac{T_1 - T_0}{(X_{v0} + X_{v1})/2} \div \Delta t$$

In cases of transient transfection and passaging of a culture, $T_0$ was assumed to be zero.

## 2.3. Preparation of compounds to treat cell culture

Compounds used in treatment/evolution of cell culture are described in chapter 5. Stock solutions were produced at a concentration informed by functional working concentrations used previously in either the literature or previously within the laboratory group.

### 2.3.1. Production of stock solutions of compounds

Stock solutions were produced by dissolving the compound in either 100% DMSO or 100% ethanol (EtOH) at 1000-5000x expected working concentrations (based on information from the literature). The vehicle solution used was determined by the required stock concentration, manufacturer information on solubility levels in different solutions and the expected volume of dissolved drug compound to be added to cell culture. DMSO and ethanol volumes were minimised so as to be below 0.15% (DMSO) or 1% (EtOH) of total cell culture volume when compounds were used at milimolar concentrations, so as to avoid any effect(s) of the vehicle compound upon cell growth or function. Prior to use, vehicle solution was filter sterilised using a syringe-tip filter (0.2 μm, Corning). Due to compounds being bought in small (milligram, mg) quantities, vehicle solution was added directly to the delivery vial and mixed by vortex until the compound was completely dissolved, thus ensuring accurate concentrations were achieved. Where required the initial solution was further diluted to the required concentration with CD-CHO. Working solution aliquots were stored in pre-sterilised light-resistant (brown) 1.5 mL microfuge tubes when prepared in DMSO, with pre-sterilised clear microfuge tubes used for EtOH and di$H_2O$-dissolved compounds. Aliquots were stored at -20 °C and thawed at 37 °C before addition to cell culture at the required working solution. Aliquot volumes were limited to approximately 200 μL so as to reduce any negative effect of excessive repeated freeze-thaw processes. Production of working stock solutions is shown in table 2-2.

**Table 2-2. Production of drug compounds for use in cell culture.**

| Compound | Manufacturer | [Stock solution] | Vehicle solution |
|---|---|---|---|
| **Brefeldin A (BFA)** | Sigma | 5 mM | 100% Ethanol |
| **Exo1** | Santa Cruz | 50 mM | DMSO |
| **Exo2** | Sigma | 10 mM | DMSO |
| **FLI-06** | Santa Cruz | 1 mM | DMSO to 10 mM, further diluted to 1 mM in CD-CHO. |

*The secretory-blocking compounds used during the project are listed, alongside the method for production of a stock solution for use in cell culture.*

## 2.3.2. Kill/dose response curves

Compounds were tested within CHO cell culture at varying levels to determine their effect upon cell growth and thus effective functional range. A likely functional range was determined from previous literature before being expanded for testing on CHO cells. For initial kill curves to test new compounds cells were seeded at 0.2 x $10^6$ cells $mL^{-1}$ in static plates with the test compound added at the required concentrations, incubated for 72-96 h before cell growth was analysed using PB. For kill curves to determine precise compound concentrations to be used in shaken culture, or to determine the stability of evolved cell lines, cells were seeded at 0.2 x $10^6$ in 550 µL total volume in 96 deep well plates with test compound(s) added at the previously determined concentrations. A 100 µL sample was taken for PB analysis to better define seeding density. After 48-72 h shaken growth (as section 2.1.2.2) a further 100 µL sample was taken so as to calculate cell growth. All kill curve concentrations were repeated at least in duplicate, in triplicate where plate space allowed. All kill curves contained untreated control cultures and were designed with at least two concentrations at which all cell growth failed in an attempt to produce a sigmoidal kill curve so as to confirm a compound's 100% lethal dose ($LD_{100}$) concentration. Media-only and media/compound wells were compared to ensure compounds did not impact upon PB analysis

### *2.3.2.1. Determination of lethal dose values*

To determine LD values, linear regression of IVCD fold change from untreated control cells was performed using Prism 7 for Mac OS X. Extreme data points from non-responsive low and fully-responsive high compound concentrations were

removed to leave only the linear responsive concentrations of the compound. As such linear regression was performed only on the region of the kill curve where a response to changes in compound concentration is seen, so as to not skew data with flatter areas of the kill curve. Regression analysis produced a linear formula from which change in growth rate from the untreated control cells (LD0) could be used to produced compound concentrations that resulted in further LD values. Through the formula, the compound concentration resulting in a 0.75-fold change in IVCD (i.e. a 25% IVCD reduction compared to untreated cells, effectively killing 25% of the cell population) was used to calculate the LD25 value, the concentration resulting in a 0.50-fold change in IVCD (a 50% IVCD reduction) the LD50 value and the concentration resulting in a 0.25-fold change in IVCD (a 75% IVCD reduction) the LD75 concentration.

## 2.4. Plasmid DNA amplification and preparation

Model Mab molecules were used throughout the project. Plasmid vectors encoding the model Mabs were used to transiently and stably transfect CHO cells so productivity levels of cell lines could be ascertained. Accessory genes with which the CHO secretory and biosynthetic pathways were engineered were also expressed from plasmid vectors.

### 2.4.1. Plasmid DNA design

All plasmid DNA used during the project was inserted into a proprietary expression vector provided by MedImmune. This vector always contained a GS gene as a mammalian selection marker, expression of which was controlled by a weak Simian vacuolating virus 40 (SV40) promoter to enhance gene amplification in stable line production. An Ampicillin (Amp) resistance gene was also present as a bacterial selective marker. A summary of the plasmids used throughout the project is shown in table 2-3. A diagram of the plasmids used is shown in Figure 2-3.

All model Mabs and accessory genes were expressed from the recombinant backbone vector. Two recombinant Mab molecules, one difficult to express (DTE) and one easy to express (ETE), were used throughout the project, in a range of vectors containing a variety of extra elements to improve gene transcription and plasmid maintenance within the cell. In all cases the heavy and light chains of the recombinant Mabs were expressed under the control of a strong Cytomegalovirus (CMV) promoter. Model Mab genes were used in the standard expression backbone for both stable and transient expression studies. For some stable-expression experiments, the same ETE and DTE Mabs were also expressed in a backbone vector that contained a transcriptional enhancing element (TEE)

ensuring more replicable transcription between transfectants (Saunders et al. 2015).

Long-term transient studies (between 4 - 10 days in length) were carried out using fed-batch culture of the MedI-Tran long-term transient cell line. Long-term transient expression was run from a CMV promoter in a standard expression plasmid containing an Epstein-Barr virus-derived OriP gene that was necessary and sufficient for plasmid retention and replication and therefore long-term expression (Daramola et al. 2014; Gahn and Schildkraut 1989; Hatton et al. 2010).

**Table 2-3. Plasmids used during project.**

| Plasmid name | Main gene expressed (under CMV) |
|---|---|
| Backbone vector | 'Empty' Backbone vector (no gene under control of CMV) with GS under SV40. |
| Standard ETE | Backbone vector expressing ETE Mab |
| ETE-TEE | Backbone vector expressing ETE Mab, also containing a transcriptional enhancing element (TEE). |
| Standard DTE | Backbone vector expressing DTE Mab |
| DTE-TEE | Backbone vector expressing DTE Mab, also containing a TEE. |
| DTE-OriP | Backbone vector expressing DTE Mab under CMV, also containing OriP for use in long-term transient culture (MedI-Tran cells). |
| Accessory-gene expressing plasmid | Backbone vector expressing range of accessory genes. Listed in Appendix 2. |
| pMAXGFP™ | Lonza pMAX™ vector expressing max GFP. |

*The plasmids used during the project are listed, with a brief summary of the main gene they were used to express, as well as brief information on any further elements present within the vector.*

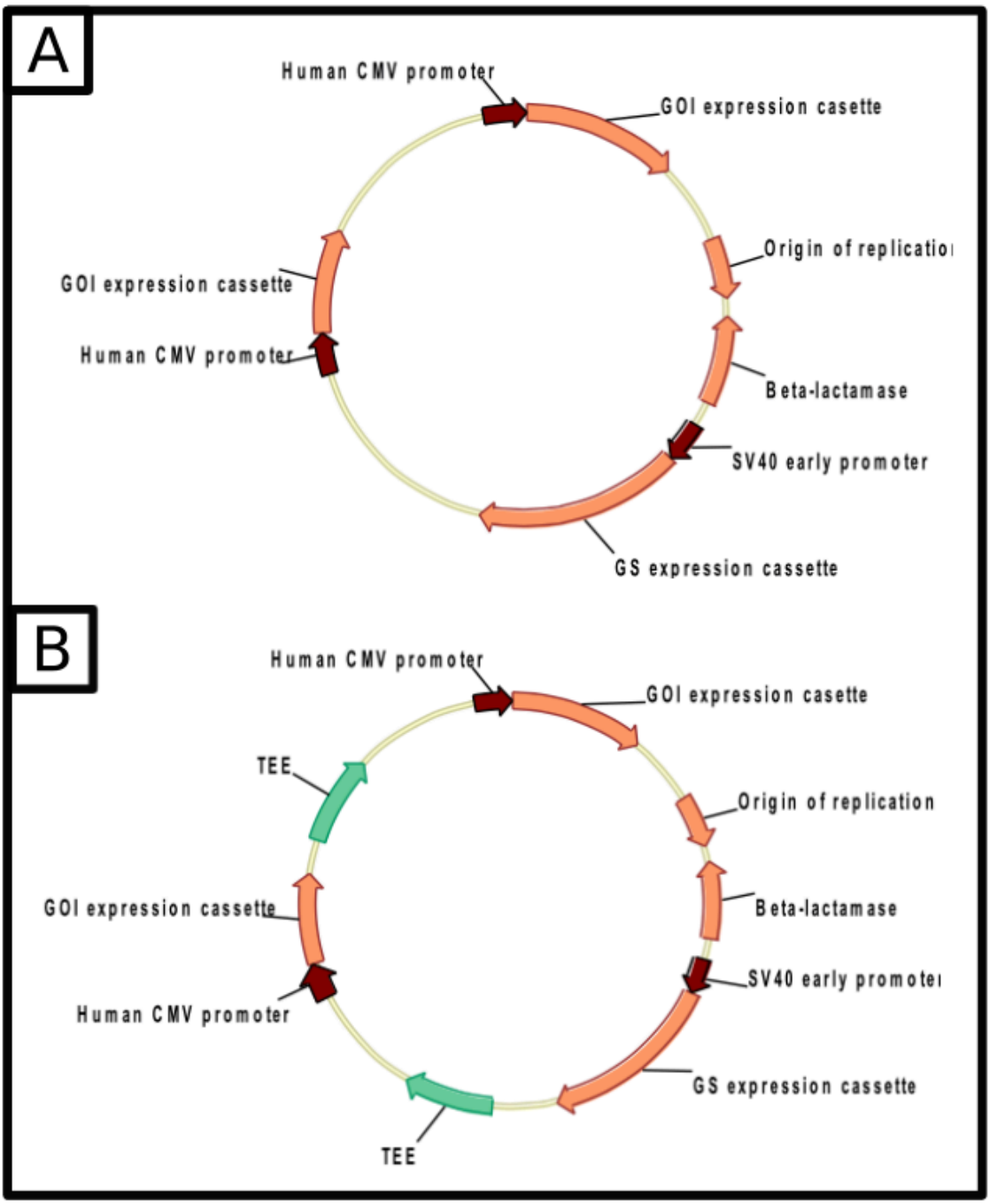


**Figure 2-3. Plasmid maps of backbone vectors used.**

*The plasmid maps of the backbone plasmids used during the course of the project are shown. In both plasmids two Gene of Interest (GOI) locations are shown under the control of a human CMV promoter. The GOI locations could include Mab or effector genes. The GS expression cassette is shown under a weaker SV40 promoter. [a] The standard backbone was used predominantly for transient expression, though some stable transfections were carried out with this plasmid. [b] this plasmid contains a transcriptional enhancement element (TEE) and was used for the production of stable pools with better control of transcriptional effectiveness. Note that image is not to scale.*

## 2.4.2. Plasmid DNA amplification

Circular plasmid DNA was used to transform DH5α competent *Escherichia coli* cells (New England Biolabs [NEB], Hitchin, UK). Growth in shaken culture amplified the plasmid DNA. Plasmid purification was carried out using Maxi prep kits.

### *2.4.2.1. Bacterial growth media*

Shaken culture of bacterial cells was carried out in liquid Luria Bertani (LB) broth (Fisher; Miller recipe: 1% Tryptone, 0.5% Yeast Extract, 1% NaCl; pH 7.0 ± 0.2). Shaken culture was carried out in a shaken incubator (Multitron, Infors) at 37 °C and 200 rpm in non-aerated vessels with at least 60% flask volume being head-space. Bacterial colonies were produced on solid LB-agar (LB broth as above, with 1.5% agar; Fisher). All media was dissolved in di$H_2O$ and sterilised by autoclaving at 121 °C for 15 mins. After cooling to <55 °C media was supplemented with requisite antibiotic selection agent for growth of bacteria transformed with recombinant plasmid DNA, normally Ampicillin (Amp, 100 μg/mL) or Kanamycin (Kan, 50 μg/mL) which was added from a 1000x stock solution (Sigma). Liquid LB-agar media containing antibiotic was poured in to 100 mm diameter petri dishes and allowed to solidify at room temperature. LB-agar plates were stored at 4 °C with an expiry date of 1 month. Prior to use plates were warmed at 37 °C.

### *2.4.2.2. Transformation of E. coli with plasmid DNA*

Competent DH5α *E. coli* cells were stored at -80 °C and thawed on ice for 10 mins immediately prior to use. Plasmid DNA (1 μg, approximately 1 μL total volume) from a plasmid stock solution was added to a 50 μL aliquot of cells, gently mixed by light flicking and incubated on ice for 10 mins. DNA was introduced to the bacteria by a heat-shock in a water bath (42 °C for 20 - 40 seconds) before a further 2 minute incubation on ice. Sterile un-supplemented LB broth (100 μL) was added to each aliquot of cells for an out-growth step in a shaking heat block for 1 h at 37 °C, 300 rpm. Transformed cells were added to LB-agar plates (one 20 μL and one 100 μL aliquot) containing selective pressure and aliquots spread with a sterile plate spreader. Single-cell derived colonies were produced by incubating plates upside-down at 37 °C. Colonies were used to inoculate shaken culture. Colony plates were stored at 4 °C, with colonies used within two weeks of transformation and storage.

### *2.4.2.3. Amplification of transformed E. coli*

Single cell colonies were used to produce a shaken starter culture. A single colony from an agar plate was used to inoculate 5 mL LB broth (with antibiotic) in a 50 mL Falcoln tube via a sterile loop. After 8 h growth in shaken conditions, 1 mL of the starter culture was used to inoculate 100 mL LB broth in an e250 Erlenmeyer flask. Amplification was carried out in shaken culture overnight (up to 17 h) at which point cells were harvested.

#### *2.4.2.4. MaxiPrep of amplified Plasmid DNA*

Plasmid DNA was extracted from *E. coli*, cleaned and purified using a Plasmid Plus MaxiPrep kit (MaxiPrep; Qiagen, Manchester, UK) according to the included protocol. The QIAvac 24 was used to allow high-throughput preparation of multiple plasmids. *E. coli* cells were harvested by centrifugation at 6000 rpm for 15 mins. Cells were resuspended, lysed and the lysate cleared before a bind-wash-elute procedure using the QIAvac apparatus. DNA was eluted and stored in either di$H_2O$ or the provided elution buffer (EB; 300 µL of 10 mM Tris-Cl, pH 8.5). DNA was stored in a 1.5 mL microfuge tubes at 4 °C (short-term) or -20 °C (long-term).

#### *2.4.2.5. Measurement of DNA concentration and purity*

Plasmid DNA concentration and purity was measured using a NanoDrop™ One$^C$ (ThermoFisher) with the buffer containing the DNA used as a blank. For each plasmid readings were taken in triplicate and an average value determined. Purity was measured using A260/A280 and A230/A260 ratios, with values of approximately 1.8 and 2.0-2.2 respectively indicating pure DNA with minimal contamination. Samples not in this range were discarded or underwent further purification.

#### *2.4.2.6. Checking of plasmid DNA by restriction digest*

Plasmid DNA was checked by restriction digest and compared to plasmid maps to ensure the correct plasmids were used. Single- or double-enzyme restriction digests were carried out on purified DNA using High Fidelity (HF) restriction endonucleases (NEB). The enzymes used depended upon the plasmid being digested, with EcoRI, BamHI, AgeI and SbfI the most common enzymes used. Restriction digests (20 µL final volume) were prepared containing 10 units of restriction endonuclease (not exceeding 10% total reaction volume) to digest approximately 1 µg plasmid DNA. The reaction mix contained 10% 10x CutSmart™ buffer (NEB), with the final volume made up with di$H_2O$. Digests were incubated at 37 °C for 2 h. The digest reaction was halted by addition of 3.33 µL 6x purple loading dye (NEB). Samples (10 µL) were loaded in to wells of an analysis gel (1.0% Agarose dissolved by microwaving in 50 mL 1x Tris, acetate, EDTA [TAE; Gibco] buffer with 0.5 µL 10000x SYBR™safe DNA gel stain added (Invitrogen; added once agarose had cooled but before gel poured). A nucleotide ladder (HyperLadder™, BioLine, London UK) was also loaded on to the gel to determine the size of DNA fragments, with a blank lane and un-digested plasmid run as a negative control. The loaded gel was run in a gel tank, containing 1x TAE buffer, at 100V for 30 mins to separate DNA bands based on size. Gels were visualised on a Benchtop UV Transilluminator (UVP, AnalytikJena, Jena, Germany) and images taken using a Biospectrum® 410 imaging system (UVP) using VisionWorks LS software (UVP).

### 2.4.3. Preparation of DNA for transient transfection of mammalian cells

Transient transfection was carried out using circular plasmid DNA. Plasmid DNA was diluted to the required concentration in nuclease free $H_2O$ (Qiagen) and concentration checked again by NanoDrop™ to ensure DNA loads were correct.

### 2.4.4. Preparation of DNA for stable transfection of mammalian cells

Stably expressing cell lines and pools were produced by transfection of CHO cells with linearised plasmid DNA. DNA used originated from the same stocks as that used for transient transfections.

#### *2.4.4.1. Linearisation of Plasmid DNA*

Plasmid DNA was linearised by restriction digest using PvuI-HF enzyme (NEB) which cut each model Mab plasmid at one restriction site situated in the bacterial resistance-encoding gene to not impact upon transcription of the recombinant genes. Gene sequences were designed so as to not contain the Pvu1 restriction site. Plasmid DNA (100 µg) was cut with 500 units of restriction endonuclease in a 500 µL digest containing 10% CutSmart™ buffer with a maximum 10% enzyme mix by volume. The digest was incubated at 37 °C for 2 h before being stopped by incubation at 4 °C. Linearisation was confirmed by running an aliquot of linearised DNA (mixed with 6x purple loading dye) on a 1% agarose gel (see section 2.4.2.6). An aliquot of non-linearised plasmid was run as a negative control.

#### *2.4.4.2. Cleaning of linearised DNA*

To prepare linearised plasmid DNA for transfection it underwent cleanup to remove salts and endonucleases used in the linearisation process, using Phenol-Chloroform extraction in conjunction with EtOH precipitation.

**Phenol-Chloroform extraction**. Linearised DNA (max 500 µL) was transferred to a pre-spun Phase Lock Gel tube (PLG; 5 Prime, QuantaBio, Beverley, MA, USA) and an equal volume of organic extraction solvent (phenol:chloroform:Isoamyl alcohol 25:24:1; Invitrogen) added and mixed by rapid inversion. Phases were separated by centrifuging at 13,000 rpm for 5 minutes. The nucleic acid containing top phase was transferred to a fresh PLG tube and an equal volume of chloroform (VWR, Lutterworth, UK) added, mixed by inversion and phases separated by spinning at 13,000 rpm for 5 minutes. The top nucleic acid containing layer was transferred to a clean microfuge tube for EtOH precipitation.

**Ethanol Precipitation.** Two volumes of 100% EtOH (-20 °C; VWR) and 3 M sodium acetate (0.1 volumes; Sigma) were added to precipitate linearised DNA, mixed by inversion and incubated at -20 °C for 30 minutes. DNA was pelleted at 13,000 rpm for 30 mins at 4 °C and the supernatant removed. The remaining DNA pellet was washed in 70% EtOH (-20 °C) and briefly vortexed to ensure uniform washing and centrifuged (13000 rpm, 10 mins, 4 °C) to re-pellet. Remaining preparation steps were performed in a sterile environment. Supernatant was

removed and the DNA pellet air-dried for 5-10 minutes. DNA was resuspended in 50 µL tissue culture (nuclease free) water to give an initial concentration of approximately 2 µg/µL. Aliquots were taken for analysis on 1.0% Agarose gel and by NanoDrop™, after which DNA volume was adjusted to give a final concentration of 0.4 µg/mL, ready for transfection for production of stable pools.

## 2.5. Transfection of CHO cells with recombinant DNA

Transient and stable transfections were used to test genetic engineering strategies and assay productivity levels of evolved and genetically manipulated cell lines. Recombinant plasmid DNA encoding model Mabs, engineering target effector genes or green fluorescent protein (GFP), were introduced to MedI-CHO- and MedI-TRAN–derived cell lines by electroporation-mediated transfection, using the Nucleofection platform (Lonza, Basel, Switzerland). Both cuvette and Amaxa 96-well Shuttle™ nucleofection technologies were utilised though the course of the project to produce cell lines expressing recombinant Mabs both transiently and stably. Polyethylenimine (PEI)-mediated transfection was trialed, as per protocols previously developed within the laboratory (Mozley et al. 2014; Thompson et al. 2012), but resulting transfection efficiency results were not satisfactory.

For all transfections cells were taken on day three of culture, early exponential phase, with a VCD of approximately 1.0 x $10^6$ cells $mL^{-1}$ and a viability above 95%. MSX and/or any other selection/evolutionary agents used in cell culture were removed from cell culture two passages prior to transfection to prevent any negative effects upon transfection efficiency. The following describes the protocols used for transient transfection of MedI-CHO (section 2.5.1) and MedI-TRAN and stably expressing (section 2.5.2) cell lines. Production of stably-expressing pools from MedI-CHO-derived cell lines is also described (section 2.5.3). Plasmid DNA used is listed in table 2-3.

### 2.5.1. Transient transfection of CHO cells – cuvette system

Transient transfection of CHO cells with plasmid DNA was performed by an electroporation protocol used previously within the laboratory (Johari et al. 2015), based upon the original manufacturer's protocol for suspension cells (Lonza 2008). This protocol was optimised for electroporation of CHO-S cells which grow faster than the MedI-CHO lines. As such some changes were made to account for the reduced growth of the cell lines used.

Transfection was conducted using the Amaxa™ Cell Line Nucleofector™ Kit V system (Lonza). The passage prior to transfection, cells were seeded as normal at 0.2 x $10^6$ cells $mL^{-1}$, with transfection being carried out on day 3 of culture. For transfection 4.5 x $10^6$ cells per cuvette were centrifuged at 130 x g for 5 mins, supernatant removed and the cell pellet resuspended in 100 µL of Nucleofector solution. Plasmid DNA (4.6 – 5.0 µg) was added and gently mixed and the cell-DNA

mix transferred to the cuvette and transfected with the Amaxa nucleofector, using programme U-024. Immediately after electroporation cells were further diluted in 600 µL pre-warmed culture medium and transferred to 10 mL pre-warmed and pre-gassed culture medium in a TPP Tubespin cultiflask, at an approximate seeding density of 0.4 x $10^6$ cells $mL^{-1}$. After 1-2 h of shaken incubation ViCell samples were taken to determine an accurate seeding density and cell viability. After 4 days' growth, cell viability and growth measurements were taken by ViCell. Cells were harvested by centrifugation at 130 x g for 5 min and supernatant samples taken for titre analysis supernatant cells. Biological and technical replicates of transfections were carried out alongside each other using the same protocol.

#### *2.5.1.1. Determining transfection efficiency: Separate GFP transfection*

Transfection efficiency was determined when new transfection protocols were used, or new cell lines being transfected, to ensure transfection repeats were comparable. To determine transfection efficiency, transfection with plasmid DNA encoding GFP (pMAXGFP™, Lonza) was performed alongside experimental transfections, using the transfection protocol described previously (section 2.5.1). After four days' growth, cell samples (at least 1 x $10^6$ cells in 1 mL growth media or PBS) were taken for analysis with an Attune® acoustic focusing flow cytometer (Applied Biosystems, Life tech, ThermoFisher). GFP was excited with a 488 nm laser, with emission detected at 530 nm to determine the percentage of the cell population that was GFP positive (therefore shown to have been successfully transfected with and expressing GFP) to determine a transfection efficiency.

#### *2.5.1.2. Determining transfection efficiency: Co-transfection with GFP*

To avoid excessive transfection and cell culture, where many different cell lines were transiently transfected, pMAXGFP™ was co-transfected alongside effector gene DNA. Due to the size difference between Mab-encoding plasmid DNA (~9500 bp) and pMAXGFP (3486 bp), effector and GFP genes were co-transfected at a 1:1 molar ratio which was calculated by dividing the size of the Mab-expressing plasmid (in bp) by that of the pMAXGFP plasmid. The two plasmids were combined in the 1:1 ratio, with a total DNA mass of 4.6 µg. Transfection was carried out as in section 2.5.1. After four days' growth samples were taken for analysis by flow cytometry as in section 2.5.1.1.

### 2.5.2. High Throughput Transient transfection of CHO cells

The Amaxa 96-well shuttle™ add-on device (Lonza) was used to perform up to 96 nucleofections in parallel, using a protocol based upon the manufacturer's suggested protocol for CHO-S derived cell lines (Lonza 2009) and further developed within the laboratory group resulting in transfection efficiency (determined by GFP expression) of above 95% (Arnall et al., manuscript in preparation). This platform was used to carry out gene screen experiments to determine the effect of varying levels of different genes, both singularly and in

combination, upon CHO cell productivity. Each gene dose permutation was repeated three times within a plate to give three biological replicates, with each plate repeated two-three times to produce technical replicates. Transfected cells were cultured in shaken square well, V-bottomed 96-well deep well plates (Greiner Bio-One) as described in section 2.1.2.2, with any changes to this protocol specific to 96-well transfection listed here.
For high-throughput gene screens the host cell (using either MedI-Tran for transient Mab expression, or a MedI-CHO-derived DTE-Mab-expressing stable cell pool), for each well 1.86 x $10^6$ cells were transfected with a total of 800 ng plasmid DNA, comprising recombinant Mab, effector and empty plasmid DNA (table 2-3). Prior to transfection, 500 µL of growth media (un-supplemented CD-CHO) was transferred to each well of 96 deep well plate, covered with a sterilised 'Sandwich Cover' lid and transferred to a shaking incubator for equilibration at 37 °C, 5% $CO_2$, 320 rpm for a minimum duration of 1 h.

DNA and cells were prepared and combined with an excess volume of 25% to allow for easier pipetting and to ensure the correct volume was used throughout. Plasmid DNA was mixed with nucleofector solution mix (a 4.5:1 ratio mix of SG Cell Line 96-well Nucleofector™ Solution:Supplement [Lonza] which were combined on the day of use) in a sterile round-bottom 96-well plate (Nunc). To avoid evaporation of small quantities of DNA, 7.5 µL nucleofector solution mix was initially added to separate wells of the round bottomed 96-well plate in the same layout of wells to be transfected. Plasmid DNA, suspended in nuclease free $H_2O$, was added to the nucleofection solution, each well containing 1000 ng total DNA (800 ng + 25% excess) in a final volume of 10 µL DNA:nucelofector solution mix per well. The final volume was made up with nuclease-free $H_2O$.

For each transfection well used, 2.33 x $10^6$ cells were centrifuged at 130 RCF for 5 mins and all supernatant removed by pipetting. For each transfection well used, the cell pellet was resuspended in 15 µL pre-mixed and pre-warmed nucleofector solution. At this stage a sample of the resuspended cells was taken to check cell density and viability were approximately the expected values.

To each well of the DNA-containing plate, 15 µL of resuspended cells was added and mixed by pipetting. Of the total 25 µL volume (15 µL cells, 10 µL DNA mix), 20 µL was transferred to the 96-well Nucleocuvette™ plate (Lonza). Bubbles were carefully removed using fine sterile needles (Microlance™, BD BioSciences, Oxford, UK) to reduce electroporation error through arcing of charge. The plate was electroporated within the Nucleofector™ 96-well Shuttle™ add-on device, using program FF-158 (CHO-S optimised). Immediately post-transfection, 80 µL pre-warmed culture medium was added to each well. The viability of three wells, selected at random, was assessed to determine an average post-transfection viability and VCD. All wells were mixed and an average of the three VCD values

was used to calculate the amount of transfected cells required to seed 575 µL growth media at 0.4 x $10^6$ cells $mL^{-1}$. The final volume made up to 575 µL per well with pre-warmed media.

Immediately after seeding, 100 µL of cell suspension from each well was transferred to flat-bottomed 96-well plates (Nunc) for PB analysis (section 2.1.4.1) to determine individual seeding density values for each well. The remaining 475 µL of cell culture was incubated as described previously, with a feed undertaken at 72 h of culture. End point cell growth and supernatant harvest samples were taken at 120 h post-seeding (as described in section 2.1.2.2).

### 2.5.3. Stable transfection of CHO cells – stable pool production

Stable pools were produced to determine the long-term productivity of different cell lines, especially to compare evolved/filtered cell lines with the control host cell line. Stable pools were created by electroporation (nucleofection) of CHO cell lines with linearised plasmid DNA. The stable transfection protocol is similar to that used for transient nucleofection (section 2.5.1) and was provided by MedImmune. Before transfection cells were grown from WCB for between 4-10 passages, the passage immediately prior to transfection lasting no more than three days with cell viability exceeding 95%. Plasmid DNA encoding the recombinant protein of interest and a MSX resistance gene was linearised by restriction digest with PvuI restriction endonuclease (section 2.4.4.1) and cleaned and resuspended in sterile nuclease free water (section 2.4.4.2), with plasmid linearisation checked on 1% agarose gel against untreated plasmid controls. Electroporation was conducted using the Amaxa™ Cell Line Nucleofector™ Kit V system. For each transfection event, 10 x $10^6$ cells were transfected with 6 µg linear DNA at a concentration of 0.4 µg/µL (15µL total DNA volume).

Cells (10 x $10^6$ per transfection) were harvested by centrifugation at 130 x g for 5 mins, with all supernatant removed by decanting and gentle pipetting so as to not disrupt the cell pellet. The dry cell pellet was resuspended in 100 µL of a pre-mixed Kit V nucleofector solution, in a 4.5:1 ratio of solution:supplement (Lonza). The resuspended cells were added to 15 µL (6 µg) linearised plasmid DNA in a sterile microfuge tube and mixed gently by pipetting. Any bubbles were carefully removed from the cuvette by lancing with a pipette tip or sterile needle (Microlance™). The cell/DNA mix was added to the nucleofection cuvette (Lonza) and electroporated using program U024 (for neonatal cells).
Transfected cells were immediately removed from the cuvette to 41 mL pre-warmed (37 °C) and pre-gassed media in an e125 flask, using the provided narrow dropping pipette (Lonza). A small amount of media was taken from the flask to further rinse the cuvette and ensure as much cell matter as possible was transferred to culture. The resuspended cells were incubated as described in section 2.1.2.1 for up to 1 h. A ViCell sample (1 mL) was taken to determine cell

viability, with the remaining 40 mL being split across 8 x 5 mL cultures in T25 tissue culture flasks (Corning). Transfectants were incubated in a static incubator as section 2.1.3. At 24 h post-seeding, MSX selective pressure was added at a final concentration of 50 μM, with each culture volume increased to 6 mL. Pool growth was maintained for up to 4 weeks, with culture progression checked regularly from ten days post transfection onwards.

As stable pools reached confluency (that is, cell culture became turbid due to cell growth), culture density and viability was measured by ViCell (section 2.1.4). Once cell density had exceeded 1.0 x $10^6$ cell $mL^{-1}$, an extra 2-5 mL MSX growth media extra (exact amount depending on cell density) was added and cells grown for a further two days, when cells were transferred to shaken culture at an initial seeding density of 0.2 - 0.4 x $10^6$ in 20 - 30 mL 50 μM MSX growth media in e125 flasks (as section 2.1.2). Initial shaken culture passages often extended from 3-7 days' duration before cell growth reached a level where progression could occur, as cells adapted to shaken culture. Once cell growth rate and viability had stabilised, cells were passaged as normal every 3 - 4 days. Cell banks of each stable pool were produced after 3 - 4 passages (roughly 10 generations) in shaken culture, when growth levels had returned close to the host level (section 2.1.5). Of the initial 8 pools produced from each transfection, growth and productivity analysis (where capacity allowed it) were used to select one pool to take forward. Each transfection was repeated three times to produce biological replicates.

## 2.6. Measurement and calculation of Mab titres

All model recombinant Mabs used were from the Immunoglobulin G (IgG) family. An IgG Mab-specific Valita™TITER fluorescence polarisation (FP) assay (Valitacell™, Dublin, Ireland) was used to measure concentrations of Mab produced by Mab-expressing CHO cells (Thompson et al. 2017). ValitaTiter (functional range of 2.5-100 mg/L) plates were used to measure Mab levels from cells stably expressing recombinant Mab. ValitaTiter High Sensitivity (HS) plates (functional range of 0-6 mg/L) were used to measure Mab levels from cell lines transiently expressing a recombinant Mab. FP values were calculated from parallel and perpendicular polarisation reads using Equation 7.

**Equation 7 –** calculating Fluorescence Polarisation for Valita™TITER assay**.**

$$FP = \frac{(F_{\parallel} - F_{\perp})}{(F_{\parallel} + F_{\perp})}$$

where $F_{\parallel}$ = fluorescence intensity parallel to the excitation plane, and $F_{\perp}$ = fluorescence intensity perpendicular to the excitation plane (Thompson et al. 2017).

### 2.6.1. Standard curve preparation

Two IgG standards were used depending upon the Mab being assayed for. A Kappa light chain (LC) standard was used for ETE model Mab (which contained a Kappa LC). A Lambda LC standard was used for the model DTE Mab (which contained a Lambda LC). Both standards were provided by Sigma. For a standard curve between 0-80 mg/L, an initial dilution of the standard was made in growth media to produce a concentration of 80 mg/L in a 1 mL volume. Six serial dilutions (500 µL initial standard mixed with 500 µL plain media, and so on) were carried out, with each standard well mixed by vortex and pipetting up and down with a clean tip, resulting in a standard curve with values at 80, 40, 20, 10, 5, 2.5 and 1.25 mg/L (ValitaTITER 2018a). For HS plates, a standard curve between 0-6 mg/L was produced. An initial dilution in growth media was produced to a concentration of 6 mg/L in 1 mL. Serial dilutions were carried out, as described above, resulting in a standard curve with values at 6, 3, 1.5, 0.75, 0.375 and 0.1875 mg/L (ValitaTITER 2018b).

For both curves a media blank was used as a 0 mg/L concentration and to normalise between plates. Standards were repeated in at least duplicate (2 wells per standard), with triplicate technical repeats (3 wells per standard) used where possible. When using multiple Valita™TITER plates across the same experiment, plates and buffer from the same batch were used, coupled with the same plate reader settings. This allowed only one standard curve could be used across the entire experiment instead of a standard curve on each plate used.

### 2.6.2. Sample preparation

Samples were harvested from cell culture using centrifugation as described in section 2.1.7. All comparative samples underwent the same amount of freeze-thaw cycles so they could be compared directly. Prior to analysis, samples were diluted so as to fit within the functional range of the ValitaTiter plate and the standard curve (section 2.6.1). All samples underwent at least a 1:1 dilution with media so as to avoid matrix effects (ValitaTITER 2018a; ValitaTITER 2018b). Dilution levels were based on culture production levels determined from previous experiments within the laboratory, depending upon whether the recombinant Mab was ETE or DTE, and whether it was being expressed transiently or stably. Frozen samples were thawed shortly before analysis was performed. All diluted samples were prepared in a shallow 96-well plate before analysis, with at least one well per plate containing a blank CD-CHO-only well.

#### *2.6.2.1. Preparation of samples from shake flask scale culture*

Samples from 10 - 100 mL cell cultures were stored in microfuge tubes at -20 °C. Samples were diluted at the required level in plain CD-CHO to a 300 - 500 µL final volume in a separate microfuge tube. The dilution was split in to three equal aliquots (100 µL for normal ValitaTitre plates; 150 µL for HS plates), each in an adjacent well of a shallow 96-well plate.

#### *2.6.2.2. Preparation of samples from high-throughput small-scale culture*

After harvesting samples were transferred from a 96-well storage plate to a dilution plate. Samples were diluted by the required amount in CD-CHO to a final volume of 100 µL (normal Valita™TITER assay) or 150 µL (Valita™TITER HS assay) in a shallow 96-well plate. Dilutions were mixed by pipetting with a multichannel pipette.

### 2.6.3. Valita™TITER analysis of supernatant samples

All ValitaTiter analysis was carried out according to the provided instructions for use (Thompson et al. 2017; ValitaTITER 2018a; ValitaTITER 2018b). Valita™ buffer (60 µL/well; containing Acetate, Bovine Serum Albumin [BSA], NaCl and 0.5 g/L Sodium azide) was added to each well of the assay plate. Using a multichannel pipette and clean tips, 60 µL of sample was transferred from the sample preparation plate to the assay plate. Once all samples had been added, all samples were mixed by pipetting up and down 100 µL six times with a multichannel pipette. The plate was incubated at room temperature in darkness for 30 mins before being read on a PHERAstar® plate-reader using specific Valita™TITER software to measure fluorescence polarisation. Gain settings were set from the well with the expected highest Mab levels (the 80 mg/L standard if used).

#### *2.6.3.1. Valita™TITER HS analysis*

For Valita™TITER HS analysis the analysis plate was prepared as described in section 2.6.3. To further remove background noise at the lower concentrations being measured, a blank plate not containing probe was prepared as per the read plate. This plate did not require a 30 mins incubation before being read.

### 2.6.4. Calculation of titre values

The standard curve (section 2.6.1) was plotted using Excel (Microsoft, Redmond, WA, USA). The known standard concentration was plotted on the *x* axis against the FP values produced by the PHERAstar® plate reader on the *y* axis. A line of best fit was fitted to the standard curve, with the *y*-intercept set to zero. Any extraneous points (those at the bottom of the curve being the most likely to be affected by noise) were removed and the line of best fit used if an $r^2$ value above 0.95 was produced. The produced equation for the line of best fit was used to convert known FP values to IgG concentrations (mg/L).

#### *2.6.4.1. Calculation of titre values from Valita™TITRE HS data*

To calculate IgG concentrations from ValitaTiter HS data, the values from the blank plate had to be taken in to account (section 2.6.3.1) by subtraction of the blank plate values from the read plate values. The parallel fluorescence value from the blank plate was subtracted from the parallel fluorescence value from the read plate to produce a new parallel fluorescence value. The perpendicular fluorescence value from the blank plate was subtracted from the perpendicular fluorescence value from the read plate, resulting in a new perpendicular

fluorescence value. These two new parallel and perpendicular fluorescence values were used in Equation 7 to calculate the correct FP values which could then be used in the standard curve equation.

## 2.7. Cell morphology analysis

To quickly visualise any changes within the CHO secretory pathway that may result in a change in secretory/productivity phenotype, antibodies raised against targets present within the mammalian secretory pathway were used to probe the cell. Immunofluorescence microscopy (IF) was used to visualise changes in organelle morphology and target and product distribution within the cell. Western blotting (WB) was used to approximately visualise abundance levels of components within a cell population. Flow Cytometry (FC) was used to better quantify expression levels of secretory pathway components and recombinant proteins across an entire cell population.

### 2.7.1. Immunofluorescence microscopy (IF)

CHO cell samples were centrifuged on to 1 cm diameter coverslips (Deckgasser, VWR) before being fixed, permeabilised, stained with primary (1°) and secondary (2°) antibodies, mounted on to a slide and imaged by epifluorescence.

#### *2.7.1.1. Mounting cells to coverslip*

Due to not being adherent, suspension-adapted CHO cells had to be mounted on coverslips before staining. Sterile coverslips were transferred to the wells of a 12-well cell culture plate (Nunc) using EtOH-sterilised tweezers. Each slide was flooded with a minimum of 500 µL 0.01% Poly-L-lysine (PLL) solution (w/v, diluted from 0.1% solution in di$H_2O$; Sigma) and incubated at 37 °C for >15 mins. PLL solution was removed by aspiration and coverslips washed in 1 mL 1x PBS which was removed by aspiration. Coverslips were left to air-dry in a sterile environment. Coated cover slips were stored in the dark at room temperature for up to 6 months.

CHO cells (1 - 2 x $10^6$ cells) were centrifuged on to PLL-coated coverslips at 200 x g for 5 min before incubation at 37 °C with 5% $CO_2$ for >1 h. During this incubation step any compounds to be tested upon cells could be added to the cells in growth medium at the required concentration. After an hour incubation cells were again centrifuged at 200 x g to ensure adherence to the coverslip prior to the fixing/staining procedure.

#### *2.7.1.2. Cell fixation*

Cell fixation was performed on both cover-slip mounted and non-mounted (pelleted) cells to allow storage of cell samples prior to analysis, as well as to allow accurate staining of cells. For mounted cell samples, fixing/washing solutions were removed after each step by aspiration. For cell pellet samples, fixing of 1 - 2 x $10^6$ cells was carried out in microfuge/15 mL Falcoln tubes. After

each step cells were pelleted by centrifugation at 200 RCF for 5 mins and solutions removed by decanting and careful pipetting to remove as much as possible without disturbing the cell pellet. Depending upon the antibody to be used, cells were either fixed by treatment with 3.7% Paraformaldehyde (PFA) fixative solution in 1x PBS (Alfa Aesar, Haverhill, MA, USA) or chilled absolute Methanol (MeOH; ThermoFisher).

**PFA fixation.** Growth media was removed from the cell pellet/mounted cells which were then washed with 1x PBS (1 mL). Fridge-cold fixative solution was added with cells at a concentration of 1 x $10^7$ cells $mL^{-1}$ PFA solution and cells incubated at 4 °C for 15 minutes. Cells were not fixed for longer than this so as to avoid over-fixation that would impinge upon cell staining. PFA solution was removed and, for slide-mounted cells to be antibody stained for immunofluorescence, 500 µL 0.1 M Glycine solution in 1x PBS added to quench the fixative before removal and rinsing and storage in 1 mL 1x PBS. If not stained immediately, PFA-fixed cells were stored at 4 °C for <6 months before use.

**Methanol fixation.** MeOH (100%) was pre-chilled in 1 mL aliquots in a 12-well cell culture plate (Nunc) at -20 °C for >20 mins. Growth media was removed from pelleted/slip-mounted cells which were washed with 1 mL 1x PBS. Cells were transferred to the pre-chilled MeOH. Cells were incubated at -20 °C for 5 minutes before MeOH was removed and cells rinsed and resuspended in 1 mL 1x PBS.

#### *2.7.1.3. Cell permeabilisation and staining*

Cells were permeabilised using a saponin-FBS permeabilisation and staining solution, with Fetal bovine serum (FBS) acting as a blocking agent (0.1% w/v saponin from Quillaja bark [Sigma]; 5% v/v FBS; dissolved in 1x PBS, sterile filtered through a 0.2 µM filter Stericup® [Millipore, Sigma] and stored at 4 °C). All staining protocols were carried out at room temperature (18-22 °C). Antibodies were stored at -20 °C and maintained at this temperature in a cold pot during use. Primary antibodies were tested by both IF and Western blotting to confirm efficacy in CHO cells prior to wider use within the project. Two primary antibodies were used in parallel if they were raised in separate organisms and therefore targeted with separate secondary antibodies.

PBS solution was removed from slide-mounted fixed cells and 300 µL (enough to flood the cover slip) of saponin-FBS solution added to each well to permeabilise cells. Saponin-FBS solution was removed and the primary antibody/antibodies added at the required concentration in 300 µL saponin-FBS solution and incubated at RT for 30 mins (See table 2-4 for antibody concentrations). The primary antibody solution was removed and cells washed with 300 µL saponin-FBS solution. The secondary antibody/antibodies, diluted at the required concentration in saponin-FBS solution, was/were added and incubated for 30

mins at RT. Secondary antibody solution was removed and cells washed a further two times with saponin-FBS solution to remove any excess secondary antibody before the slip was flooded with 1x PBS. A drop of RT ProLong™ Gold Antifade Mountant (Invitrogen™) - containing the nuclear DNA stain DAPI - was placed on a pre-labeled glass microscope slide (VWR). Cover slips were removed from PBS and dried by blotting before being placed, cell-side down, on the mountant and pressure gently applied to steady the clover slip and remove bubbles. Excess mountant was removed by aspiration and slides sealed with four small drops of clear nail varnish. Slides were stored in darkness at 4 °C.

*Table 2-4 (next page) shows all antibodies used to stain cells for Western blotting and immunofluorescence purposes through the course of the project.*

**Table 2-4. Antibodies used during course of project.**

| Primary Antibodies – Immunofluorescence and Western Blotting use. | | | | | |
|---|---|---|---|---|---|
| **Antibody** | **Dilution.** | **Target.** | | **Raised in** | **Manufacturer** |
| GM130 monoclonal | 1/1000 (IF) | Golgi marker GM130 | | Mouse | BD biosciences |
| Calreticulin monoclonal | 1/250 (IF) | ER marker Calreticulin | | Rabbit | Thermo Fisher |
| GBF1 polyclonal | 1/200 (IF); 1/2000 (WB) | BFA target GBF1 | | Rabbit | Abcam®, Cambridge, UK |
| BiP monoclonal | 1/2000 (WB) | ER stress response protein BiP | | Rabbit | Abcam® |
| Actin-b | 1/1000 | Cytoskeletal β-actin (housekeeping) | | Mouse | Abcam® |
| **Secondary antibodies – Immunofluorescence use.** | | | | | |
| **Antibody** | **Dilution/ concentration** | **Excitation max (nm)** | **Emission max (nm)** | **Emission colour** | **Manufacturer** |
| Anti-Mouse AlexaFluor 488 | 1/1000 (IF) | 496 | 519 | Green | Invitrogen |
| Anti-Rabbit AlexaFluor 594 | 1/750 (IF) | 590 | 617 | Red | Invitrogen |
| DAPI | n/a | 350 | 450 | Blue | Invitrogen |
| **Secondary antibodies – Western Blotting using Li-Cor system.** | | | | | |
| AlexaFluor 680 goat anti-rabbit IgG | 1/10,000 | 680 | 704 | Red | ThermoFisher |
| AlexaFluor + 800 goat anti -mouse IgG | 1/10,000 | 496 | 519 | Green | ThermoFisher |
| **Fluorescent tag conjugated Mab Heavy/Light chain antibodies.** | | | | | |
| Goat anti-Human IgG Fcγ-APC | 5 µg/mL | 650 | 660 | Red | Jackson Immuno Research (Cambridge, UK) |
| Kappa-FITC Goat anti-Human | 3 µg/mL | 495 | 519 | Green | Southern Biotechnology (Birmingham, AL, USA) |
| Lambda-FITC Goat anti-Human | 3 µg/mL | 495 | 519 | Green | |

#### *2.7.1.4. Imaging of stained cells*

Imaging work was carried out at the Wolfson Light Microscopy Facility, university of Sheffield Biomedical sciences department. Images were taken using a motorised Olympus upright epifluorescence BX61 system running a Hamamatsu Orca monochrome camera through Volocity imaging software (Olympus, Shinjuku, Tokyo, Japan). A FITC laser/filter combination was used to image cells stained with AlexaFluor 488-conjugated secondary antibodies. A TX-Red laser/filter combination was used to image cells stained with AlexaFluor 594-conjugated secondary antibodies. A DAPI blue/cyan filter was used to image cell nuclei stained with DAPI present in cell mounting solution.

#### *2.7.1.5. Organelle tracking*

Staining of CHO cells with live-cell organelle stains for use with microscopy and IF was optimised, based on the provided manufacturer protocol (MolecularProbes™ 2003; MolecularProbes™ 2005) and previous use of these dyes in CHO cells (Tigges and Fussenegger 2006), so as to maximise correct organelle staining whilst minimising off-target background staining, particularly that of the plasma membrane. For microscopy, staining was carried out in a 12-well plate with cells mounted on coverslips. For cell pellets, staining was carried out in 1.5 mL microfuge tubes, with solutions removed by centrifugation at 200 x g for 5 mins.

**ER tracker staining**. A 1.0 mM ER-Tracker™ Green glibenclamide BODIPY® FL (Molecular Probes™, Invitrogen) stock solution was produced in DMSO. Optimisation resulted in a protocol where a 1 µM working solution was produced in growth media and pre-warmed to 37 °C. Cells (2 x $10^6$) were washed in 1 mL 1x PBS. Working solution (300 µL) was added to the cells, mixed gently and incubated at 37 °C for 30 mins. Staining solution was removed and cells resuspended in 1 mL 1x PBS.

**Golgi tracker staining.** A 1.0 µM BODIPY® FL $C_5$-ceramide (Molecular Probes™, Invitrogen) working solution was produced in 1x PBS. Cells (2 x $10^6$) were washed in 1 mL 1 x PBS. Protocol optimisation resulted in the following staining protocol: Working solution (300 µL) was added to the cells, mixed gently and incubated at 4 °C for 30 mins. Staining solution was removed and cells washed twice with ice-cold 1x PBS before resuspension in 1 mL 1x PBS.

For flow cytometry, stained cells were stored on ice and analysed within 30 mins to minimise expulsion of the stain from the cell. For IF, cells were mounted on to a slide as section 2.7.1.3.

### 2.7.2. Flow Cytometry

Flow cytometry analysis was used to analyse GFP-expressing cells, organelle tracker stained cells or cells stained with fluorescent-tagged antibodies. Analysis

was performed using an Attune acoustic focusing cytometer (Applied Biosystems, Foster City, CA, USA) at the University of Sheffield Flow Cytometry Core Service. An ImageStream®X imaging flow cytometer (Amnis®, Merck, Kenilworth, NJ, USA) was used for analysis of organelle-tracker stained cells at MedImmune, Cambridge. Cells (up to 4 x $10^6$ in 1 mL of culture media or 1x PBS) were harvested/stained and stored on ice prior to analysis. For all analyses an unstained/non-transfected negative control sample was used with which to set negative control gates. The cell population was first gated to remove debris and cell doublets according to forward scatter (FSC) and side scatter (SSC) ratio. Fluorescence level and intensity were then gated according to the negative control, with a positive value determined as anything above 99% of the negative population. Green-fluorescing cells were analysed using a 488 nm laser with emission detection at 530 ± 30 nm (BL1). Red fluorescing cells were analysed using a 633 nm laser with emission detection at 660 ± 20 nm (RL1).

#### *2.7.2.1. Intracellular staining of Mab HC and LC*

CHO cells stably expressing recombinant Mabs were stained so the intracellular levels of Mab heavy and light chain (HC, LC respectively) could be determined across the population. Staining followed a MedImmune protocol. MeOH fixed cells (2 x $10^6$ cells; section 2.7.1.2) were washed with 5 mL 1x PBS then 5 mL blocking solution (1x PBS with 1% lyophilised BSA [Biowest Nuaillé, France])) on ice. Cells were stained with goat anti-Human IgG Fcγ-APC antibody (Jackson ImmunoResearch, West Grove, PA, USA) to stain heavy chain, and Lambda-FITC goat anti-Human or Kappa-FITC goat anti-Human antibodies (both Southern Biotechnology, Birmingham, AL, USA) to stain Mab light chain, depending on whether the cells expressed kappa or lambda LC. Cells were resuspended in 2 mL blocking solution containing 5 µg/mL anti-HC and 3 µg/mL anti-LC antibodies and incubated for 30-60 minutes. Staining solution was removed by centrifugation and decanting. Cells were washed twice with 5 mL blocking solution with stained cells resuspended in 1 mL blocking solution prior to analysis by flow cytometry. All washing and staining/incubation steps were carried out on ice or at 4 °C with cells pelleted at 130 x g for 5 mins. All staining solution preparation and cell staining was carried out in the dark to minimise chances of fluorescent tag bleaching.

### 2.7.3. SDS-PAGE and Western blotting

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) was used to separate proteins extracted from CHO cells, with western blotting used to highlight the presence of specific proteins within the CHO proteome.

#### *2.7.3.1. Whole cell protein sample extraction*

Whole Cell Protein (WCP) samples for separation by SDS-PAGE and analysis by western blotting were produced by lysing of cell cultures in Pierce™

Radioimmunoprecipitation (RIPA) cell lysis buffer (ThermoFisher) supplemented with a 1/250 dilution of Protease inhibitor cocktail set III, EDTA free (EMD Millipore Corp.-Calbiochem, Billerica, MA, USA). Cells were harvested at 300 RCF for 5 mins at a concentration of 2.5 x $10^7$ cells per 1 mL of RIPA-protease inhibitors to be used. Supernatant was removed by aspiration and the cell pellet washed in 1 mL 1x PBS at room temperature. After centrifugation and PBS removal by aspiration, cells were resuspended in the required volume of pre-chilled (4 °C) RIPA-protease inhibitor buffer and incubated on ice for 30 mins. Cells were further disrupted by homogenisation using a QIAshredder (Qiagen), the sample being passed through the shredder 1-3 times at 13,000 x g for 1 min. A refrigerated centrifuge was used at 4 °C to maintain a chilled cell lysate. After homogenisation, lysate was incubated on ice for a further 30 mins before centrifuging at 10,000 RCF for 10 mins at 4 °C. The supernatant, containing the whole cell lysate, was transferred to a Lo-Bind microfuge tube (Eppendorf, Stevenage, UK). Protein samples were stored at -20 °C (short term; up to 2 months) or -80 °C (long term). All comparative samples underwent the same freeze-thaw processes for reproducibility purposes.

#### *2.7.3.2. Sample separation by SDS-PAGE*

RIPA-prepared samples (from section 2.7.3.1) were diluted with plain RIPA buffer to a concentration of 10 x $10^6$ cells $mL^{-1}$. Samples were prepared for SDS-PAGE by mixing one volume RIPA sample with 2 volumes di$H_2O$ and one volume 4x NuPAGE LDS sample buffer (Invitrogen, ThermoFisher), resulting in a concentration of 2,500 cells $mL^{-1}$ of sample. Samples were incubated at 75 °C for 5 mins to break disulphide bonds. If not analysed immediately samples were stored at -20 °C. Chameleon Duo protein standard ladder (Li-Cor, Lincoln, NE, USA) was heated at 95 °C for 2 mins.

Protein separation gels (NuPAGE 4-12% Bis-Tris, 1.5 mm x 15 well; Invitrogen) were rinsed with di$H_2O$, the comb and plug tape removed and wells rinsed out with 1x NuPAGE MOPS SDS running buffer (from 20x stock by dilution with di$H_2O$; Invitrogen). Gels were placed in to a gel running tank and the upper buffer chamber filled with 1x running buffer to ensure there were no leaks, with the outer tank then part-filled with 1x running buffer. Samples (20 µL equating to 50,000 cells) or 4 µL Chameleon duo protein ladder were loaded per well. Gels were run at 200 V for 50 mins. During running the gel tank was cooled with ice packs.

#### *2.7.3.3. Protein transfer from SDS gel to membrane*

Transfer of proteins from gel to membrane was carried out using an iBlot® 2 dry blotting system (Invitrogen). The protein gel was removed from the gel cassette and rinsed with di$H_2O$ before sandwiching in an iBlot® gel transfer stack with nitrocellulose membrane (Invitrogen). Each layer of the gel-membrane and filter paper stack were gently rolled to remove bubbles before loading in to the iBlot®

system. Proteins were transferred using a 7 - 10 minute transfer programme for small and large (>150 kDa) proteins respectively, depending on which was to be blotted for.

#### *2.7.3.4. Western blotting*

Blocking and antibody staining were carried out in blocking buffer (1x PBS, 1% lyophilised BSA). All incubation and washing steps were carried out on a rocking plate at RT. Membrane edges were removed and the membrane incubated in 20 mL blocking buffer for 1.5 - 2 h. The membrane was rinsed with blocking buffer (1x PBS, 0.1% Tween 20 [BioChemica, AppliChem, Barcelona, Spain]) before 1 h incubation on a rocking plate at RT with 1° antibodies in 2 mL blocking buffer (for 1° antibody concentrations see table 2.3) in a heat-sealed pouch. After incubation the membrane underwent 3 x 5 min washes in 15 mL washing buffer. Staining with 2° antibodies was carried out in 10 mL blocking buffer for 1 h before 3 x 5 min blocking buffer washes. Membrane was dried on chromatography paper and stored in the dark before analysis (blots were stable for up to 6 months). Blots were scanned and imaged using an Odyssey SA fluorescent blot imager (Li-Cor), with image correction and quantitative analysis performed using ImageStudio software (Li-Cor). Antibodies used are shown in table 2-3.

## 2.8. PCR analysis of cell lines

The Polymerase Chain Reaction (PCR) was used to analyse mRNA transcripts within the CHO cell. Reverse transcription PCR (RT-PCR) was used to determine mRNA transcript levels of specific gene within the CHO cell. Standard PCR was also used to amplify transcripts of a specific gene, with the PCR product sequenced to check for mutations within the gene in CHO samples. To reduce the incidence of nuclease contamination all RNA prep and RT-PCR work was carried out in a sterile environment, the contents of which was cleaned stepwise with 70% ethanol, 10% bleach (sodium hypochlorate 10-15% active chlorine; Acros Organics, ThermoFisher) and RNaseZap® (ThermoFisher). All plastic ware used was certified nuclease free, with filter-tip pipette tips (SLS) also used to minimise contamination

### 2.8.1. RNA sample preparation

RNA samples from CHO cells were extracted using an RNeasy® minikit (Qiagen®) using the supplied protocol with a microcentrifuge. CHO cells (1 x $10^6$ were harvested and stored at -80 °C prior to RNA extraction. Samples were disrupted and homogenised by passing through a QIAshredder at 13,000 RCF. Ethanol was added to the lysate and the sample applied to the RNeasy mini spin column to bind DNA. RNA was washed before being eluted in 40 µL nuclease free water (Qiagen)

in to a nuclease-free microfuge tube and sterility maintained throughout use. An aliquot was taken to measure RNA concentration and purity by nanodrop.

### 2.8.2. Production of cDNA by reverse transcription

RNA was converted into complementary DNA (cDNA) by reverse transcription using the QuantiTect® Reverse Transcription kit (Qiagen). RNA (800 ng) from each sample was converted to cDNA, with a non-reverse transcriptase negative control also produced for each sample. RNA samples were thawed on ice and genomic DNA (gDNA) eliminated with a wipeout buffer at 42 °C for 2 mins. Reverse transcription was then performed, with reverse transcription buffer, primer mix and reverse transcriptase added to the gDNA-free RNA sample. The reverse transcription mix was incubated at 42 °C for 15 mins before three mins at 95 °C to inactivate the reverse transcriptase. The produced cDNA was stored at -20 °C prior to use.

### 2.8.3. PCR primer design

Primers were designed to amplify the Sec7 domain of the GBF1, ArfGef1 and ArfGef2 (BIG1/BIG2 synonyms) genes. These primers were used for both sequencing and RT-PCR. Primers were also designed to amplify sections of BiP, ATF6, CREB3L2 and XBP1 genes for RT-PCR analysis. Design of forward and reverse primer pairs was carried out using primer3 through the NCBI Primer-BLAST (basic local alignment search tool) software (Koressaar and Remm 2007; Untergasser et al. 2012; Ye et al. 2012). Primers were designed against gene sequences taken from the CHO-K1 genome 2014 update (Lewis et al. 2013; Xu et al. 2011). Primers were designed to be between 17 - 25 bases in length with an optimum of 20 bases. The percentage of GC base content was set at between 40 - 60% with the melting point ($T_m$) of each primer set between 52 - 62 °C, with each primer from a pair set to be with 5 °C of its partner primer. The 3' GC clamp was set such that a maximum of three G and C residues were present in the final 5 bases at the 3' end of the primer. Single and di-nucleotide repeats/runs were limited to a maximum of four in a row to limit incorrect binding. With the upstream (5') forward primer binding the mRNA sequence this primer was the same as the DNA sequence of the annealing site. The downstream (3') reverse primer was the reverse complement of the DNA sequence at the annealing site. Four primer pairs were designed for each gene with one primer pair chosen after primer efficiency tests. Primers were synthesised by Integrated DNA Technologies (IDT; Leuvan, Belgium). Primer stock solutions were prepared by resuspension in nuclease-free water, with stock solutions prepared at a concentration of 100 µM as described by the manufacturer. Primer pairs and sequences are shown in table 2-4.

The housekeeping genes FKBP1a and MMADHC were used to normalise samples for cDNA levels. These are historically used within the laboratory group as reference genes for qPCR- and RT-PCR-based techniques having been selected from MedI-CHO RNA-seq transcriptomic data as having highly stable expression levels across many MedI-CHO growth conditions (Brown et al. 2018).

### 2.8.4. RT-PCR of CHO transcripts

For each RT-PCR reaction cDNA was amplified by forward and reverse primers to determine transcript levels of specific genes. For each gene to be analysed each RT-PCR reaction was repeated in triplicate. Reactions containing no-reverse transcriptase cDNA and no template (cDNA replaced with water) were used as negative controls for each primer set. Primer pairs were diluted in nuclease-free water such that each primer was present in the RT-PCR reaction at 2 μM concentration.

RT-PCR was carried out using QuantiFast® SYBR® Green PCR Kit (Qiagen). For each RT-PCR reaction 2.5 μL of primer pair was mixed with 12.5 μL SYBR Green and 1.1 μL cDNA sample (or negative control substitutions). The final reaction volume was made up to 25 μL with 8 μL nuclease-free water. All RT-PCR reactions were prepared in sterile nuclease-free microfuge tubes. SYBR safe was added to the reaction mixture last and was not exposed to the light. For technical replicates the mix was prepared together before being split across three wells. All RT-PCR reactions were prepared with a minimum 10% excess per well. After preparation reaction mixture was transferred to a MicroAmp® Fast optical 96-well reaction plate (Applied Biosystems), with each technical replicate put in a separate well. The plate was sealed with a MicroAmp™ optical adhesive film (Applied Biosystems) and transferred to an Applied Biosystems 7500 Fast Real-Time PCR system for execution of the RT-PCR experiment. RT-PCR was prepared and controlled using 7500 fast system sequence detection software v 1.4.0 (Applied Biosystems), with the same software used for results analysis. Thermal cycles for RT-PCR experiments are shown in table 2-5.

**Table 2-5. Primer sequences for PCR/RT-PCR.**

| Primer | Sequence (5'-3') | Length | Start base | Stop base | $T_m$ (°C) | GC % | Product length |
|---|---|---|---|---|---|---|---|
| **GBF1** | | | | | | | |
| Fwd | AGGAAGCCGCCTCGATTTTC | 20 | 2026 | 2045 | 60.7 | 55 | 739 |
| Rev | GATCATAGCTGCCAGGAGGT | 20 | 2764 | 2745 | 58.9 | 55 | |
| **ArfGef1** | | | | | | | |
| Fwd | TGGCAGCTACAGTACACAGATG | 22 | 2241 | 2262 | 60.1 | 50 | 770 |
| Rev | GAACATGGCTCACTGCTTCC | 20 | 3010 | 2991 | 59.2 | 48 | |
| **ArfGef2** | | | | | | | |
| Fwd | GGGACCCAGACAACCATTCA | 20 | 1900 | 1919 | 59.6 | 55 | 735 |
| Rev | AGCTTTGGCCGTTTTAGCCA | 20 | 2634 | 2615 | 60.8 | 50 | |
| **ATF6** | | | | | | | |
| Fwd | TGGCAAAGCAGCAATCATCG | 20 | 608 | 627 | 59.8 | 50 | 185 |
| Rev | TGGCACCACATTGACCACAT | 20 | 792 | 773 | 60.2 | 50 | |
| **BiP (HSPA5a)** | | | | | | | |
| Fwd | GCCACGAATGGAGACACTCA | 20 | 885 | 904 | 60.0 | 55 | 148 |
| Rev | CCTTTTCCACCTCACGACGA | 20 | 1032 | 1013 | 60.0 | 55 | |
| **CREB3L2** | | | | | | | |
| Fwd | GGGGTCTTTGGTCTCCAACAA | 21 | 386 | 406 | 60.1 | 52 | 199 |
| Rev | GAGACAGGCTGAAGGGATGC | 20 | 584 | 565 | 60.5 | 60 | |
| **XBP1s** | | | | | | | |
| Fwd | TCTGTCAGTGGGGACCTCAT | 20 | 694 | 713 | 59.9 | 55 | 154 |
| Rev | GCTGAGAGGTGCTTCCTCAA | 20 | 847 | 828 | 60.0 | 55 | |
| **FKBP1a** (housekeeping control – designed by Adam Brown) | | | | | | | |
| Fwd | CTCTCGGGACAGAAACAAGC | 20 | 36 | 55 | 58.6 | 55 | 95 |
| Rev | GACCTACACTCATCTGGGCTAC | 22 | 130 | 109 | 59.6 | 55 | |
| **MMADHC** (housekeeping control – designed by Adam Brown) | | | | | | | |
| Fwd | TGTCACCTCAATGGGACTGC | 20 | 344 | 363 | 61.7 | 55 | 145 |
| Rev | CAGGTGCATCACTACTCTGAAAC | 23 | 488 | 466 | 60.5 | 48 | |

*Table 2-5 shows the sequences of the primers designed for RT-PCR analysis of specific genes in evolved and non-evolved CHO cells.*

RT-PCR data was produced as a cycle time (Ct) value, with higher numbers indicating that more RT-PCR cycles were required to reach a certain fluorescence threshold, indicating lower initial transcript levels. Ct values are measured on a log scale and as such are not linear (a change of 1 relating to a 10-fold change in expression level). Analysis threshold was reduced to 0.02 amplify low levels expected with endogenous genes. Each Ct value was converted to an expression fold change value by comparing the raw value of the test subjects and primers against the control samples and housekeeping gene primers using the double-delta Ct analysis technique according to Equation 8 (Kannan 2018; Livak and Schmittgen 2001; Schmittgen and Livak 2008). These values were then averaged across biological and technical replicates.

**Equation 8 -** Double-delta analysis for RT-PCR analysis. Calculating the change of gene transcript level between the test and control, whilst normalising against housekeeping controls

$$\Delta CTE = TE - HE$$
$$\Delta CTC = TC - HC$$

$$expression\ fold\ change = 2^{\Delta CTE - \Delta CTC}$$

Where TE = test cells, test gene; TC = control cells, test gene; HE = test cells, housekeeping gene; HC = control cells, housekeeping genes.

#### *2.8.4.1. RT-PCR primer efficiency tests*

Primer efficiencies were tested to ensure primer pairs functioned as designed. Each primer pair was tested at six cDNA concentrations. cDNA concentrations were prepared as 1 in 3 serial dilutions, from a 3-fold dilution down to a 729-fold dilution. Each concentration with each primer pair was repeated in triplicate. For each primer pair two negative controls were prepared: one each of a no-reverse transcriptase reaction and a no template DNA reaction. The primer efficiency test was carried out using the RT-PCR protocol described in section 2.8.4, using the disassociation step shown in table 2-5.

Primer efficiency was calculated by plotting the RT-PCR Ct scores against the $\log_{10}$ value of the dilution factor. A trendline was added to this plot and the slope value of the trendline used to calculate the primer efficiency using ThermoFisher's qPCR efficiency calculator (ThermoFisher 2018). For primer pairs to be suitable for RT-PCR use their efficiency had to be between 95-105%.

**Table 2-6. Temperature cycling for RT-PCR.**

| Stage | Temperature (°C) | Duration (min) | repeats |
|---|---|---|---|
| **1** | 50.0 | 02:00 | 1 |
| **2** | 95.0 | 15:00 | 1 |
| **3** | 94.0 | 00:15 | 40 |
| | 60.0 | 01:00 | |
| **Disassociation step: for primer efficiency test only.** | | | |
| **4** | 95 | 0:15 | 1 |
| | 60 | 1:00 | |
| | 95 | 0:15 | |

**Table 2-7. Standard PCR component make-up.**

| Component | Volume |
|---|---|
| Q5® high-fidelity 2x master mix | 12.5 µL |
| 10 µM forward primer | 1.25 µL |
| 10 µM reverse primer | 1.25 µL |
| Template cDNA | 0.5 µL (approx. 1000 ng) |
| Nuclease-free water | 9.5 µL (to 25 µL final volume) |

**Table 2-8. Standard PCR theromcycling protocol.**

| Stage | Temperature (°C) | Duration (min) | repeats |
|---|---|---|---|
| 1: Denaturation | 98 | 00:30 | 1 |
| 2: Amplification | 98 | 00:10 | 35 |
| | 67 | 00:30 | |
| | 72 | 00:30 | |
| 3: Final extension | 72 | 02:00 | 1 |
| 4: Hold | 4 | ∞ | 1 |

*Tables 2-6 and 2-8 show the PCR cycling settings used to amplify DNA sequences for RT-PCR (2-6) and standard PCR (2-8) experiments.*
*Table 2-7 shows the make-up of the standard PCR experiment protocol.*

### 2.8.5. PCR and sequencing

Forward and reverse primers were designed to sequence the Sec7 domain of endogenous CHO GBF1. Primers were designed such that at least 50 bases were present between the primer site and the start/end of the target Sec7 sequence (within a maximum of 180 base pairs of the Sec7 sequence outer limits). The Sec7 protein sequence was found using NCBI BLAST with BLAST tools used to locate it within the whole GBF1 protein sequence (Altschul et al. 1990). From this the region of the GBF1 gene encoding the Sec7 domain was elucidated, with sequence translation software (EMBOSS transeq) used to confirm the correct region of DNA had been selected (table 2-8; Li et al. 2015; McWilliam et al. 2013; Rice et al. 2000).

RNA samples were prepared from CHO samples as in section 2.8.1. These samples were then converted to cDNA as per section 2.8.2. The required gene section to be sequenced was amplified by PCR using Q5® high-fidelity DNA polymerase (NEB). cDNA concentration was measured by NanoDrop. The PCR reaction mixture was prepared in 200 µL domed cap PCR tubes (StarLab, Milton Keynes, UK) as shown in table 2-6. Tubes were loaded in to a Veriti 96 well thermal cycler (Applied Biosystems) and the PCR amplification reaction carried out as in Table 2-7.

After PCR 5 µL 6x purple loading buffer (NEB) was added to the 25 µL PCR reaction and mixed. Samples (25 µL) were run on a 1% agarose gel (as section 2.4.2.6) at 100 V for 20 mins. PCR amplified DNA was visualised using a UV visualiser and excised from the gel using an ethanol-sterilised blade. PCR product was purified using QIAquick gel extraction kit (Qiagen) – agarose gel was dissolved at 50 °C before PCR product was bound to a QIAquick DNA binding tube and washed before elution in nuclease free water. Two-way sequencing was performed by the University of Sheffield sequencing service using a 3730 DNA analyser (Applied Biosystems).

**Table 2-9. Sequence details of Sec7 domain in GBF1 gene/protein.**

| **Protein sequence** | |
|---|---|
| Protein NCBI reference sequence | XP_007646849.1 |
| Protein Length (amino acids; aa) | 1856 |
| Sec7 protein sequence length (aa) | 186 |
| Sec7 start (aa) | 697 |
| Sec7 end (aa) | 882 |
| **Gene sequence** | |
| Sec7 start base | 2089 |
| Sec7 end base | 2646 |
| Sec7 DNA length | 557 |
| NCBI Gene ID | 100689421 |

*Table 2-9 shows the details of the Sec7 domain that is the functional region of the Arf-GEF family of proteins that are involved in initiating vesicle budding in the Golgi body.*

## 2.9. Statistical analysis techniques

Data was curated, manipulated and analysed using Excel®:Mac 2011 version 14.5.3 (Microsoft®, Redmond, WA, USA). Graphs were plotted and statistical analysis performed using Prism 7 for Mac OS X version 7.0c (GraphPad, Ca, USA).

### 2.9.1. Basic statistical analysis techniques

#### *2.9.1.1 Average (median/mean)*

Average values for data were for the most part shown as the mean average of an entire data set – the total of all results divided by the mean of the results. For flow cytometry data average values were shown as the median value – that being the middle value once the data set has been arranged in order of numerical value – so as to remove any outliers from the average calculation. As flow cytometry data should fit a normal distribution little difference between mean and median values was expected.

#### *2.9.1.2. Standard deviation and standard error of the mean*

Standard deviation (SD) quantifies the spread of data points about the mean, whilst SEM quantifies the precision of the mean, displaying how far the sample mean is likely to be from that of a true population mean. The standard error of the mean SEM is smaller than the standard deviation. All error bars plotted through the course of the project show SEM and not standard deviation. All average and data spread values were calculated from all biological and technical replicate data points to provide more power and accuracy to the data and conclusions drawn from it. As such, the use of SEM instead of standard deviation is viable so as to better compare the means of biological replicates. SEM and standard deviation values were calculated by Prism software. SD was calculated as in equation 9 and SEM as in equation 10.

**Equation 9 -** Standard Deviation.

$$SD = \sqrt{\frac{\Sigma(x - X)^2}{n - 1}}$$

Where Σ = sum of, $x$ = each individual value, $X$ = population mean and $n$ = number of population.

**Equation 10 -** Standard error of the mean.

$$SEM = \frac{SD}{\sqrt{n}}$$

#### *2.9.1.3. Student's T-test with Welch's correction*

To compare data to the control in a small experiment (less than a total of six comparisons required) an unpaired parametric (assuming Gaussian distribution) Student's T-test was used, using Welch's correlation as equal SDs were not assumed. T-tests were performed by Prism to determine whether results were statistically significant. Throughout the project statistical significance was shown as P-value scores, shown on graphs as asterisks as follows:

- n/s – $P > 0.05$ (data is not statistically significant with 95% confidence – 19 out of 20 times).
- * - $P \leq 0.05$ (data is statistically significant with 95% confidence).
- ** - $P \leq 0.01$ (data is statistically significant with 99% confidence).
- *** - $P \leq 0.001$ (data is statistically significant with 99.9% confidence).
- **** - $P \leq 0.0001$ (data is statistically significant with 99.999% confidence)

### 2.9.2. Dunnett's multiple comparison test

For experiments with large numbers of parameters to compare a T-test could not be used as with this there is a 5% (1 in 20) chance of error and as such is not suitable in larger data sets, especially those with 20 or more comparisons to be made. Instead a Dunnett's multiple comparison test was used to compare many test values to the single control and to determine whether any showed statistical significance. Prism was used to perform Dunnett's tests with significance shown as described in section 2.9.3.

### 2.9.3. Design of Experiment experimental design and analysis

The Design of Experiment (DoE) approach was used to plan and perform experiments to determine cause and effect relationships with respect to multigene engineering of CHO cells. Two-level factorials were used with the effect of gene dose levels compared at upper and lower levels, where upper levels were data driven and lower levels set to zero. Design-Expert® version 10 software (Stat-Ease, Minneapolis, MN, USA) was used to plan factorial experiments and analyse the data collected so as to determine factors impacting upon experimental results (Anderson and Whitcomb 2007). Statistical significance was calculated by this software using analysis of variance (ANOVA).

# 3) Secretory pathway mapping and identification of CHO engineering targets

Several datasets were used during this chapter. The in-lab CHO transcriptomic data set was produced by Dr. Darren Geoghegan (ex-Prof. David James [DCJ] group). The CHO proteomic data set was produced by Dr. Joseph Longworth (ex-DCJ group). B cell/Plasma cell data set produced by Nabila Rahmen (Dr. Andrew Peden group). The initial gene target data base was designed in collaboration with Stefan Krämer (ex-DCJ group).

*A systematic process was required to select gene targets with which to engineer the CHO cell secretory pathway with the aim to increase productivity. To build upon the knowledge of pathways and protein families functional in the mammalian cell secretory pathway gained through the literature review (section 1.5), more specific knowledge of the activity and prevalence of these genes and proteins was required. Analysis of transcriptomic and proteomic data (referred to combined as 'omic) from CHO cells and from high-secreting plasma cells was used to build upon this literature-driven understanding. In this chapter:*

- *Data from relevant CHO cell lines produced within the DCJ laboratory group was used to highlight genes important in CHO productivity and growth.*
- *Plasma-cell data sets produced by the Peden laboratory highlighted genes important in conferring enhanced productivity and secretory specificity upon plasma cells.*
- *Analysis of 'omic data from previous CHO literature allowed comparisons between different CHO lineages to be made as well as some consensus as to genes functional in providing an enhanced CHO productivity phenotype.*

*These analyses allowed the formation of a gene targets database which was subsequently used to inform selection of genetic engineering targets with which to enhance the CHO biosynthetic and secretory pathways.*

## 3.1. Mapping of the CHO secretory pathway

An in-depth, well annotated map of the CHO secretory pathway is not available within the literature due in part to the current incomplete nature of the CHO genome (Lewis et al. 2013; Xu et al. 2011). However, resources mapping the general ER protein folding and secretory pathways within other mammalian cells, including from CHO-relevant cell lines of a rodent lineage, are available from the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa and Goto 2000; Kanehisa et al. 2014). CHO 'omic data has been combined with prior mapping and literature-based knowledge of the genetic and protein interactions that make up the CHO cell biosynthetic and secretory pathways.

This combination of background secretory knowledge and specific 'omic data has provided a working model of CHO biosynthetic and secretory systems. Combined with knowledge of secretory bottlenecks and previous CHO engineering strategies, this informed decisions as to specific regions of the secretory pathway to target with engineering in an attempt to increase cellular machinery, alleviating biosynthetic and secretory bottlenecks to increase flux though the secretory pathway, aiming to ultimately increase CHO cell productivity (Becker et al. 2008; Hansen et al. 2017; Peng et al. 2011; Peng and Fussenegger 2009b). The rational behind this approach has been shown with the use of 'omic data already being shown to improve the performance of mammalian cell lines in biologic production by allowing a rational engineering approach (Lewis et al. 2016; Vishwanathan et al. 2015).

### 3.1.1 CHO transcriptomic and proteomic data sets

Transcriptomic and proteomic data sets produced from MedImmune CHO cell lines by the DCJ laboratory group were used to enhance secretory pathway mapping within the CHO cell. CHO 'omic data was used to better inform selection of targets for genetic engineering of the CHO biosynthetic and secretory pathways. The 'omic data provided information on genotypic and proteotypic changes occurring within CHO cells at different times within cell culture and in cells with different Mab productivity level. Using transcriptomic and proteomic datasets to compare Mab-producing and non-producing cell lines highlighted pathway alterations and changes in genetic expression levels that occur upon turning a parental CHO cell into a mammalian cell factory. This allowed links to be made between genotype, proteotype and phenotype. That many genes and proteins shown to be enhanced in producing CHO cells have previously been linked to enhanced productivity in both CHO and other cell lines (e.g. UPR transcription factors and response elements) further highlighted the close link between various different CHO cell lines and other non-CHO production cell lines. This provided confidence that other mammalian cell lines could be used as a model upon which to base CHO cell genetic engineering.

#### *3.1.1.1. CHO transcriptomic data set*

Transcriptomic profiling is used to study the transcriptome of an organism, providing information on the sum of RNA transcripts and therefore the control of gene expression levels (Mantione et al. 2014). The abundance of RNA transcripts only shows the level at which genes are transcribed and whilst this does not necessarily correlate with cellular protein levels it does intimate at protein levels whilst providing an insight into regulatory and control-based changes within the cell at a specific point in time. Transcriptomic profiling includes non-coding and small RNAs in the dataset, whilst gene splicing, active/inactive forms of proteins and PTM status are not accounted for. Therefore the exact functional stoichiometry is not known.

If knowledge of the transcription levels of a small number of genes is required, quantitative PCR (qPCR) can be used to compare against cellular and housekeeping controls. For large-scale global transcriptomic analysis of a cell line/organism the two prevalent techniques are microarrays, which quantify a set of pre-determined sequences (e.g. from a certain organism or cell pathway), and RNA sequencing (RNA-seq), in which all transcripts are recorded by high-throughput sequencing. Whilst more costly than microarray analysis, RNA-seq provides a more high-throughput, wider ranging, precise and accurate quantification of a larger number of transcripts, including information on isoforms and splicing (Mantione et al. 2014; Wang et al. 2009b; Zhao et al. 2014).

In-lab transcriptomic data sets used throughout this project showed gene transcript levels as Fragments per Kilobase of transcript per Million of mapped reads (FKPM). This shows the relative expression level of a specific transcript (i.e. the higher the FKPM, the more transcripts that are present). This transcript abundance data was converted to a fold-change of transcript abundance in the test subject (e.g. higher producing cells) compared to transcript abundance in the control (e.g. parental cells), allowing the relative change in expression levels to be determined (i.e. the larger the fold change, the bigger the change in abundance level). Fold change data was the most commonly used to mine differences between producing and non-producing cell lines. However, attention still had to be paid to the absolute FKPM value, as a high fold-change value could be produced from a combination of a very low control FKPM value paired to a modest test FKPM, resulting in an amplified fold change value over-estimating the effect this gene's transcription level change would have upon the cell phenotype.

Raw transcriptomic data of MedI-CHO cell lines, produced by RNA-seq, was provided by Darren Geoghegan of the DCJ group (Geoghegan et al. 2018). Data compared host MedI-CHO with a MedI-CHO cell line stably expressing an easy to express Mab, with both cell lines grown in bioreactor culture. RNA profiles of each cell line were taken, in duplicate, at days four and eight of culture to provide

information on transcript profile during growth and stationary growth phases. The data was used here to calculate abundance fold-changes. The difference in transcript abundance between producing and non-producing cells was calculated by dividing the value of the test subject (e.g. abundance in producing cells at day 8 of culture) by the comparative control value (abundance in host cells at day 8 of culture). Comparisons were also made between producing cells at days eight and four of culture to compare changes that occur within a producing CHO cell as it changes from prioritising cell growth to producing higher levels of recombinant protein during stationary phase.

#### *3.1.1.2. CHO proteomic data set*

Proteomic profiling provides information as to cell population's proteotype - that being the abundance of all proteins present within the cell - at a certain time point. Whilst transcriptomic data provides information as to regulatory control within the cell at the sampled time point, proteomic data provides a quantitative insight to exact protein levels as well as providing information on present PTMs. This provides insight to the molecular levels, interactions and turnover rates that in turn directly inform and impact upon a cell's function and phenotype at a certain stage in the cell cycle or growth curve (Larance and Lomond 2015).

Raw proteomic data of a clonally derived CHO-S cell line stably producing an ETE Mab (Herceptin) was produced using mass spectrometry of labeled cells (pulsed stable isotope labeling of amino acids in cell culture; pSILAC). Samples were taken at days three (growth phase) and eight (stationary phase) of bioreactor culture, with biological duplicates taken of each sample (Schwanhaeusser et al. 2011). Proteomic data was produced by Dr. Joseph Longworth of the DCJ laboratory and was used here to calculate abundance fold change between the two different stages in cell growth.

## 3.2. Plasma cell transcriptomic data

An RNA-seq data set comparing transcript levels between naïve B cells and terminally differentiated plasma cells was provided by Nabila F Rahman of the Peden laboratory group, who also carried out fold-change calculations and analysis. Data was presented as a $\log_2$ fold-change in abundance levels between the two stages of cell development (as in Shi et al. 2015). The dataset provided insight to the regulatory changes that result in the high-producing and secreting plasma cell phenotype. All plasma cell data sets were taken from publically available data. Transcriptomic analysis was performed by Nabila Rahman. Here data was sorted and mined to highlight genes linked to plasma cell differentiation and possibly the enhancement of a productivity phenotype that could potentially function in CHO cells to enhance their productivity.

### 3.2.1. Analysis of plasma cell transcription factors

Due to limits regarding plasmid size and DNA loads possible in multigene engineering of the CHO cell, difficulty in upregulating lots of separate genes individually and the likely lack of effect the engineering of a few specific molecules will have upon the cell due to the imposition of stoichiometric imbalances, it is likely that transcription factors will play an important role in multigene engineering of the CHO cell biosynthetic and secretory pathways. As such expression of a single or small selection of genes activating signaling cascades that instigate myriad changes within the cell should allow fine and balanced engineering of the CHO biosynthetic and secretory pathways. Plasma cell transcriptomic data was mined to investigate transcription factors with increased transcriptional abundance in plasma cells which could therefore be considered to have a functional role in increasing cellular productivity levels.

The plasma cell transcriptomic data set was narrowed to show only genes known to function as transcription factors. The 50 genes with the highest fold-change in abundance between B and plasma cells were selected, before further narrowing down to remove those artificially amplified (those with very low starting values) and linked to pathways not involved in cell growth, proliferation, cargo protein biosynthesis, folding, PTM, secretion or ER stress response. These genes are listed in table 3-1 with pathway and function notes taken from their gene ontology (GO) terms.

As these transcription factors play a role in the differentiation of B cells into Mab-producing factories they could, depending upon the pathways they regulate, prove to be worthwhile targets for enhancing the secretory and productivity capacity of CHO cells (Dinnis and James 2005). However, due to the terminally differentiated nature of plasma cells some plasma cell specific transcription factors could direct cells towards terminal differentiation and halt cell growth. These would therefore be unsuitable for CHO cell engineering. Some genes that are not classical transcription factors (e.g. elongation factors) were included within the dataset due to their global role in transcription and translation enhancement.

As high-productivity plasma cells mature from naïve B cells they undergo an increase in ER and Golgi volumes alongside an increase in a wider array of secretory and biosynthetic pathway components (Shaffer et al. 2004; Shapiro-Shelef and Calame 2005). It is therefore not unexpected that transcription levels of many transcription factors involved in the UPR see an increase in plasma cells. XBP1, IRE1, ATFs 4, 5 and 6 and DDIT3/CHOP are all UPR-linked transcription factors that increase in abundance in plasma cells, further enhancing their reputation as targets for engineering of protein folding and secretion capacity in CHO cells (Nishimiya et al. 2013; Pybus et al. 2014; Tigges and Fussenegger 2006). TRIP11 may play a role in Golgi stress response as well as Golgi structure

maintenance, whilst CREB3 family members are known to regulate secretory capacity in secretory-specific cells, whilst also activating secretory genes when expressed in non-secretory specific HEK and *Drosophila* cells (Barbosa et al. 2013; Fox et al. 2010).

PSIP1 and HMGN3 both function in stem cell differentiation, but their known function in pancreatic cells alongside upregulation in plasma cells shows potential stress-response and general secretion roles respectively. Some of the listed transcription factors (PRDM1, ARID3A) may be plasma-cell specific with functions leading to terminal differentiation and as such they are unlikely to be feasible targets within CHO manufacturing processes. With roles in promoting homologous IgG gene expression in plasma cells it is feasible that IRF4, TCF4 and POU2AF1 could increase transcription levels of recombinant IgG genes, although this would likely require the transgene to be under control of a promoter containing plasma cell specific TFREs. Whilst this is not feasible within the scope of this project it could be an area of development in synthetic promoter/ transcription factor development (Brown and James 2016). Likewise, the elongation complex members ELL2 and MLLT3 enhance transcription/ translation rates so are interesting targets but not specific to the scope of this project.

*Table 3-1 (next page) shows the most highly upregulated transcription factors in plasma cells. From the initial RNA-seq data set genes were narrowed down as follows: Genes listed as transcription factors were retained; genes that were considered artificially amplified were removed; to maintain focus on cell growth and productivity, genes involved in cell growth, proliferation, cargo protein biosynthesis, protein folding, PTMs, secretion or ER stress response were retained. The transcriptional fold-change of each gene in plasma cells is shown, alongside brief notes of each gene's potential function taken from gene ontology.*

**Table 3-1. Upregulated transcription factors in plasma cells**

| Gene | Uniprot ID (mouse) | Abundance Fold Change | Function/pathway notes (from GO terms) |
|---|---|---|---|
| **HID1** | Q8R1F6 | 9.37 | Intracellular protein transport, Brefeldin A response. |
| **PRDM1** | Q60636 | 7.49 | Transcriptional repressor, drives B-lymphocytes maturation to IgG secreting cells. |
| **FOS** | P01101 | 7.00 | Cell proliferation and differentiation. Activates Phospholipid synth in growing cells. |
| **CREB3L2** | Q8BH52 | 6.13 | Promotes transport and secretion and ER biogenesis. |
| **XBP1** | O35426 | 5.20 | Binds ERSE, leading to increase in ER chaperones, CHOP. |
| **ERN1 (IRE1)** | Q9EQY0 | 5.00 | Serine/Threonine kinase. Senses unfolded proteins leading to enzyme activation and splicing of XBP1 DNA to active form. |
| **ATF5** | O70191 | 4.53 | Cell survival, proliferation & differentiation. |
| **ELL2** | Q3UKU1 | 4.31 | Elongation Factor. Increases catalytic rate of RNApolII transcription (suppresses transient pausing). Role in IgG production, influences mRNA processing: Poly(A) site choice, exon skipping & Ig HC processing |
| **IRF4** | Q64287 | 3.03 | Binds MHC class I promoter. IgG LC transcription enhancer (alongside PU.1). |
| **HMGN3** | Q9DCB1 | 2.90 | Affects insulin/glucagon levels and modulates expression of pancreatic genes involved in insulin secretion. |
| **ARID3A** | Q62431 | 2.65 | Involved in B cell differentiation. Binds promoter for induced HC transcription |
| **ATF6** | F6VAN0 | 2.28 | Initiates UPR under ER stress |
| **MLLT3** | A2AM29 | 2.05 | Super-elongation complex member – increases rate of RNApolII transcription |
| **TCF4** | Q60722 | 1.77 | Binds Ig enhancer mu-E5/KE5 motif |
| **PSIP1** | Q99JF8 | 1.62 | Involved in stem cell differentiation/ neurogenesis. Possible protective role during stress-induced apoptosis. |
| **ATF4** | Q06507 | 1.45 | ER stress response: DDIT3/CHOP |
| **POU2AF1** | Q64693 | 1.40 | Activates Histone H2B and Ig genes. |
| **DDIT3 (CHOP)** | P35639 | 1.25 | ER stress response |
| **CREB3** | O43889 | 1.23 | ER-based UPR |
| **TRIP11** | Q15643 | 1.20 | Potential role in maintaining cis-Golgi structure. |

### 3.2.2. Comparison of plasma cell and producing CHO cell transcriptomics

The CHO cell RNA-seq data set was compared to that of the plasma cell. Genes within the CHO cell were ranked based on the FKPM fold change seen between the parental and producing cell lines on day eight (stationary phase) of bioreactor culture to elucidate genes that undergo the greatest increase in activity during recombinant protein production in CHO cells. After removal of artificially amplified genes (i.e. those with a top FKPM value ≤1, an arbitrary selection), the top 50 ranked CHO genes were searched for within the plasma cell data set to compare the process by which CHO cells and plasma cells adapt to Mab production. Of the 50 CHO productivity-linked genes only six genes were present within the plasma cell transcriptomic data set. These genes, along with abundance fold-change data from both producing CHO and plasma cells, and functional notes from GO terms, are shown in table 3-2.

That only six of the top-50 highest overexpressed genes in producing CHO cells (when compared to parental cells) are present in plasma cells may be due to poor annotation of the CHO genome but may also be due to the specificity of human plasma cells and the difference between them and CHO cells with regard to Mab productivity levels. Of the six genes present in both lists only one, ATP8a2/ATP8b2, involved in phospholipid transport, appears to be linked to the secretory or biosynthetic pathways, with a potential role in organelle biogenesis and expansion. Both these points highlight that CHO cells are not specified for enhanced productivity, even when expressing a recombinant Mab.

As such, looking in to CHO transcriptomic data may not necessarily aid in deciphering potential engineering targets with which to enhance CHO productivity, with secretory specific cell lines likely to provide more novel targets. With many of the gene hits being involved in molecular and mechanical interactions with other CHO-resident proteins in the process of biosynthesis and secretion, these targets may require CHO homologues to function within CHO cells, whereas CHO-derived targets will be known to function correctly. As such, use of non-CHO derived targets is a potential high-risk strategy but with potential high rewards with the knowledge that they are known to function in high-producing cell lines. The plasma cell transcript abundance levels of genes with a direct role in protein folding and secretion (e.g. SNAREs, protein folding chaperones, vesicle proteins) were also used to better inform gene selection. However, due to the likelihood of more precise stoichiometric balance between these components, they were not analysed as closely as the more global transcription factors, but their abundance data was used to back-up literature and CHO'omic data during target selection.

**Table 3-2. Comparison of CHO and plasma cell transcriptomic data**

| Gene | CHO fold change (FKPM values) | Plasma fold change (FKPM values) | Functional notes (from GO terms) |
|---|---|---|---|
| **Glul (glutamine synthetase)** | +83.46 | -3.80 | Glutamine synthesis (presence due to inclusion on expression plasmid.) |
| **Nedd9** | +8.50 | -3.50 | Docking protein, coordinating role of tyrosine kinase signaling involved in cell adhesion and growth |
| **Slc14a** | +5.68 | -3.50 | Urea transporter |
| **ATP8a/b2** | +4.33 | +4.35 | Phospholipid-transporting ATPase. Catalytic component of P4-ATP flipase. |
| **SP100** | +4.00 | -4.50 | Tumour suppressor |
| **Fos** | +3.65 | +7.00 | Fos/Jun pathway role in cell proliferation, differentiation. Soma apoptotic connotations |

*In Table 3-2 CHO and Plasma cell RNA-seq data sets are briefly compared. Genes within the CHO cell were ranked based on the FKPM fold change seen between the parental and producing cell lines to elucidate genes that undergo the greatest increase in activity during CHO recombinant protein production. After removal of artificially amplified genes the top 50 ranked CHO genes were searched for within the plasma cell data set . This allowed comparison of the process by which CHO cells and plasma cells adapt to Mab production. Of the 50 CHO productivity-linked genes only six genes were present within the plasma cell transcriptomic data set.*

## 3.3. Mapping of CHO 'omic data on to biosynthetic pathway maps

Transcriptomic and proteomic datasets comparing producing and non-producing CHO cells, and CHO cells in exponential and stationary growth phases, were used to assay the 'omic status of the CHO cell biosynthetic and secretory pathways in comparison with widely available molecular maps of mammalian cell protein processing in the ER. Pathways relevant to ER processing and secretion were taken from KEGG and were specified for Chinese hamster (Kanehisa and Goto 2000; Kanehisa et al. 2014). CHO and plasma cell 'omic data used to augment these gene pathway/interaction maps.

Mapping focus was directed towards pathways more specifically involved in protein transport into the ER, protein folding/misfolding, the ER stress response/UPR and COPII vesicle formation (figure 3-1), as well as SNARE interactions at the ER, the *cis*-, *medial* and trans-Golgi and the TGN (figure 3-2). CHO and plasma cell 'omic data for the genes listed within the maps were taken from the requisite data sets and plotted to show the fold-change in transcript/protein level of these genes between producing and non-producing CHO cells (transcriptomic data); the stationary and growth phases of producing CHO cells (transcriptomic and proteomic data); and naïve B and plasma cells (transcriptomic data). As well as providing information on the activity and abundance of specific genes and proteins, these comparisons also provided:

- Comparison of biosynthetic- and secretory-specific transcriptomic differences between producing and non-producing CHO cells.
- Comparison of 'omic differences between producing cells in growth and stationary phases.
- Comparison of transcriptomic differences in the protein processing machinery and capacity of CHO and plasma cells.
- Ability to better define the functional levels, and therefore better asses the importance, of different pathways linked to cellular productivity.

The KEGG 'protein processing in the ER' and 'SNARE interactions in vesicular transport' molecular pathway maps are shown in figure 3-1 and figure 3-2, with relevant sections of the map highlighted. For some sections (e.g. translocon) the entire pathway is not shown on the map but components of this sub-pathways were taken from the requisite KEGG-linked pathway and are included in the overall mapping exercise for completeness. Data from the transcriptomic and proteomic data sets was used to calculate the fold-change in gene transcript number or protein abundance as follows:

1. CHO transcriptomic data – comparison of expressing and non-expressing CHO cells: Comparing fold-change of FKPM abundance levels in stably-producing MedI-CHO-ETE to that of non-producing MedI-CHO. Both during stationary growth phase.
2. CHO transcriptomic data – comparison of cells in the exponential growth (day 4) and static growth/production (day 8) stages: Comparing fold-change in FKPM in stably-producing MedI-CHO-ETE in stationary producing phase to that of the same cell line in exponential growth phase.
3. CHO proteomic data – comparison of cells in the exponential growth (day 4) and static growth/production (day 8) stages: Comparing protein abundance fold-change seen in stably-producing CHO-S in stationary phase to that of the same cell line in exponential growth phase.
4. Plasma cell comparison: Comparing FKPM transcript abundance values in terminally differentiated, production specific plasma cells, to those in non-differentiated low-producing naïve B cells.

Fold-change data is shown in figure 3-3 and figure 3-4. The KEGG pathway maps were the same for both CHO and human processes, suggesting that these maps show general pathways that are conserved across mammalian cells. Gene/protein synonyms were taken from uniprot.org and were combined when feasible. Where 'omic data for a specific gene could not be found in the data sets used no bar is presented in the expression charts. Whilst fold change values are often the average of at least two biological replicates, error bars are not shown as only average values, without standard deviations or errors, were provided in the raw datasets.

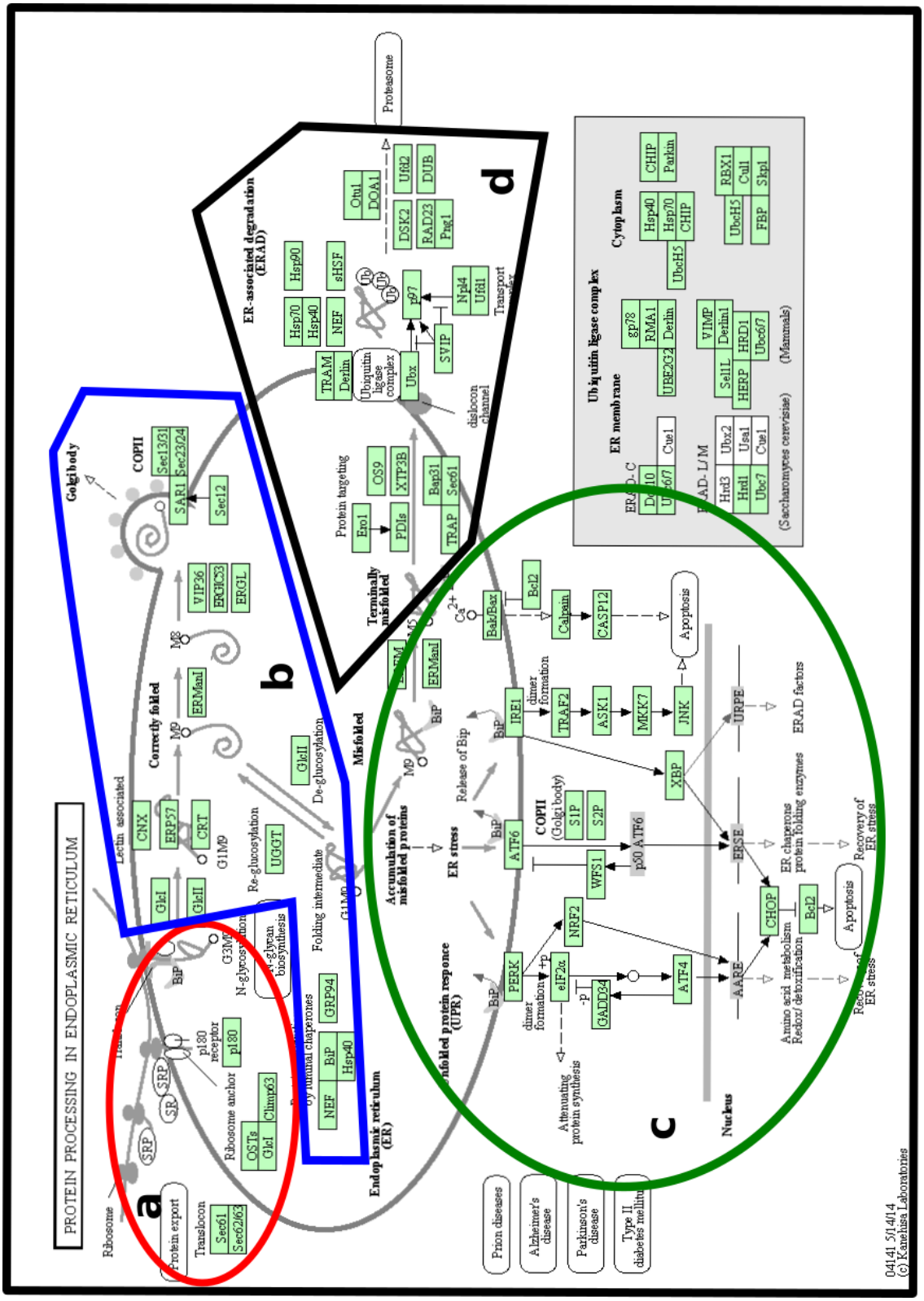


**Figure 3-1. Protein processing in the CHO ER.**

*A molecular map showing proteins with known involvement in ER protein processing, specific to CHO cells, was taken from KEGG. These were split to show [a] Protein import to the ER, [b] protein folding and export from the ER, [c] ER stress response and [d] ER-associated degradation and response to terminally miss-folded proteins. Pathway map taken from KEGG 'Protein processing in ER' in Cricetulus griseus (Kanehisa and Goto 2000). See also Figures 3.3/3.4.*

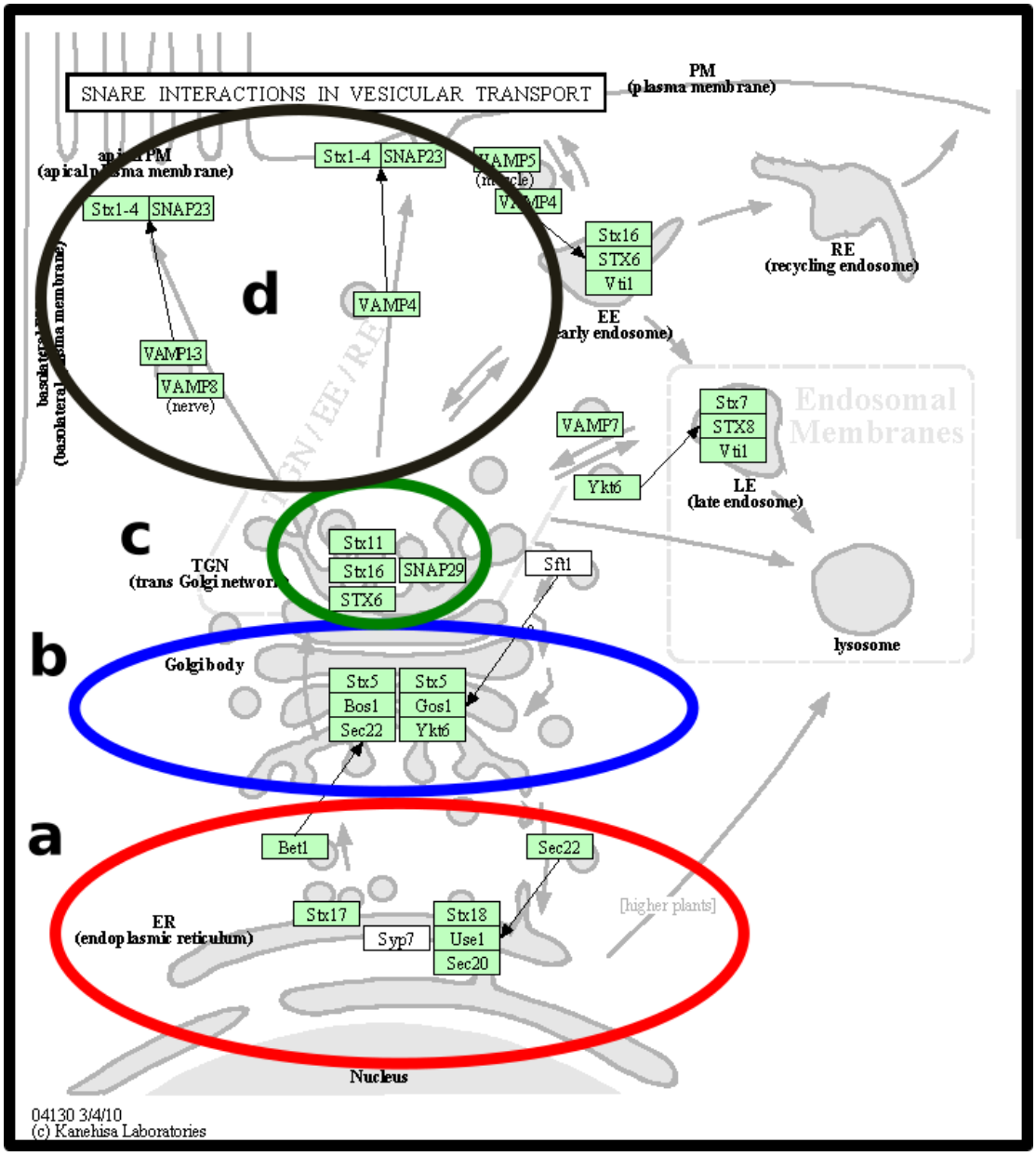


**Figure 3-2. SNARE interactions in vesicle transport.**

*A molecular map showing known SNARE interactions within CHO cells was taken from KEGG. These were split to show [a] ER, [b] Golgi, [c] TGN and [d] plasma membrane associated SNAREs. Pathway map taken from KEGG 'SNARE interactions in vesicular transport' for Cricetulus griseus (Kanehisa and Goto 2000). See also Figure 3.3.*

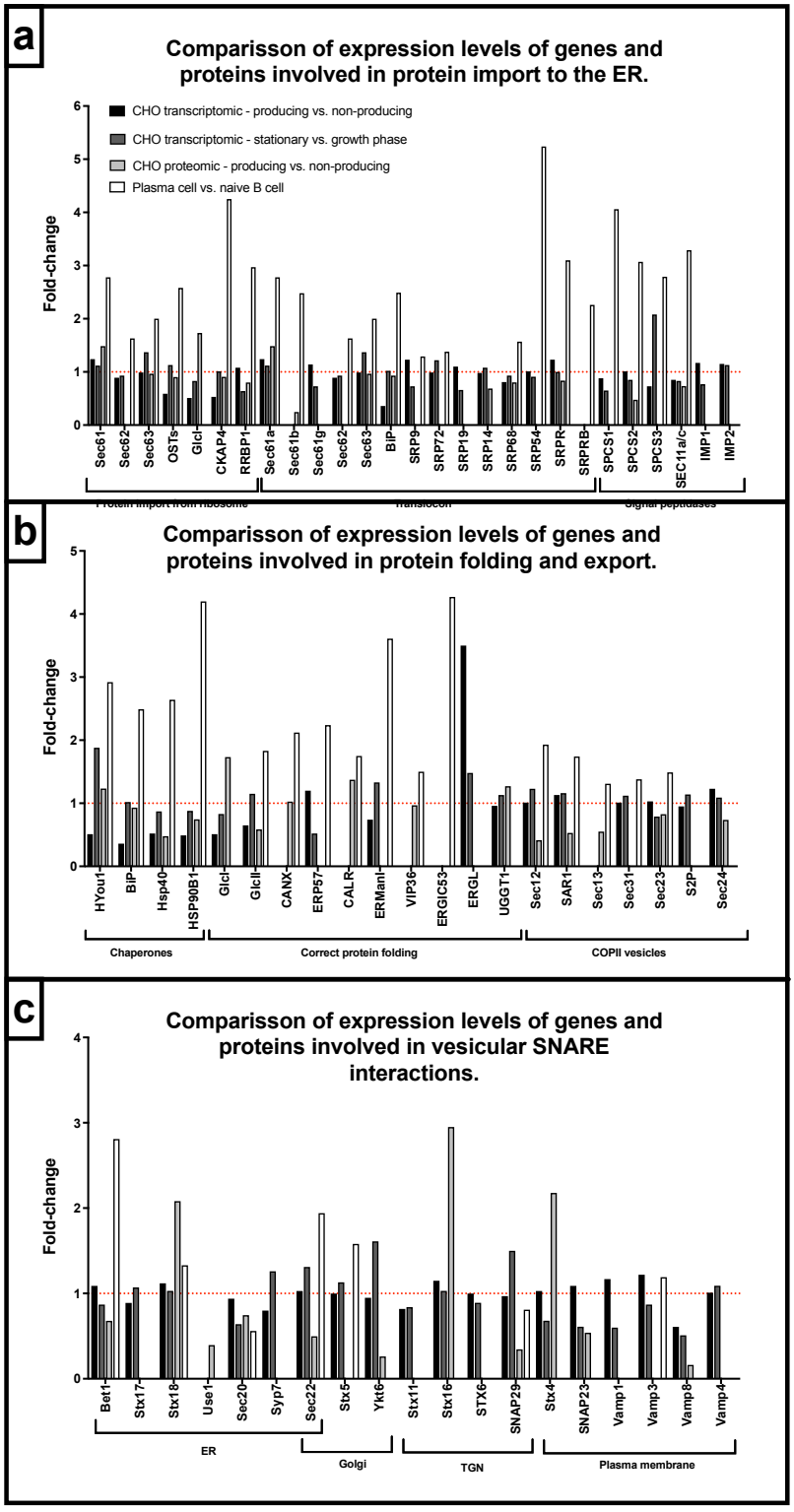


**Figure 3-3. 'omic analysis of CHO cell biosynthetic pathway components.**

*CHO and plasma cell 'omic data was used to analyse the CHO biosynthetic pathway, with fold changes plotted as described in section 3.3. Fold changes show the difference in expression levels between producing and non-producing CHO cells (transcriptomic); producing CHO cells during exponential and stationary growth phases (transcriptomic and proteomic); and naïve B and plasma cells (transcriptomic). [a] Genes with a function in protein transport into the ER, including ribosome docking, translocon and signal peptidases (figure 3-1 a). [b] Genes involved in protein folding and formation of COPII vesicles (figure 3-1 b). [c] Genes involved in SNARE interactions in vesicular transport (Figure 3-2 a-d).*

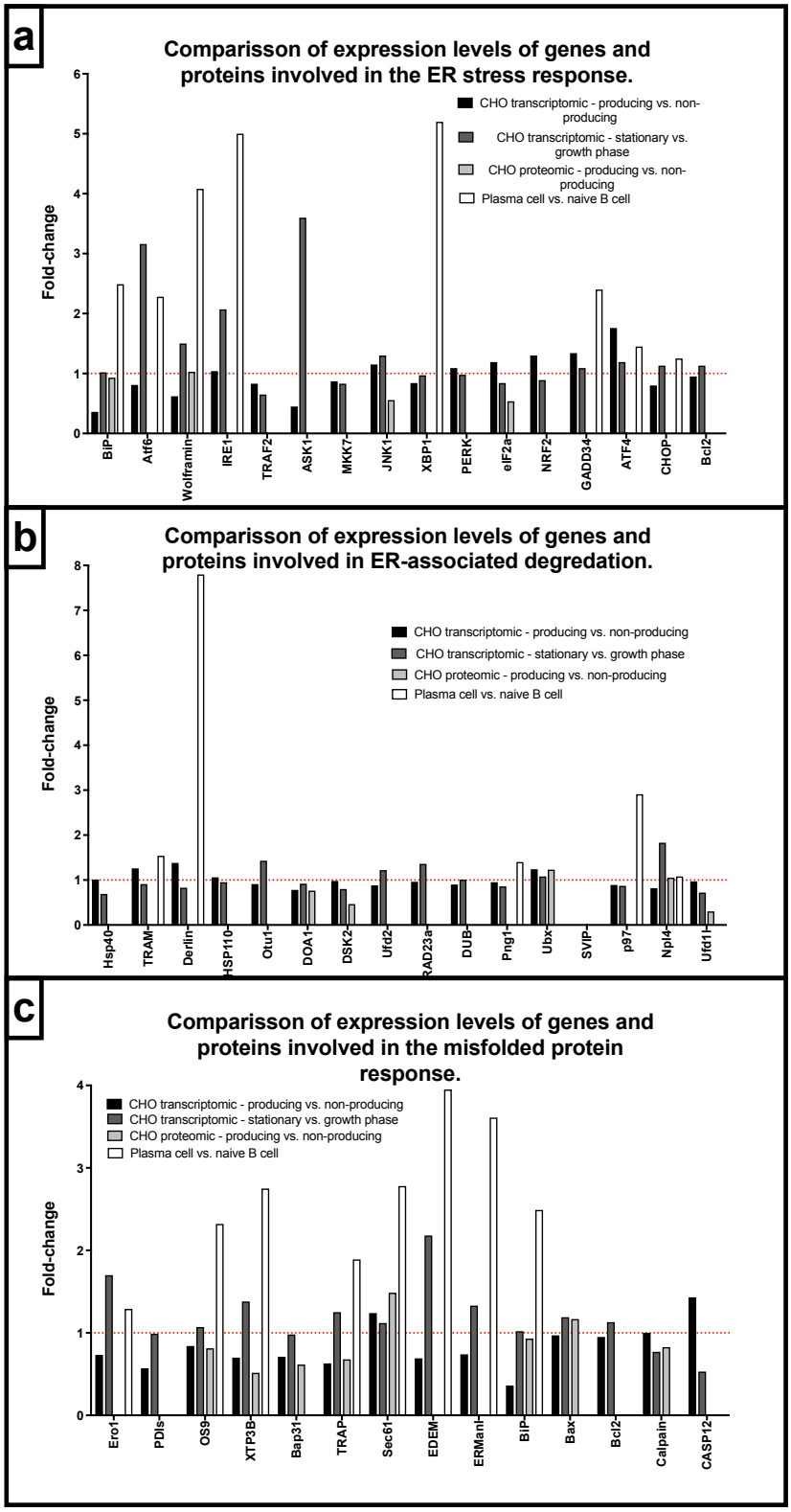


**Figure 3-4. 'omic analysis of CHO ER stress responses.**

*CHO and plasma cell 'omic data was used to analyse the UPR in CHO cells. Columns show fold changes as described in figure 3-3 legend. [a] Genes with a function in the UPR (Figure 3-1 c). [b] Genes involved in ERAD (Figure 3-1 d). [c] Genes involved in processing of terminally misfolded proteins (Figure 3-1 d).*

### 3.3.1. Analysis of protein processing gene levels in Plasma cells

Transcriptional activity changes occurring during the terminal differentiation of naïve B cells to plasma cells can be expected to include genes involved in ER structure, the ER stress response, protein folding, the UPR and secretion. Genes involved in protein import from the ribosome, such as CKAP4 (Climp63) and RRBP1 (p180) see a 2-4 fold increase in transcript abundance in plasma cells, whilst many translocon subunits (e.g. Sec61) see a smaller 1.5-3-fold increase. This shows the adaption of plasma cells to increased numbers of ribosomes and translational levels. The ER stress marker BiP sees a close to 3-fold increase, highlighting the role it has in UPR activation and mediating protein processing and secretory capacity during plasma cell maturation (figure 3-3 a; white bars).

Likewise, genes encoding proteins involved in the correct folding of proteins, such as chaperones, increase in abundance in plasma cells. Hyou1 (NEF), BiP, Hsp40 and HSP90B1 (Endoplasmin) all see a 2-4-fold increase in abundance. Whilst the ERES forming proteins Sec12 (PREB) and Sar1 both see a close to 2-fold increase in abundance, COPII coat subunit genes (Sec 23/24, 13/31) see only a slight increase, suggesting that ERES formation and not vesicle budding is a rate-limiting step in cargo export from the ER in plasma cells (figure 3-3 b). SNARE genes, when present in the plasma cell data set, generally see an increase in expression in plasma cells across all secretory pathway organelles except the TGN. Bet1 and Sec22 see a 2-fold or higher increase in expression levels, whilst Stx18, Stx5 and VAMP3 see smaller increases. However, Sec20 and SNAP29 do see a decrease in expression levels (figure 3-3 c). The loss of data for some SNAREs suggests that the overall CHO SNARE interactions model is not indicative of plasma cell SNARE interactions.

Genes involved in ER stress-response and UPR activation (BiP, IRE1, ATF6, XBP1 and Wolframin, which has a role in response to ER overload) see a large (up to 5-fold) increase in RNA transcript abundance in plasma cells, backing up knowledge that the UPR response does play a role in altering organelle and cell phenotype to one of enhanced productivity (Moore and Hollien 2012; Walter and Ron 2011). UPR effector genes with roles in apoptosis activation (ATF4, CHOP) see a slight increase in abundance, highlighting that the UPR response may need to direct cells to apoptosis, but that this occurrence is not much more common than in normal, non-producing cells (figure 3-4 a, see section 1.5.3). Likewise, many proteins involved in the ER associated protein degradation (ERAD) of misfolded proteins are not present in plasma cells except for Derlin which has a functional ERAD role and sees a 7.8-fold increase in gene activity (figure 3-4 b). Genes involved in recognition of misfolded proteins also see an up to four-fold upregulation, suggesting that pathways that overcome protein misfolding are activated, rather than proteins being directed towards degradation, or the cell towards apoptosis.

### 3.3.2. Analysis of protein processing gene levels in CHO cells

The transcriptomic and proteomic response of CHO cells to recombinant Mab expression sees more subtle changes when compared to that of maturing plasma cells. Most of the genes involved in protein import into the ER were shown to decrease in transcriptomic level in producing CHO cells (figure 3-3 a; black bars). The same pattern was observed for protein folding and misfolded protein response genes, with the exception of ERGL (an LMAN1 isoform), whilst the transcription levels of COPII vesicle components remained relatively unchanged (figure 3-3 b, figure 3-3 c). Whilst BiP and ER chaperones see a decrease in expression levels (to approximately half the level in parental cells), the UPR activators IRE1, PERK, ATF6 and CHOP all remain at a similar level. The only exception amongst this family is ATF4, which sees an approximately 2-fold increase in producing CHO cells compared to non-producers. Over-expression of ATF4 has also been shown to activate apoptosis-directed pathways (section 1.5.3; Gardner et al. 2013).

The transcription levels of ER/Golgi SNAREs and genes involved in protein folding stay at a similar level or decrease slightly when compared to parental cells, although PM-located SNARE levels increase by a small amount (figure 3-3 b, c). Whilst there is little change in expression levels of genes involved in ERAD, there appears to be a decrease in genes that respond to protein misfolding, suggesting that CHO cells do not have the capacity to fully quality control their biosynthetic pathway upon it being overloaded (figure 3-4 b, c).

There are more transcriptomic and proteomic changes seen between CHO cells in exponential and stationary growth phases than between non-producing and producing CHO cells, although again these are still fairly minimal. UPR activator BiP remains at a similar proteomic and transcriptomic level along with the UPR signalers PERK and CHOP, whilst ATF6 and IRE see an increase in transcript abundance, suggesting that these capacity-increasing pathways do see some upregulation during the enhanced productivity levels of the stationary growth phase. However, there is no proteomic data to confirm this pattern carries through to a molecular level (figure 3-4, a; dark grey bars).

Most ER protein import, ER chaperones and protein-processing genes/proteins see little change or a slight decrease in transcript and protein levels, although gene expression of Hyou1 does increase by nearly 2-fold during stationary phase (figure 3-3 a, b; dark and light grey bars). Transcription levels of SNAREs see little change between growth and stationary phases, although protein levels of some Syntaxin family members (Stx-18, -16 and -4) increase by 2-3 fold, suggesting their functionality is controlled post translation (figure 3-3, c). There is little change in levels of ERAD and misfolded protein response genes or proteins (figure 3-4 b, c).

Transcriptomic and proteomic data suggests that there is little increase in the gene transcription and protein levels of genes linked to protein productivity in CHO cells when comparing growth and stationary phases. As the stationary phase is the period in which CHO cells prioritise production over growth, this may be because the cells have already adapted (during the growth phase) to the productivity requirements placed upon it by the transgene.

Comparative 'omic data highlights the transcriptomic difference between CHO cells and plasma cells, with large levels of overexpression of protein folding and UPR effectors detected as B cells mature to plasma cells. Meanwhile CHO cells see little positive change (and in some cases a decrease) in transcriptional levels of genes involved in the UPR and protein processing in response to recombinant Mab expression. Overall, this suggests that there is wide scope for engineering to improve protein processing within the CHO cell ER as, when compared to secretory-specific plasma cells, CHO cells do not appear to respond to their increased metabolic burden or alter much at a molecular level during a production phase of growth. However, this could be due to use of a ETE-expressing cell line for dataset production, as low recombinant expression and increased ER stress levels could have proved different with a DTE Mab, affecting the 'omic data produced.

Due to the only minor changes occurring in producing CHO cell 'omic profiles no obvious gene engineering targets to enhance CHO production levels have come to light from CHO datasets. However, with many large transcriptional changes seen in plasma cells there are more potential targets. ER stress response proteins (e.g. XBP1s, ATF6, BiP) have previously been shown to enhance CHO productivity but may still be useful in a multigene or control capacity (Pybus et al. 2014; Tigges and Fussenegger 2006). Many proteins involved in polypeptide import into the ER (e.g. RRBP1, CKAP and SRPs) are upregulated, with SRPs having already been shown to enhance CHO productivity levels (Le Fourn et al. 2014). Meanwhile chaperones and protein-folding genes are also widely upregulated (e.g. HYou1, HSP90B1, Calnexin, Calreticulin), with some of these having already been shown to enhance mammalian cell productivity, along with related chaperone proteins (Borth et al. 2005; Chung et al. 2004; Ishaque et al. 2007; Josse et al. 2010). Coupled with plasma-cell transcription factors these provide a selection of potential targets with which to engineer the CHO biosynthetic and secretory pathways.

## 3.4. Comparative analysis of wider CHO'omic literature

There are many examples within the literature where transcriptomic and proteomic differences between high- and low-producing CHO cell lines (or

producing and non-producing cell lines), as well as the differences within the CHO cell at different points in cell growth (and therefore productivity level) have been elucidated. As with section 3.3, this allows a view into the molecular changes within the cell that lead to increased productivity, as well as changes the CHO cell undergoes to change from a growing cell (in exponential phase) to one more targeted towards production and secretion of recombinant genes (during stationary phase). A meta-analysis of CHO transcriptomic data has recently been published, performed to map a transcriptomic profile of a high-producing CHO cell. However this found little concordance across the available data sets except for genes involved in the cell cycle and lysosome pathways (Tamosaitis and Smales 2018). Here (with work performed before the publication of the previously referenced paper) genes and proteins that are up- and down-regulated within the datasets selected were sorted into functional groups so as to provide a more directed view as to what is required by the CHO cell to have a enhanced productivity phenotype. This provided guidance towards potential CHO engineering strategies and targets.

Data from nine CHO 'omic papers were combined to find a consensus of genes and/or proteins enriched in high-producing CHO cells which could therefore be considered as functional in increasing CHO cell productivity. Many of these papers took a global approach, providing transcriptomic and proteomic data across all gene types, correlating gene/protein expression levels to high productivity. To better focus data towards the secretory and biosynthetic pathways comparisons were made between data related to transcription factors; protein export; protein folding; ER stress and the UPR; organelle biogenesis; structure and capacity (e.g. lipid metabolism) and secretory pathway elements. The papers used all listed 'omic data for specific genes, generally considered productivity biomarkers. A précis of the literature used for this exercise is shown in table 3-3. Data from three further papers investigating proteomic and transcriptomic changes in producing murine myeloma NS0 cells was also analysed to provide supplementary information on gene function within a producing cell-line other than CHO (Charaniya et al. 2009; Seth et al. 2007; Smales et al. 2004).

To reduce double hits a Uniprot ID (from mouse) was found for all genes to help combine gene synonyms. A threshold of an increased transcript/protein abundance seen in at least three datasets was set to sort for genes that were likely conserved in having a large impact upon CHO cell productivity. From the nine papers listed in table 3-3, 319 genes and proteins with biological functions linked to protein biosynthesis and secretion were found to be upregulated in high-producing CHO cells. Of these, only 10 genes/proteins were upregulated in at least three separate CHO data sets, or upregulated in two separate CHO data sets and at least one NSO or hybridoma MAK dataset. Of these genes, none were shown to be upregulated in the Clarke et al. 2011 paper and as such data from this is not

shown. A table containing all genes for which there is a consensus as to upregulation providing a function in a high production capacity is shown in table 3-4, listing the papers in which they are shown to be upregulated. The fold-change in protein/transcript abundance seen in each paper is shown when available.

Given the pivotal role it plays in activation of the UPR in response to ER stress, BiP (HSPA5, GRP78) is unsurprisingly shown to be upregulated and correlated to enhanced CHO productivity in four different datasets. This is despite some of its increased function initiating protein signaling cascades that can result in apoptosis activation. However, BiP levels are seen to reduce in CHO cells at 33 °C compared to those at 37 °C, a pattern also seen in hybridoma cells that do not increase productivity at 33 °C (Yee et al. 2009). Reasons for this deviation from the norm are unclear but reduced cell growth at 33 °C may lead to a reduction in ATP levels, altering BiP's response to unfolded proteins and the UPR activators it normally binds.

As expected in cells producing Mabs (which contain four disulphide bonds per molecule), there is a consensus across the 'omic datasets that disulphide forming PDIAs 1, 3 and 4 are upregulated in producing CHO and NS0 cells. The folding chaperones HSP90B1 (Endoplasmin) and Hyou1 are also upregulated, as is the related heat-shock protein HSPD1, although this appears to function more in interactions with mitochondrial proteins than general ER cargo. The ER housekeeping proteins Calreticulin and Calnexin, both with roles in protein sorting, folding, oligomer assembly and quality control within the ER, are both upregulated in several different analyses showing their importance in ER maintenance. Upregulation of ACAA2 is explained by its role in halting apoptosis, thus allowing cells to survive the higher ER stress levels associated with recombinant protein production. EEF1a1 enhances elongation of nascent polypeptide chains as they exit the ribosome, speeding up the process of translation, the second (after transcription) potentially rate-limiting step in transgene production.

**Table 3-3. CHO proteomic and transcriptomic datasets used from literature.**

| 'Omic method | Cell types compared | Specific subsets of data used | Reference |
|---|---|---|---|
| Transcriptomic (Microarray) | High IgG1-producing CHO-S/CHO-K1 compared to low-producing lines (3-fold worse) | Transcription factors, Protein export and processing, SNARE interactions. | (Harreither et al. 2015) |
| Proteomic (shotgun) | High producing CHO-DG44 (fusion protein) compared to non-producing parental line. | Protein Metabolism | (Carlage et al. 2009) |
| Proteomic (iTRAQ) | High producing (Mab) CHO-DG44 in exponential phase compared to same cell line in stationary phase. | Chaperone/ protein folding, lipid metabolism and transport | (Carlage et al. 2012) |
| Transcriptomic (Microarray) | High-producing CHO-DUKX producing human bone morphogenic protein (DTE), compared to parental CHO-DUKX | ER and Golgi resident proteins | (Doolan et al. 2008) |
| Proteomic (shotgun) and Transcriptomic (microarray) | High-producing (Mab) CHO compared to low-producing CHO. | Genes with positive correlation to productivity | (Kang et al. 2014) |
| Proteomic | CHO cells with high sustained Mab productivity compared to those with low sustained Mab productivity. | Proteins with potential roles in sustaining high productivity. | (Meleady et al. 2011) |
| Proteomic and Transcriptomic | High-producing CHO compared to low-producing CHO. Both expressing dhfr-GFP fusion. | Protein metabolism, carbohydrate metabolism, 'others'. | (Nissom et al. 2006) |
| Transcriptomic | IgG high-producing CHO cells at 33 °C, compared to same cell line at 37 °C (lower producing). | Lipid/cholesterol metabolism, protein folding, intracellular protein transport, 'others' | (Yee et al. 2009) |
| Transcriptomic | Wide range of CHO cells grown under different conditions expressing variety of products (Mab, fusion proteins, therapeutic factors) | Lipid metabolism, vesicle/secretion. | (Clarke et al. 2011) |

*Nine papers from the literature which studied 'omic changes within high and low producing CHO cells/CHO cell growth stages, were compared to see if there was any consensus as to genes involved in enhanced CHO productivity. This table highlights the nine papers used, what cell lines they were comparing and the 'omic method used.*

**Table 3-4. Consensus of upregulated genes in high-producing CHO cell lines from available literature data sets.**

| Gene name | Uniprot ID | Harreit-her et al. 2015 | Carlage et al. 2009 | Carlage et al. 2012 | Doolan et al. 2008 | Kang et al. 2014 | Meleady et al. 2011 | Nissom et al. 2006 | Yee et al. 2009 | Function |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Maximum fold-change in/correlation to production in high-producing cells | | | | | | | | |
| **BiP** | P20029 | + | +2.8 | +3.45 | + | | | | -1.5 | ER stress response; UPR activation |
| **Pdia3** | P27773 | + | | +2.28 | | | | | | Protein disulphide isomerase; also upregulated in 2 NS0 cell analyses. |
| **Pdia4** | P08003 | + | | +2.36 | + | | | | | Protein disulphide isomerase. |
| **HSP90B1** | P08113 | + | | +2.63 | + | | + | | +1.3 | Endoplasmin, molecular chaperone; upregulated in 2 NS0 analyses |
| **Hyou1** | Q9JKR6 | + | | +2.25 | + | | | + | | Molecular chaperone; upregulated in 2 NS0 analyses |
| **Calr** | P14211 | + | | +2.55 | | | | | | Promotes folding, assembly and quality in ER. Upregulated in 1 NS0 analysis |
| **CANX** | P35564 | + | | +2.07 | + | | | | | Calnexin, ER protein sorting, molecular chaperone |
| **PDIA1** | P09103 | + | | +1.95 | | | | | | Protein disulphide isomerase; upregulated in 2 NS0 analyses |
| **ACAA2** | Q8BWT1 | | | +2.20 | | + | | | +1.5 | Abolishes BNIP3-mediated apoptosis. |
| **HSPD1 (HSP60)** | P63038 | | | | | | + | + | | Mitochondrial protein import and folding. Upregulated in 1 NS0 analysis. |
| **EEF1a1** | P10126 | | | | | | | gene >1.5 protein >1.2 | | Nascent peptide elongation factor. Upregulated in 1 NS0 analysis. |

*Table 3-4 (previous page) shows the 11 genes that were upregulated across at least 2 of the 8 'omic papers analysed and highlighted in table 3-3. Consensus of genes upregulated across several different CHO'omic data sets comparing high and low producing cell lines is shown. If a gene was shown to be upregulated in high producing cells it is shown with a "+", with the fold-change increase in expression level shown if available. Likewise, if a decrease in expression level was seen in high producing cells, a "–" is shown.*

It is somewhat of a surprise that there are only ten genes across the nine papers' combined data that show a consensus in upregulation in producing/high-producing CHO cells. This further highlights the variation there is across different CHO cell lines, sub-populations, recombinant products and production strategies. It is unsurprising that upregulation of several protein chaperones and PDIAs is seen in high-producing cells, especially in the production of Mabs. Of the analyses compared, only one (Harreither et al. 2015) looked at expression levels of transcription factors (XBP1, ATF6, IRE1, PERK) and as such these do not appear in this data comparison.

## 3.5. Generation of gene target list.

A list of genes that could be used for genetic engineering of the biosynthetic and secretory pathways of the CHO cell was produced. This provided an easily accessible and understandable way to capture scientific and data-based knowledge on a wide range of potential gene engineering targets. It allowed comparison of different genes as well as basic information on the gene's function and information on any previous use in genetic engineering to enhance CHO productivity. Genes were selected from literature, pathway mapping and 'omic analysis, including genes mentioned in sections 3.2, 3.3 and 3.4. The gene list also collected GO terms and likely functions linked to the gene to better inform target selection, as well as protein sequences. The gene list is shown in Appendix 1 and is summarised in table 3-5, with genes grouped together based upon their approximate cellular function. In total there are 93 genes in the Appendix 1 list. The gene list was used to inform gene engineering targets that were evaluated in Chapter 4. For each gene listed a *pro forma* was also produced, providing more in-depth information on gene targets, including a short literature review covering its molecular role within the cell. This especially focused upon potential roles in protein production and secretion, as well as any previous use within mammalian cell factories and more in-depth 'omic data covering both CHO and plasma cells. Cellular location of proteins was also taken into account using the Compartments web tool to ensure genes functioned within the expected cellular location in related mouse and rat cells (Binder et al. 2014). The gene *pro forma* are not shown, with as much information as possible captured within the Appendix 1 gene list.

**Table 3-5. Summary of gene target database shown in full in Appendix 1.**

| Target area | Gene targets | Total |
|---|---|---|
| **Chaperones and protein folding** | HSP70; Tor1a; HSP90B1; FKBP11; PDIAs (various); Calnexin; Calreticulin; TXNDC5; DNAJC1; PPIB; HYou1; TMEM59 | 12+ |
| **Transcription factors/Transcriptional control** | TFE3; CREB3L2; XBP1s; MLLT3; BLIMP1; ELL2; YY1 | 7 |
| **Vesicle formation: Cargo loading/sorting** | Sar1; Sec12; Sec16; TMED2; GBF1; SRP family (including translocon); Clec3b; HID1; ASAP2; AP3S1; EMP47; VIP36; Sec23; Sec24d; COPE1; RAB40B | 16+ |
| **Vesicle transport: Tethering and docking** | Sly1; Munc18c; Stx5; VTI1a; SNAP29; YKT6; Rab1; Rab11a; Rab8; Rab26; RabD; Sec3; SCAMP2; SytI; SytII; Bet1; Bet3; SCFD2; P115; VPS45; Gorasp2; Golph3; VAMP4 | 23 |
| **Miscellaneous plasma cell hits** | CRELD2; FNDC3A; FNDC3B; RRBP1; TBL2; NBAS; CAB45; ERGIC53; MCFD2; IRF4; ARID3A; CopZ2 | 12 |
| **Lipid biosynthesis** | PRKD1; CERT; SGMS1; FADS2; PItp; PLD3 | 6 |
| **ER stress response** | Ero1Lα; MZB1; PERK; ATF6; IRE1; eIF2α; ATF4; BiP; CHOP; TRIB3; TMEM214; mTOR; NRF1/2; onco-KIT; GCLM; PRDX4; | 16 |
| **Miscellaneous** | Largen | 1 |
| | **Total** | 93 |

*An overview of Appendix 1 is shown here, showing the 93 genes present on this database and the gene/protein families from which they come from. The gene targets which were chosen to engineer CHO cells were selected from this list.*

## 3.6. Selection of gene engineering targets

Table 3-5 shows the initial gene list from which genes were selected to engineer the CHO cell . A more complete database, containing brief descriptions of each gene is shown in Appendix 1. An online database was produced to better collate all literature, 'omic and GO information on each target..

### 3.6.1. Gene selection strategy and procedure

Accessory/effector genes with which to engineer the CHO secretory pathway were selected using information on their function and pathway gleaned from literature, pathway mapping and transcriptomic and proteomic data from both CHO and plasma cells. These techniques were also used to select genes with which to more directly engineer non-secretory regions of protein processing within the ER. Whilst this will be explained and analysed here, the main focus will be on specific secretory pathway engineering.

Accessory genes were selected to focus engineering on four defined areas of the secretory pathway:

1. Early secretory pathway: Selection and loading of cargo into COPII vesicles at the ERES; vesicular transport through the ERGIC; vesicle fusion at the Golgi.
2. Late secretory pathway: Transport of cargo through the Golgi stack and on to the Plasma Membrane (vesicle formation, transport and fusion).
3. Transcriptional control: Transcription factors expected to have a role in the expansion of the secretory pathway.
4. Lipid Biosynthesis: Engineering of the CHO lipid biosynthesis and transport systems with a view to increasing organelle volume and transport requirements.

Genes were selected (for the most part) upon two premises. Firstly, due to the expectation they would directly impact upon CHO productivity when expressed singularly. Secondly, based upon pathway mapping and prior knowledge, to function alongside other genes within/across a specific pathway (in an attempt to maintain stoichiometry) to enhance productivity. Furthermore, effective genes could be co-expressed to see if engineering of separate pathways and bottlenecks resulted in a stacking effect of improvements, further elevating CHO productivity. A number of genes previously shown to enhance productivity were also selected both to act as a positive control but also to further test them in a multigene system, test sequences taken from a different origin and to see whether results from literature could be replicated. Some 'wildcard' genes were also selected based on literature suggesting they aided a phenotype linked to

enhanced productivity (e.g. cell size) despite little previous use or knowledge about their exact function.
The origin of an accessory gene sequence has an important role in the ability for it to function within a CHO cell expressing a recombinant humanised antibody. Whilst biosynthetic and secretory machinery is relatively well conserved across mammalian cells differences in molecular machinery do exist. It is possible that these could have an impact upon accessory gene function within the cell. If the function of the accessory gene's product involves direct interaction with the recombinant biologic it stands to reason that a human-derived gene sequence was used to better facilitate this interaction and its function.

As such, genes encoding ER chaperones and catalysing disulphide bond formation, as well as those that interact with these effectors, were encoded by human-derived gene sequences. Many of these proteins will of course also interact with other CHO ER machinery but human sequences were selected based upon their primary interaction with the recombinant product. Engineering targets known to function as transcription factors or within the CHO secretory pathway interact directly with CHO molecular machinery and so were expressed from CHO-derived sequences. Some of the target genes (especially those involved in protein processing and the early stages of vesicle formation) interact with both CHO machinery and the humanised recombinant product. In these cases a decision as to what sequence to use was made based upon the perceived importance of each interaction, though with the back-up that conservation of components between mammalian cells would allow some level of functionality to remain. For genes selected from plasma cells if a CHO homologue was not present the human sequence was used.

### 3.6.2. Selection of targets for secretory pathway engineering

A total of 16 gene targets were selected with which to engineer the CHO secretory pathway. The selection process utilised the gene target database produced in chapter 3 and shown in Appendix 1. The gene database list was narrowed to a testable number of gene targets by further literature research on the gene's function. However, not all gene targets are listed on the database as this was mainly used to gather known information from literature and 'omic data. As such, some genes known to play a key role in the secretory pathway but not yet shown to have a potential link to productivity have been selected as part of a selection strategy driven by both specific gene and overall pathway knowledge, as well as a desire to test some unknown targets.

Here the genes selected for CHO engineering are briefly reviewed to explain their selection, being highlighted in bold. Selections are summarised in tables 3-6 and 3-7 in a more concise manner than the information shown in Appendix 1.

#### *3.6.2.1. The early secretory pathway*

The early secretory pathway, transporting proteins from the ER to the Golgi via the ERGIC, is reviewed in depth in section 1.4.1. Some of the reasons for the selection of specific genes are discussed and listed in table 3-6

**Cargo loading at ER.** Correct sorting of cargo into COPII vesicles initiates the secretory pathway. Increasing the amount of ERES and therefore COPII vesicle budding events would increase transport from the ER, alleviating the burden of increased protein levels within the ER. The GEF **PREB** (the homologue of the yeast Sec12 gene) activates the Sar1 protein which acts as a focal point for ERES/COPII vesicle formation, with Sec12 required for ER-Golgi protein transport in yeast (Jensen and Schekman 2011; Nakano et al. 1988). Increased PREB concentration at ERES also allows secretion of proteins (such as collagen) normally too large for COPII vesicles (Saito et al. 2014). It has, alongside Sar1, been associated with increased CHO production capacity and is up-regulated in plasma cells (Harreither et al. 2015). As such increasing PREB levels should increase Sar1 activation levels and therefore the number of ERES and therefore try and enhance ER-Golgi vesicle numbers. It may also aid the formation of larger transport vesicles which could potentially aid production of large recombinant proteins.

**ER-Golgi vesicular transport.** Anterograde transport between the ER and Golgi occurs via the ERGIC (sections 1.4.1.2, 1.4.1.3). The transmembrane lectin **LMAN1** (ERGIC53) has been linked to increased productivity in CHO cells (Harreither et al. 2015). It is present in a 1:1 stoichiometry with **MCFD2**, a luminal interaction partner. Together they form a cargo receptor complex (Nyfeler et al. 2006). The two proteins are linked to the transport of blood factors, glycoproteins and lysosome-resident enzymes, with both genes upregulated in plasma cells (Appenzeller et al. 1999; Nyfeler et al. 2008; Vollenweider et al. 1998; Zhang et al. 2005).

**Vesicle tethering and fusion at the ERGIC/Golgi.** Tethering of vesicles at the *cis*-Golgi brings target and vesicle SNARE proteins into close proximity, allowing vesicle fusion. GEF-activation of **Rab1** by **Bet3** allows it to interact with ERES, recruiting **P115** (USO1) tethers to COPII vesicles (Allan et al. 2000a). P115 promotes tethering at the ERGIC and *cis*-Golgi via interaction with target GRASP/GM130 tethers. P115 knockdown inhibits cargo trafficking and it also has an activating role in the vesicle docking and fusion process (Sztul and Lupashin 2006). Rab1 sees an increase in plasma cells and has roles in mediating other SNARE and tether interactions (Sztul and Lupashin 2009). The **Bet1** protein is a SNARE known to function in correct protein sorting in yeast (Morsomme et al. 2003). It interacts with P115 and alongside Syntaxin 5 (**Stx5**) facilitates the fusion of COPII vesicles to each other (to form transport

intermediates) and the *cis*-Golgi membrane. Stx5 can be activated by P115 and also interacts with the COG tethering complex (Hong and Lev 2014; Szul and Sztul 2011). Bet1 and Stx5 have been shown to play a role in increased CHO cell productivity (Harreither et al. 2015).

#### *3.6.2.2. Late-secretory pathway targets*

Transport through the Golgi and on to the plasma membrane is not as well defined as ER-Golgi transport. The **Rab11a** protein can not be found in the CHO or plasma cell datasets used in this project, although some interacting proteins are. It is associated with both constitutive and regulated secretion and is upregulated in tissues with high secretion levels (Urbe et al. 1993). It is involved in post-Golgi exocytosis and endocytosis, interacting with the exocyst tethering complex and motor proteins. However, it also plays a role in trafficking from endosomes (Junutula et al. 2004; Sztul and Lupashin 2006; Welz et al. 2014; Zhang et al. 2004).

#### *3.6.2.3. Transcription factors and plasma-cell derived targets*

Plasma cell transcription data was mined for novel targets to enhance the secretory pathway, with transcription factors particularly of interest (section 3.2.1). **CRELD2** is ER- and Golgi-located and highly expressed in endocrine tissues as well as plasma cells. It is ER-stress inducible, with suggested functions downstream of ATF6. It has been shown to specifically regulate trafficking of acetylcholine receptor subunits (Maslen et al. 2006; Oh-hashi et al. 2009; Ortiz et al. 2005). It has been shown to have some PDI-like qualities as well as interacting with chaperones and foldases (Hartley et al. 2013). The **CREB3L2** transcription factor is also upregulated in plasma cells. It binds the enhancers of secretory pathway genes and, along with related family members, has been shown to up-regulate secretion in non-secretory insect and HEK cells (Barbosa et al. 2013; Fox et al. 2010; Panagopoulos et al. 2007). The **TFE3** transcription factor has a role in the Golgi stress response. Activated (by dephosphorylation) by the onset of Golgi stress, it binds the GASE and drives translation of Golgi-related genes involved in glycosylation, structure and vesicle transport. However TFE3 has yet to be expressed recombinantly outside of a human or mouse cell line (Oku et al. 2011; Sasaki and Yoshida 2015; Taniguchi et al. 2015).

The spliced version of human **XBP1** has previously been used to enhance CHO productivity, but a CHO-derived version of XBP1 has not yet been used (Becker et al. 2008; Cain et al. 2013; Tigges and Fussenegger 2006). Use of a CHO-derived XBP1s sequence may improve the interactions it can make with gene promoter regions, further enhancing the effect it has upon CHO-cell productivity. A human-derived XBP1s gene will also be used to test this hypothesis and to act as a control.

#### *3.6.2.4 – Lipid biosynthesis pathway targets*

Expansion of organelle volume should result in an increased capacity for both protein folding and modification as well as providing a larger surface area from which vesicle budding can occur. Expression of an activated form of the ceramide transfer protein (**CERT**) helped increase CHO productivity through enhancing lipid transport to the Golgi (Florin et al. 2009; Rahimpour et al. 2013). Sphingomyelin synthetase (**SGMS1**) converts ceramide in to DAG within the Golgi, the conical shape of DAG aiding vesicle budding. DAG also recruits **PRKD1** to the Golgi, which helps control the budding of tubules and vesicles at the TGN. A positive feed-back loop through lipid kinases further enhances DAG production (Bard and Malhotra 2006; Fugmann et al. 2007).

### 3.6.3. Selection of targets for biosynthetic pathway engineering

Fourteen genes were selected with which to directly engineer the CHO biosynthetic pathway. Some of these genes had been previously tested in CHO cells and as such were selected to both confirm previous results within our platform, but also with a view to their use as part of multigene combinations. Some of the reasons for specific gene selection are discussed with information taken from literature. Selected targets are listed in table 3-7 and are included in Appendix 1.

#### *3.6.3.1. Chaperones*

Increasing the level of chaperones that aid and increase the rate of protein folding should increase the rate at which complicated recombinant proteins can be produced within the CHO cell. **HSP90B1** (also known as GRP94/Endoplasmin) aids in the processing of secretory proteins. It is upregulated in plasma cells and has been linked to increased production capacity in CHO cells in three separate studies (Harreither et al. 2015; Meleady et al. 2011; Yee et al. 2009). **HYou1** has a cyto-protective role in response to oxygen deprivation as well as a potential role in protein folding. It is upregulated in plasma cells and high-producing CHO cells (Nissom et al. 2006). Torsin A (**Tor1a**) is an ATPase with chaperone function. Its overexpression in CHO cells has already been shown to increase recombinant protein production depending upon the model protein used (Josse et al. 2010).

**Calreticulin** and **Calnexin** are ER-located protein-folding proteins that are both increased in plasma cells and linked to increased productivity in CHO cells (Harreither et al. 2015). Their co-expression in CHO cells resulted in an increase in total product levels within the cell, although secreted levels did not increase (Chung et al. 2004). **FKBP11** accelerates protein folding during protein synthesis and is upregulated in plasma cells (Shaffer et al. 2004). **CypB** is a molecular chaperone with some roles in fatty acid metabolism. Co-expression with BiP and PDIs can increase CHO productivity (Johari et al. 2015; Pybus et al. 2014).

Overexpression of **HSPA1a** (HSP70) has been shown to increase productivity of large blood factors in BHK cells by 50% through secretion enhancement and increased apoptosis resistance (Ishaque et al. 2007).

#### *3.6.3.2. PTM enzymes*

Post-translational modifications such as disulphide bond formation and glycosylation are essential for the production of a functional product. Enhancing the cellular machinery that performs these modifications should increase the cell's ability to produce high-quality product. Protein disulphide isomerases such as **PDIA4** catalyse the formation of the disulphide bonds essential for Mab structure and function. All PDIs are upregulated in plasma cells and PDI overexpression has been shown to increase CHO productivity (Borth et al. 2005; Johari et al. 2015; Pybus et al. 2014). PDIA4 is linked to increased productivity in CHO cells (Harreither et al. 2015). **Ero1-Lα** transfers disulphide bonds to PDIs (Hussain et al. 2014). Co-expression alongside XBP1s resulted in an increase in CHO productivity (Cain et al. 2013).

#### *3.6.3.3. Transcription factors*

Overexpression of ER-based transcription factors should allow global control of many pathways involved in protein folding. As discussed in section 4.2.2, **XBP1s** up-regulates both secretory and protein processing components (table 4-1). **YY1** is believed to alter the epigenetic profile of cells and its overexpression in various CHO and human cell lines has increased productivity by up to six-fold (Tastanova et al. 2016). The transcription factor **Largen** does not specifically target the ER or protein processing. It is thought to be involved in regulation of cell size and mitochondrial respiration and overexpression in mouse cell lines has resulted in increases in cell size (Yamamoto et al. 2014). Were it to have a similar effect in CHO cells it could increase cell capacity and thus productivity. It has been shown that CHO cells with a larger volume (produced thorough hypothermic adaption) having been shown to have increased productivity levels compared to smaller cells (Syddall et al., manuscript in preparation).

#### *3.6.3.4. Miscellaneous targets*

Members of the signal recognition complex family aid ribosome docking at the ER and transport of nascent polypeptide chains from the ribosome into the ER lumen via the translocon. **SRP14** has been linked to increased productivity in CHO cells and is upregulated in plasma cells. SRP14 overexpression increases CHO productivity seven-fold, with co-expression alongside other SRPs further enhancing productivity (Harreither et al. 2015; Le Fourn et al. 2014). Perioredoxin peroxidases catalyse the hydrolysis of hydrogen peroxide, protecting cells against oxidative stress. **PRDX4** is upregulated in plasma cells under the control of XBP1 (Shaffer et al. 2004).

**Table 3-6. Overview of secretory pathway accessory genes for CHO engineering.**

*Table shows accessory genes chosen to engineer the CHO secretory pathway. More detail is available in Appendix 1. H/C = human or CHO derived sequence used; H = human; C = CHO.*

| Gene | Function | Interactions | Chosen from | H/C |
|---|---|---|---|---|
| **Early secretory pathway** | | | | |
| **PREB (Sec12)** | Activates Sar1 in ERES formation | CHO machinery | Secretory pathway mapping; increased in plasma cells. | C |
| **LMAN1 (ERGIC 53)** | ERGIC receptor | Cargo, membranes, MCFD2 | Transcriptomic increase in plasma and producing CHO cells. | H |
| **MCFD2** | ERGIC receptor | Cargo, cellular machinery | Transcriptomic increase in plasma cells | H |
| **Rab1a** | Tether recruitment to ERES | ERES, tethers, membranes | Secretory pathway mapping | C |
| **P115** | Vesicle tether | Membranes, SNAREs, tethers, Rab1 | Transcriptional increase in plasma cells | C |
| **Bet3** | Rab1 activator; MSC tether subunit | Rab1 | Secretory pathway mapping | C |
| **Bet1** | SNARE protein | Stx5, COPII vesicles | Increased in plasma cells and producing CHO cells. | C |
| **Stx5** | SNARE protein | Bet1, P115, COPII vesicles | Transcript levels increased in plasma cells and producing CHO cells. | C |
| **Late-secretory pathway** | | | | |
| **Rab11a** | Golgi-PM transport | Exocyst components | Secretory pathway mapping | C |
| **Plasma cell-derived targets/Transcription factors** | | | | |
| **CRELD2** | ER stress, signaling cascades, chaperone interaction | Chaperones, signaling molecules | Large transcript increase in plasma cells | C |
| **TFE3** | Transcription factor regulating Golgi stress response. | DNA (GASE) | Shown to regulate Golgi stress response | C |
| **CREB3L2** | Transcription factor binding secretory gene regulators | DNA, ER/Golgi membranes | Shown to activate secretory genes | C |

| Gene | Function | Interactions | Chosen from | H/C |
|---|---|---|---|---|
| **XBP1s** (2 forms) | Transcription factor regulating ER stress response | DNA | Overexpression shown to increase CHO productivity | C/H |
| | **Lipid Biosynthesis** | | | |
| **CERT** | Ceramide transport protein. | Lipids, Golgi | Shown to increase CHO productivity | C |
| **SGMS1** | Converts ceramide in to DAG (and sphingomyelin) | ceramide, DAG | Lipid transport pathway mapping | C |
| **PRKD1** | Controls vesicle budding at TGN | DAG, membrane. | Lipid transport/secretory pathway mapping | C |

Table 3-7. Overview of protein processing accessory genes for CHO engineering.
*Table shows accessory genes chosen to engineer the CHO biosynthetic/protein processing pathway. More detail is available in Appendix 1. H/C = human or CHO derived sequence used; H = human; C = CHO*

| Gene | Function | Interactions | Chosen from | H/C |
|---|---|---|---|---|
| | | **Chaperones** | | |
| **HSP90B1 (GRP94)** | Processing/transport of secreted proteins | Recombinant proteins, other ER machinery | Increased transcript abundance in high-producing CHO and plasma cells | H |
| **HYou1** | Cell protection; Protein processing | Recombinant proteins, other ER machinery | Increased transcript levels in plasma cells | H |
| **Tor1a** | Protein folding | Recombinant proteins other ER machinery | Previous expression in CHO cells increased productivity | H |
| **Calreticulin** | Protein folding, oligomeric assembly and quality control | Recombinant proteins, calcium, other ER machinery | Expression increased CHO productivity. Transcriptomic increase in plasma cells. | H |
| **Calnexin** | Protein folding, quality control | Glycoproteins, calcium, other ER machinery | | H |

| Gene | Function | Interactions | Chosen from | H/C |
|---|---|---|---|---|
| **Chaperones (contd.)** | | | | |
| **FKBP11** | Accelerates protein folding | Recombinant proteins, other ER machinery | 4-fold increase in transcript abundance in plasma cells | H |
| **CypB** | Chaperone; fatty acid metabolism | Recombinant protein; lipid synthesis machinery | Previously shown to increase CHO productivity | H |
| **HSPA1a (HSP70)** | Protects against apoptosis through interaction with caspase cascades | Apoptotic signaling cascades | Previous expression in BHK cells increased productivity | C |
| **PTM enzymes** | | | | |
| **PDIA4** | Protein disulphide isomerase | Recombinant proteins, other ER machinery | Shown to increase CHO productivity levels | H |
| **DAD1** | Glycosylation | Recombinant proteins, other ER machinery | | C |
| **Ero1-Lα** | Transfers disulphide bonds to PDI | PDI, other ER machinery | Shown to increase CHO productivity; | H |
| **Transcription Factors** | | | | |
| **YY1** | Epigenetic alterations | DNA promoter regions | Expression saw 6-fold increase in CHO factors | C |
| **PRR16 (Largen)** | Cell size /mitochon-drial function | DNA promoter regions | Shown to regulate cell size | H |
| **Miscellaneous** | | | | |
| **SRP14** | Signal recognition; import of nascent polypeptides to ER lumen | Ribosomes, nascent polypeptides | Shown to increase CHO productivity | C |
| **PRDX4** | Hydrogen peroxide reduction | Hydrogen peroxide | Transcriptomic levels increase in plasma cells | H |

## 3.6. Discussion and conclusions.

Analysis of plasma cell transcriptomic data has shown that many molecular components involved in secretion see an upregulation at the transcript level as B cells mature into plasma cells. The data further backs up existing literature suggesting the UPR plays a role in enhancing the biosynthetic and secretory pathways of secretory-specific cells, with genes involved in its initiation (BiP) and signaling cascades (XBP1, IRE1, ATF family members and CHOP) all seeing an up-to a 5-fold increase in transcriptional level in plasma cells (Moore and Hollien 2012; Shaffer et al. 2004; Shapiro-Shelef and Calame 2005).

More specific analysis of transcription factors showed that of those upregulated the majority were linked to the secretory and biosynthetic pathways (table 3-1, figure 3-3, 3-4; white bars). That some of these factors have previously been shown to enhance transcription of secretory pathway components provides a solid foundation upon which to base an engineering strategy to enhance CHO cell productivity (Barbosa et al. 2013; Fox et al. 2010). It can therefore be hypothesised that utilisation of a small number of transcription factors will allow upregulation of a large number of secretory pathway genes, with each one regulated at a level determined by the cell itself. This would ensure that stoichiometric ratios within the cell remained balanced. Furthermore expression of a single transcription factor would reduce the metabolic burden upon the cell compared to overexpression of many single effector genes.

Comparison of non-producing and stably-producing CHO cells suggests that there is generally little difference in gene expression between the two when comparing the biosynthetic and secretory pathways. This was not entirely unexpected as it is known CHO cells are not specialised for excessive protein production and secretion and as such at most a moderate increase in protein processing and vesicle transport factors was hypothesised. However, with regards to the UPR and ER stress response this was somewhat unexpected. It was hypothesised that expression of a recombinant Mab would overload the ER, resulting in a tangible ER stress response. It was also surprising that the ER stress sensor BiP and ER chaperones were downregulated (figure 3-4 a). This lack of a response could be due to the recombinant protein in question being easy to express and therefore not overly burdening the ER or its quality control pathways.

It is also possible that long term stable production of a recombinant Mab places the cell under a level of chronic ER stress that results in a feedback suppression of the ER stress response (Gomez and Rutkowski 2016). It could also be the case that, during the clonal selection process by which the producing cell line was made, the high-producing clone selected had a higher tolerance to ER stress

before activating the UPR, which in part resulted in its high productivity. When comparing the transcriptomic differences CHO and plasma cells undergo upon initiation of Mab synthesis, it is clear that CHO cells undergo few changes to aid Mab production and are far away from being specialised in this process.

Comparison of producing CHO cells during exponential and stationary phases of growth shows more 'omic changes than are seen between parental and producing CHO cells. The increased expression of some UPR response genes (e.g. ATF6, IRE1) show that during the stationary phase there is an increase in the burden placed upon the ER. This increase in expression correlates with the most productive part of the cell's growth cycle, although it is interesting that some other UPR activators (XBP1, CHOP, PERK) do not see a similar increase. Likewise there is no concerted increase in abundance of UPR effectors (chaperones, protein folding proteins) or vesicle transport proteins.

This data suggests again that when required to produce a recombinant protein the CHO cell is able to do so but is unable to make sweeping transcriptomic and proteomic changes to enhance overall productivity, again highlighting CHO's lack of specificity for production, especially when compared to plasma cells. The lack of large 'omic changes within producing CHO cells (either due to Mab expression or due to growth stage) suggest that there is little value in choosing genetic engineering targets due to their preponderance in producing CHO cells. However this data could be used to rectify 'omic deficiencies seen in CHO cells that are not present within plasma cells with the intention of widening biosynthetic and secretory bottlenecks.

These CHO and plasma cell data in combination show that there is some scope to increase in CHO cells the levels of genes known to function within the UPR, protein processing and vesicle transport, as they do not already see large increases in expression levels upon production of a recombinant Mab. As such it could be considered that there is no one specific bottleneck within the CHO biosynthetic and secretory pathways inhibiting productivity. Instead all levels of the protein production process, including organelle volume; UPR response; chaperone density; and the levels of vesicle components, combine to produce a productivity phenotype that as a result of many bottlenecks in a series leads to a production level lower than it could be.

The knowledge that the UPR initiates an increase in ER and Golgi volume, protein chaperones and folding machinery levels, resulting in increased productivity in plasma cells, highlights the role that global transcription factors can have in enhancing cellular productivity. Utilisation of these factors is a strong candidate for improving cell productivity. However, there are likely to be only a handful of blockbuster single-gene targets that can effect the desired large-scale changes

upon the CHO cell to increase its productivity. As such, engineering of specific parts of the biosynthetic and secretory pathways, informed by knowledge of molecular interaction that occur within them, is still a valid strategy to increase capacity and productivity, as has been shown with SRPs, SNAREs, SNARE effectors and lipid transport engineering (Florin et al. 2009; Le Fourn et al. 2014; Peng et al. 2011; Peng and Fussenegger 2009b; Rahimpour et al. 2013). However, due to limitations in the number of direct effector genes that can be ectopically expressed at once, the pathway regions chosen for specific engineering will have to be selected based on a small number of proteins having a large effect upon the pathway. Plasma cell transcriptomic data can be used to inform these decisions, alongside secretory pathway knowledge.

This chapter produced information on the CHO cell's 'omic response to recombinant Mab production, comparing this to that of the secretory specific plasma cell. Combinatorial analysis of other CHO 'omic analyses within the literature further enhanced this insight. The information presented in this chapter and Appendix 1 will be used in proceeding chapters to help the knowledge- and data-driven selection of effector genes with which to engineer the CHO cell biosynthetic and secretory pathways.

# 4) Genetic engineering of the CHO secretory pathway

Some of the work in the chapter was carried out alongside Claire Arnall (CA) and Joe Cartwright (JC). The high-throughput gene screen platform was developed by CA. The single-gene screen, in both transient and stably-expressing cell lines, was carried out alongside JC and CA. Some gene selection and design work (specifically those targeting the non-secretory functions of the ER) was carried out by CA independently of the gene target list produced in chapter 3 by NB. Genes were designed so as to work in the single gene and MGEV expression vector platforms designed by Yash Patel (YP) and CA.

*Genes were selected with which to engineer the CHO biosynthetic and secretory pathways. An initial transient single-gene screen of 32 genes specifically targeting the biosynthetic and secretory pathways was carried out to determine the effect each gene and their dose level had upon CHO production of a DTE Mab being both transiently- and stably-expressed. This also allowed comparisons to be made between the response of the two expression systems to genetic engineering, especially of the unfolded protein response. Accessory genes were then co-expressed in two/three-gene combinations that were selected based on both the single gene screen data as well as scientific knowledge of the secretory and biosynthetic pathways.*

*The two single-gene screens showed that the effect of accessory genes in cells transiently expressing a model DTE Mab is not repeated in cells stably expressing the same Mab. However some genes did enhance production levels of DTE-Mab being expressed transiently and stably. Multigene engineering of the CHO biosynthetic and secretory pathway in a cell line transiently expressing a DTE-Mab showed that some gene combinations can concurrently increase CHO productivity, but that specific targeting of a molecular pathway involved in constitutive secretion is unlikely to bring about significant increases in production.*

## 4.1. Introduction

Recombinant protein expression places an increased metabolic burden upon the CHO cell. Engineering to enhance transcription and translation increases the amount of proteins required to be processed by the cell's biosynthetic and secretory pathways, further increasing the metabolic burden upon the cell's processing and secretory pathways. That the relationship between mRNA transcript level (for both $F_c$ fusion and EPO model proteins) and Qp of CHO cells is non-linear highlights the presence of a post-transcriptional bottleneck in the production and secretion of recombinant proteins (Hansen et al. 2017; Johari et al. 2015; Ku et al. 2008). As such, enhancing cellular levels of transcription and translation may well increase cellular productivity, but increased flux in these early stages of protein production will likely further clog any current bottlenecks within the protein processing and secretory system (Brown et al. 2014; Quax et al. 2015).

### 4.1.1. Previous engineering of the CHO cell biosynthetic and secretory pathways

Engineering and improvement of host CHO cells, to a certain extent, has likely resulted in all super-high producers, with directed evolution strategies and genetic enhancements used to improve cell viability, growth levels, the ability to carry transgenes and utilisation of selection strategies (Fischer et al. 2015). To cope with an increased flux of proteins to process and secrete, several genetic engineering strategies have been undertaken with which to augment the CHO biosynthetic and secretory pathways and increase productivity. That the Qp of professional secretory plasma cells can be up to four-fold higher than that of CHO cells further highlights that the physiological limit of CHO cell production capacity has not been reached (Hansen et al. 2017).

The recombinant overexpression of effector genes, either CHO genes functioning within these pathways or more novel genes taken from other organisms, has been shown to change organelle morphology and enhance processing and transport processes, resulting in an overall enhanced productivity phenotype. Genetic engineering of the UPR has seen an increase in cellular productivity, but if incorrectly facilitated can lead to apoptosis. Ectopic overexpression of transcription factors can allow the fine alteration of many requisite molecular components through enhancing the cell's usual transcription mechanisms, ensuring molecular stoichiometry remains balanced and functional. For example, overexpression of ATF6 results in a 2.5-fold increase in Mab productivity, whilst BiP overexpression resulted in a 1.5-fold increase (Johari et al. 2015; Pybus et al. 2014). Overexpression of spliced human XBP1 has produced a five-fold increase in Mab production (larger increases were seen using SEAP and SAMY models),

with ER volume also seen to increase (Becker et al. 2008; Cain et al. 2013; Tigges and Fussenegger 2006).

Engineering of ER stress response can also be approached by increasing cellular levels of glutathione, an antioxidant that protects DNA, proteins and metabolites against oxidative damage, as well as helping the formation of disulphide bonds and maintaining ER oxidoreductases in a reduced state. Increasing levels of GCLC (a gene involved in the first step of glutathione synthesis) in CHO cells increased Mab Qp and titre by up to 1.95-fold (depending upon growth stage), as well as increasing the number of high-producing clones by 75% (Orellana et al. 2017; Orellana et al. 2015). The use of chemical chaperones and regulators can also be used to enhance the biosynthetic pathway, with small molecule proteostasis regulators shown to increase the activity of endogenous ATF6, or at least match the ATF6-activated phenotype without actually activating it (Plate et al. 2016).

More directed engineering of the biosynthetic and secretory pathways has also been shown to increase CHO productivity. Overexpression of SRPs and translocon elements improved the transfer of nascent polypeptides into the ER, increasing DTE productivity by up to seven-fold (Le Fourn et al. 2014). Increasing levels of ER-resident chaperones, foldases and disulphide bond forming enzymes has generally resulted in productivity increases. Torsin 1A overexpression increased levels of secreted luciferase (2.5-fold) and Mab (1.3-fold), highlighting the differing effect one effector gene can have on two different types of protein product (Josse et al. 2010). Stable co-expression of Calnexin and Calreticulin (via an inducible system resulting in a 2.9- and 2.8-fold increase of the effector genes respectively) saw a 1.9-fold increase in thrombopoietin production levels, showing the positives of overexpressing more than one gene at once (Chung et al. 2004).

Overexpression of disulphide bond forming PDIs resulted in a 37% increase in Mab production rate, supporting previous experiments performed in bacteria, yeast and insect cells, showing that the use of other cell lines as models is a viable practice (Borth et al. 2005; Davis et al. 2000; Nishimiya et al. 2013). In BHK cells, expression of the chaperone HSP70 increased blood factor VIII productivity by 50% and also increased cell viability (Ishaque et al. 2007).

Overexpression of specific SNARE proteins and their effector SM proteins saw an up-to 6-fold increase in SEAP and SAMY production levels in CHO cells (Peng et al. 2011; Peng and Fussenegger 2009b), whilst engineering of lipid transport to the Golgi via expression of an activated version of the ceramide transport protein (CERT) saw an up-to 45% increase in Mab and tPA production levels (Florin et al. 2009; Rahimpour et al. 2013).

Highlighting the role targets taken from outside of the CHO cell biosynthetic and secretory pathways can have, engineering of CHO cells with a permanently active oncogenic mutant of the tyrosine receptor kinase KIT (onco-KIT), shown to function in cell differentiation and cancer cell survival, increased CHO productivity of a stably-expressed GFP-$F_c$ fusion protein by up to 2-fold. Onco-KIT also increased cell proliferation and growth, promoting mitochondrial activity and cellular resistance to hypoxia through enhancing the UPR (Mahameed and Tirosh 2017). Likewise, expression of the chromatin remodeller and transcription factor YY1, known to have functions in embryonic development, cell differentiation, apoptosis, oncogenic transformation and B-cell maturation, increased antibody titre in various human and CHO cell lines by up to 6-fold (Tastanova et al. 2016).

However, there is also some contradiction within the literature as to the impact certain effector genes have upon CHO cell productivity. ATF4 overexpression has resulted in both a decrease (Mab) and increase (Antithrombin and Mabs) in CHO productivity (Haredy et al. 2011; Haredy et al. 2013; Ohya et al. 2008; Rajendra et al. 2015). Overexpression of BiP has also been shown to result in a decrease in specific secretion levels whilst its repression has increased secretion levels by up-to 3-fold. This suggests BiP has an important role in ER stress management but not necessarily protein secretion (Borth et al. 2005; Dorner and Kaufman 1994; Dorner et al. 1988; Morris et al. 1997).

It is possible that these inconsistencies may arise courtesy of experimental differences. Utilisation of different transfection methods of the effector or production genes (stable or transient); different recombinant products having distinct responses to specific effectors dependent upon their molecular requirements; or contrasts in process and testing between different experiments could all result in results that are difficult to compare. Whilst these examples show that any positive data taken from genetic engineering experiments may not be entirely global in its application, the opposite may be true and that, if the science behind a particular target is solid, it may serve to improve productivity for some model recombinant products.

The dose level of a certain effector gene can also have an effect upon its function and therefore its impact upon CHO cell productivity (Hansen et al. 2017). Exceeding the optimal levels of an ER-resident enzyme such as a PDI – which catalyses disulphide bond formation - can overwhelm other parts of the protein folding machinery (Davis et al. 2000). Overexpression of CHOP at a high level saw a decrease in Mab expression, whilst a low-level overexpression saw a slight increase in productivity (Rajendra et al. 2015). This could be explained by the excessive activation of CHOP leading to activation of pro-apoptotic UPR cascades (Gardner et al. 2013). Some explanation of this may come from co-expression of

CHOP with BiP and PDIs resulting in a 2-fold increase in Mab production levels. The protein folding capacity of PDIs may go some way to alleviating ER stress, suggesting that this dampens the CHOP-activated pro-apoptotic cascades whilst not affecting those involved in non-apoptotic ER stress recovery (Nishimiya et al. 2013). Furthermore, lower gene doses of XBP1s and ATF6 have been shown to increase Mab productivity (Johari et al. 2015), possibly again due to UPR activation levels that do enough to relieve the cell of ER stress, but not so much that apoptotic pathways are activated.

The aim of this work was to use genetic engineering of the CHO cell with a wide range of effector genes to enhance CHO productivity of a DTE model Mab. The effect of target genes upon transient and stable DTE-Mab production was tested to a) determine the effect of genes upon Mab production; b) allow comparison of how CHO cells react when using the two separate expression systems. A DTE-Mab was used as an expression model as it is more industrially relevant to try and enhance the production levels of products that are DTE, as ETE-Mab production levels are generally perceived to be satisfactory. Furthermore, expression of a DTE-Mab would place the CHO biosynthetic and secretory pathways under increased stress levels, allowing effector gene effect to be amplified. It is possible that ETE-Mab production has been optimised and as such engineering would have minimal effect upon CHO productivity.

## 4.2. Single gene engineering of CHO cell biosynthetic and secretory functions

An initial gene screen was performed to determine the impact effector genes and their dose levels have upon CHO cell productivity. Using a high throughput, long-term fed-batch transient expression system, MedI-Tran cells were transiently co-transfected with a vector expressing a DTE Mab (with both HC and LC expressed from the same plasmid) alongside another vector expressing the effector gene to be tested (listed in tables 3-6 and 3-7). A DTE Mab was used as a model system due to a larger drive across industry to enhance DTE expression levels. It is likely that a DTE will have more protein processing or secretory bottlenecks to solve to enhance its production. As such any overall improvement in protein processing or secretion will be amplified more greatly with a low-producing DTE compared to a high-producing ETE for which productivity is already close to maximal.

### 4.2.1 Gene design

Effector gene sequences derived from CHO genes were taken from CHOgenome.org, utilising the CHO-K1 2014 genome sequence, release 101 of the original 2011 genome sequence (Hammond et al. 2012; Lewis et al. 2013; Xu et al. 2011). For human-derived sequences, protein sequences were taken from

uniprot.org with the relevant gene sequence sourced from NCBI gene. Comparison of the selected protein and nucleotide sequences against other the genomes and proteomes of other mammalian species were made using BLAST (NCBI) to investigate homologue similarities between CHO and human-derived genes and proteins. This also ensured the sequence chosen was correct. As genes were to be expressed in CHO cells, all sequences were codon optimised so as to maximise translation levels.

#### *4.2.1.1. Mutation of certain effector genes*

To maximise the functionality of some proteins mutated versions were used. In CERT, phosphorylation of the serine 132 residue reduces its lipid transfer ability. To reduce phosphorylation with the aim of increasing CERT's lipid transfer ability, the serine 132 residue was replaced with an alanine residue (S132A), with the specific codon encoding residue 132 deciphered and swapped for an alanine-encoding codon (Florin et al. 2009). Likewise with the TFE3 transcription factor, phosphorylation of the serine 108 residue anchors TFE3 in the cytoplasm, rendering it unable to enter the nucleus and activate transcription. Again the serine was replaced with an alanine residue (S108A) so as to produce a constitutively active mutant which could enter the nucleus without any upstream activation required (Taniguchi et al. 2015).

The CREB3L2 transcription factor is made up of lumenal, trans-membrane and N-terminal cytosolic domains. Under normal conditions the entire protein is present across and on both sides of the ER membrane. Upon transportation to the Golgi membrane the protein is cleaved by an S2P integral membrane protease proximal to the cytosolic face of the lipid bilayer (Barbosa et al. 2013; Rawson 2013). This releases the transcriptionally active cytosolic domain, allowing it to enter the nucleus and induce transcription of secretory genes (Fox et al. 2010). For this reason only the active cytosolic domain was used for CHO cell engineering. A hydrophobicity plot (ExPASy ProtScale) highlighted the transmembrane domain as starting around residues 310/320 (Gasteiger et al. 2005; Kyte and Doolittle 1982). This region was searched for the target sequence of the S2P cleavage domain (TCLMV) that resides just within the Golgi lipid bilayer (Barbosa et al. 2013). This was found to start at amino acid residue 316 and as such the DNA sequence downstream of this point (encoding the transmembrane and lumenal domain) was removed to leave only the transcriptionally active N-terminal cystolic region.

#### *4.2.1.2. Effector gene sequences and plasmid design*

The effector gene sequences used are shown in Appendix 2. Short 5' prefix and 3' suffix sequences were added to the DTE Mab gene and each effector gene to ensure their compatibility with the MedImmune single gene expression platform and MGEV systems respectively. These adjustments comprised of a 5' sequence

14 bases in length immediately upstream of the ATG start codon. The 5' sequence contained a restriction enzyme cut site, 5' untranslated region (UTR) section and a Kozak sequence. The 3' 8 base pair downstream sequence was placed immediately after the TGA stop codon and contained a restriction enzyme cut site. All recombinant genes were under the control of a CMV promoter. All plasmids contained an OriP sequence to allow long-term transient expression (Daramola et al. 2014; Gahn and Schildkraut 1989; Hatton et al. 2010).

### 4.2.2. Single gene screen of effector genes in long-term transiently-expressing CHO cells

Effector genes were transiently transfected alongside a transiently expressed DTE-Mab. Effector genes were expressed at four separate levels so as to determine the dose response of CHO cells to each effector gene. The use of four levels provided better elucidation of dose-response trends than would have been provided by three dose levels. Cells were transfected using the Lonza Amaxa 96-well shuttle™ protocol optimised by CA, as described in section 2.5.2. The gene dose levels were defined as a percentage of total plasmid DNA transfected (800 ng), with 10.0, 5.0, 2.5 and 1.25% levels used. The gene levels and DNA loads by mass used are shown in table 4-1. The XBP1s (human), CypB, Largen and Ero1-Lα effector genes were only transfected at one level (5.0%) as they had all previously been screened.

Transfected levels of plasmid DNA were measured by mass and not plasmid copy number. Due to the variance in size of the genes expressed this meant that copy numbers were not the same across all effector genes. However this experiment was designed to compare dose levels of the same gene to determine the most effective gene dose levels for each gene separately. As such the variance in gene dose level across the different effector genes was not seen as problematic with comparisons being made to the Mab-only control and within cells transfected with the same plasmid. Post-transfection cells were seeded in a 96 deep-well plate and grown for 120 h, with a feed at 72 hours. Cell viability samples were taken at the start and end points to determine cell growth, with titre samples taken at the end point (section 2.1.2.2).

The gene screen was spread across four 96-well plates. Each plate carried seven genes at each of the four dose levels, one of the extra 5.0% genes from previous studies, and control (Mab-only) transfections. Each transfection (for each gene at each dose level) was repeated three times within each plate to provide biological replicates, with each plate repeated three times, producing three technical replicates. In the case of a well obviously not growing or becoming contaminated the results for each of these were removed from all data sets. This resulted in a maximum of nine data points being available for each effector gene (three wells

in each of three plates) and 36 data points for the control (three wells in each of in twelve plates). This holds true for all data sets – cell growth, overall titre and specific productivity. Data was normalised against the average of all control across all plates such that it could be displayed as a fold change of the control. A one-way ANOVA, utilising a Dunnett's multiple comparisons test (a many-to-one comparison) was used to compare the combined data from each transfection condition to that of the controls and calculate statistical significance. The overall results of the single gene screen in cells transiently expressing DTE-Mab are shown in figure 4-1.

**Table 4-1. Gene dose levels for single- and multi-gene engineering of CHO cells**

| DNA | % of total DNA load | % of Mab load | Actual mass (ng) |
|---|---|---|---|
| **Gene screens with transient Mab expression (single and multi-gene)** | | | |
| Total DNA load | *100* | - | *800* |
| Mab | 66.6 | 100 | **533** |
| Effector 10% | **10** | 15 | 80 |
| Effector 5% | **5** | 7.5 | 40 |
| Effector 2.5% | **2.5** | 3.75 | 20 |
| Effector 1.25% | **1.25** | 1.88 | 10 |
| **Gene screen with stable Mab expression** | | | |
| Total DNA load | *100* | *(Percentage of 533 ng transient Mab load)* | *800* |
| Effector 30% | **20** | 30 | 160 |
| Effector 17.5% | **11.7** | 17.5 | 93.6 |
| Effector 5% | **3.3** | 5 | 26.7 |
| | | | |

*The levels of gene dose used to engineer CHO cells is shown.*
*When cells were transiently expressing effector genes alongside transiently expressed Mab, the gene dose level is described as a percentage of the overall DNA load, with effector genes transfected at 10%, 5%, 2.5% and 1.25% of the total DNA load of 800 ng.*
*When cells were transiently expressing effector genes alongside a stably expressing Mab, the gene dose level again is described as a percentage of the overall DNA load of 800 ng, with effector genes transfected at 20%, 11.7% and 3.3% of total DNA load.*
*Total DNA load was made up of Mab DNA (if required and requisite effector DNA. The total volume (800 ng) was made up with empty backbone vector containing all the constituent parts of the gene-containing vectors, but not the gene itself.*

#### *4.2.2.1. Results of gene screen in transiently Mab-expressing cells*

The effects of ectopic gene expression upon cell growth, titre and specific productivity is assayed. The results are shown in figure 4-1.

##### *4.2.2.1.1. Cell Growth*

Across the gene screen ectopic gene expression does not appear to have had a large impact upon cell growth when compared to the average IVCD of the control across all 12 plates (50.1 RFU day $mL^{-1}$; figure 4-1 a). Some genes (1-DAD1, 15-CREB3L2) produce a slight increase in IVCD when compared to the control, whilst others (10-TFE3, 11-FKBP11, 26-HSP90B1) see a slight decrease in growth levels, but none of these are statistically significant. Furthermore, across the entire screen there does not appear to be an obvious dose response in terms of growth (with the exception of 13-Bet3 and 26-HSP90B1, suggesting these two genes may have a role in determining cell growth. However given the size of the gene screen a small number of genes with a correlation effects to gene dose is to be expected,) further highlighting the lack of effect of all the effector genes upon cell growth.

##### *4.2.2.1.2. Titre*

Titre was measure using the Valita™TITER high-sensitivity assay. Across the gene screen there are varying levels of effect upon overall DTE Mab titre compared to the level in control Mab-only transfected cells. which averaged 4.12 mg/L across the twelve plates. A small selection of genes appear to improve overall titre whilst many appear to have little effect, or actually instigate a decrease titre levels (figure 4-1 b). Expression of both XBP1s genes (27-XBP1s-CHO, 29-XBP1s-Human) increase titre levels significantly. The CHO isoform results in a 2.45-fold titre increase at a 10% gene dose level, with 1.5- and 1.4-fold increases seen at 5% and 2.5% doses respectively. The response of CHO cell productivity levels to XBP1s-CHO level highlights an obvious dose response with higher XBP1s levels resulting in higher titres.

That a large titre leap is seen between 5% and 10% gene doses suggests that there may still be some headroom in to which XBP1s-CHO can extend in terms of enhancing productivity. At a 5% dose level human XBP1s saw a 2-fold increase in titre compared to the control, which is larger than the 1.5-fold increase seen at the same gene dose levels with the CHO isoform. As both genes are of similar sizes (CHO – 1113 bp; Human – 1131 bp; Appendix 2) it can be considered that the gene copy number of both isoforms is very similar and as such a direct comparison between the two results can be intimated at.

There is 84% sequence homology between the two XBP1s isoforms (from BLAST, data not shown). With XBP1s interacting with promoter regions and TFREs within the CHO genome it was hypothesised that expression of a CHO-derived

isoform would provide enhanced interaction with promoter elements and therefore increased transcriptional activation of biosynthetic and transcriptional enhancing genes, increasing productivity. This would suggest that enhancing effector gene activity is not necessarily achieved by the origin of the gene sequence and highlighting the similarity of transcription factor activity across species.

The other genes to significantly increase DTE-Mab titre were 4-PRDX4 and 16-CRELD2, which saw a 1.5- and 1.4-fold increase in titre respectively at the highest 10% gene dose. One other gene – 15-CREB3L2 at 10% - produces an equal to or greater than 1.4-fold increase in titre, though this is not statistically significant. That XBP1, CRELD2 and CREB3L2 expression all result in titre increase shows that transcription factor engineering can play an important role in enhancing transient Mab production. That PRDX4 expression significantly increases cell productivity suggests that better control of oxidative stress aids Mab productivity.

That the ER chaperones and PTM producing enzymes 1-Dad1 and 3-HSPA1a also produce a slight increase in productivity (1.2-fold, though neither statistically significant) shows that enhancement of ER processing may enhance productivity, though 20-Tor1a and 24-Calreticulin both instigate a slight insignificant decrease. Both of these genes have similar ER chaperone functions. With the exception of 5-PREB, 12-P115, 13-Bet3 and 22-MCFD2, many genes involved in trafficking and secretion result in an insignificant but noticeable change in titre. In many of these cases there also does not appear to be much of a dose-response pattern.

#### *4.2.2.1.3. Specific Productivity (Qp)*

The average Qp value for the control across all 12 plates was 0.088 pg cell$^{-1}$ day$^{-1}$ (figure 4-1 c). Specific productivity data for the single gene screen highlights that any increase in titre levels is as a direct result of an increase in Qp, due to cell growth being generally unaffected by effector gene expression (figure 4-1 a). Effectively there is little difference between the graphs showing Qp and titre.

The single gene screen in a transiently-expressing cell line show that there is scope for enhancing CHO transient DTE-Mab productivity levels. This is particularly prevalent in the case of transcription factors, especially XBP1s but also the plasma-cell highlighted CREB3L2 and CRELD2. However some transcription factors (TFE3, YY1, Largen) had little or no positive effect upon productivity and titre levels. Single-gene expression of molecular components of the biosynthetic and secretory pathways had a less marked effect than that of transcription factors. Whilst some ER-based genes (Dad1, PRDX4) did have a slight impact upon productivity, many had none. With many of these secretory

components being selected to work in combination with other effector genes it was not hypothesised that they would have a large positive impact upon productivity when expressed on their own (see section 4.3.4).

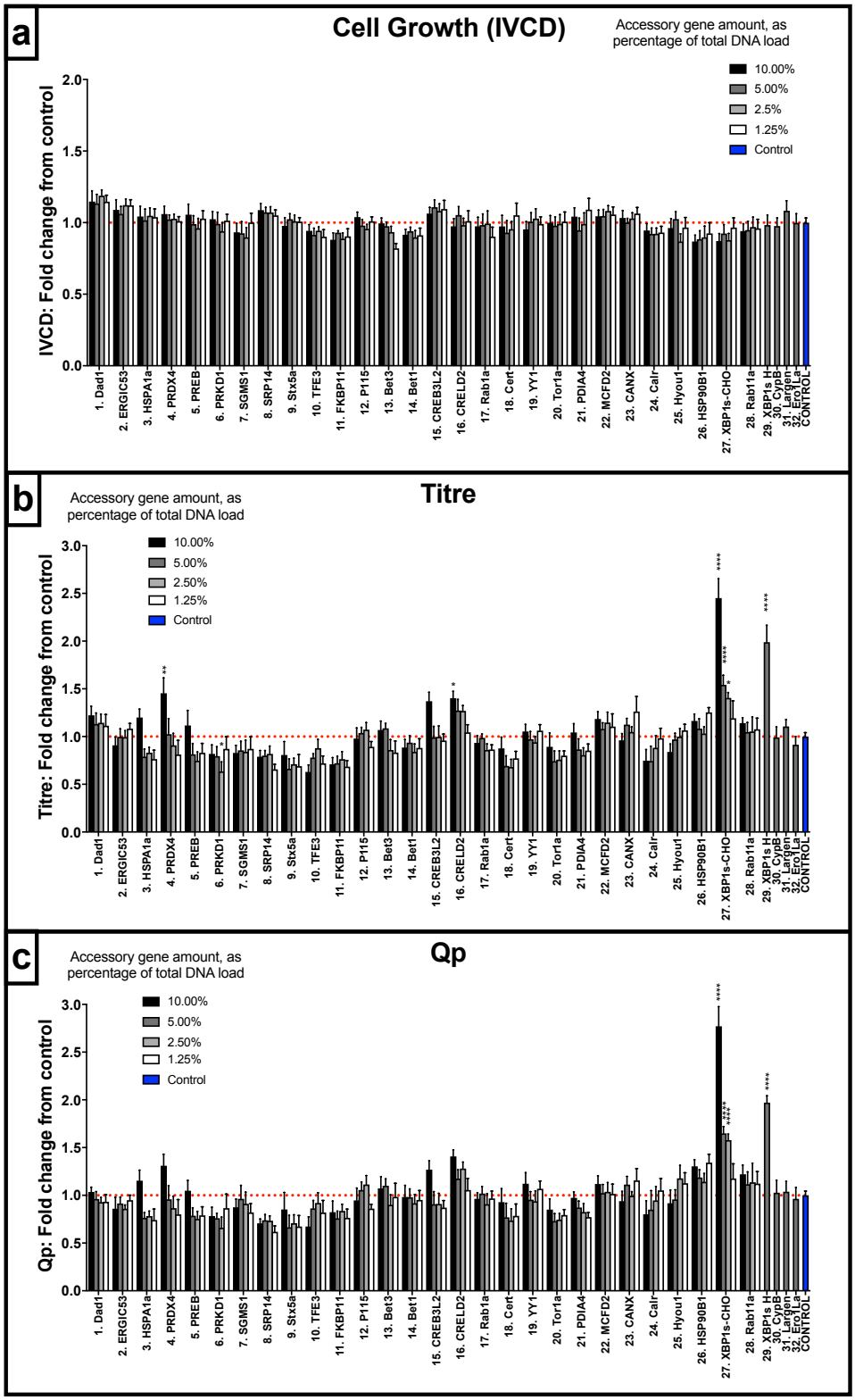


**Figure 4-1. Gene screen of transiently-expressed single-effector genes in a CHO cell transiently expressing a DTE-Mab.**

*.MedI-TRAN cells were transiently transfected with a DTE-Mab and a range of effector genes, each at 4 different dose levels. After 120 h incubation [a] growth, [b] titre and [c] specific productivity were assayed and displayed as fold-change of level of Mab-only expressing cells. Columns show mean of three technical replicates of three biological replicates (n = 9; n = 36 for control cells). Error bars show SEM. Statistical significance calculated by Dunnett's test is shown by asterisks. DNA loads explained in Table 4-1.*

### 4.2.3. Single gene screen of effector genes in stably-expressing CHO cells

*Note that this section differs to the previous. Here, effector genes are still transiently expressed, but are done so in a pool of cells stably expressing the same DTE Mab used previously. As a pool was used the population was heterogeneous.*

Genetic engineering of CHO cells transiently expressing a product provides useful information as to the function of each gene within the cell as well as providing information for any long-term transient production systems (Daramola et al. 2014). Ultimately however genetic engineering strategies would mainly be used to try and enhance productivity in stably expressing cell lines. As such carrying out a gene screen in cells stably expressing a recombinant product will provide more tangible information as to a genes' potential in enhancing productivity in a more industrially relevant scenario.

The increased levels of Mab being produced in a stably-expressing cell is much greater than when expressed transiently (section 5.4.2; Daramola et al. 2014). As such the CHO cell's secretory and biosynthetic pathway is likely to be under a greater burden when stably expressing that under transient gene expression. Whilst the stable selection process results in a pool in which all cells can both produce product and grow well, use of a pool instead of a clonally-derived cell line allows for some heterogeneity across the entire cell pool population. As such some bottlenecks will be present across the population and not all cells will already be maximally producing recombinant product, as would more likely be the case in a clonally derived population. As the pools have not been pre-selected as high producers there is also more head-space into which they can increase performance. As such any impact of expressed effector genes will be amplified and thus easier to spot.

Using the same high throughput transfection and growth platform for the transient production effector gene screen, another gene screen was performed using a pool of cells stably expressing the same DTE-Mab used in the transient screen (see section 5.2.4.4 for pool production; sections 4.2.3, 2.5.2 and 2.1.2.2 for gene screen platform). Growth and productivity profiles for this DTE-expressing stable pool (henceforth referred to as MedI-CHO-DTE) were performed as in section 5.2.4.4 and are shown in figure 4-2.

Titre data was produced by Valita™TITER with three technical replicates. Genes were tested at three levels to better streamline the gene screening process, reducing it to three 96-well plates. The gene dose levels were altered in an attempt to better accommodate the increased productivity of the stable cell line. This was made possible as the majority of the 800 ng DNA load per transfection was no-longer made up of DTE-encoding plasmid. The upper gene dose level was

set at 20% of the total DNA load with the remainder made up of non-coding backbone DNA (160 ng effector gene). The medium level was set at 11.7% (93.6 ng) and the lower level at 3.3% (26.4 ng; see table 4-1).

The majority of the genes used were the same as those in section 4.3.2. The exceptions were Tor1a, XBP1-human, Dad1 and Largen. Tor1a and Largen were removed due to the lack of impact they had upon CHO productivity. Human XBP1s was removed and only CHO carried forward. This was despite the human isoform outperforming the CHO isoform at the same level (5%) in the transient form, a decision made to maintain CHO isoforms of all transcription factors used. Dad1 was removed as, although it produced a slight increase in titre in transiently expressing CHO cells this was due to a slight increase in growth and not Qp.

Due to industrial constraints the screen was set to try and improve titre as a function of enhanced Qp. This is because a as a large increase in cell growth can cause problems in both upstream (increased chance of shear forces in bioreactors, increasing cell breakage and possibly increasing HCP levels) as well as downstream processing (more cell debris to purify product from). Furthermore use of Dad1 in another screen did not enhance titre or productivity (data not shown).

Human isoforms of the transcription factor ATF6αc and protein disulphide isomerase PDI were added to the gene screen, both having previously been shown to improve CHO productivity (Borth et al. 2005; Johari et al. 2015; Pybus et al. 2014). A human ATF6c isoform was used, despite it being a transcription factor, as this was the sequence used previously and so was retained to allow comparison to previous studies.

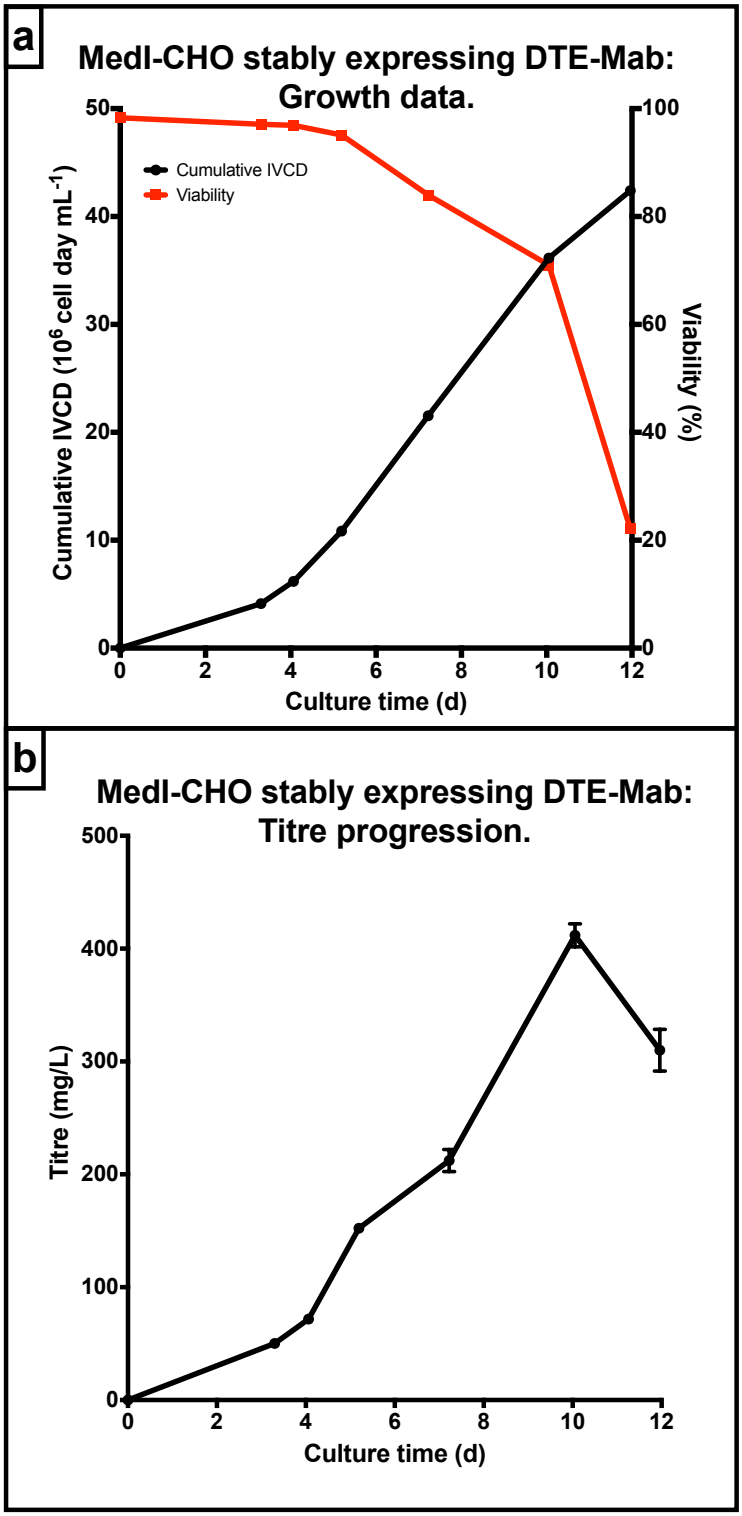


**Figure 4-2. Growth and productivity profiles of MedI-CHO-DTE stable pool.**

*MedI-CHO were stably transfected with DTE-Mab to produce a heterogeneous stable pool, referred to henceforth as MedI-CHO-DTE. A 12-day growth and productivity profile was performed to assay the performance of MedI-CHO-DTE. [a] Growth data shows that cells did not reach the same density as parental cells but that defined lag, exponential, stationary and death phases were present. Cell viability started to decrease around days 5-6 of culture. [b] Cell titre continued to build across culture time, reaching a high of approximately 400 mg/L on day 10. Growth data shows one point from a single ViCell read. Titre data shows three technical replicates. Error bars show SEM.*

#### *4.2.3.1. Mechanistic normalisation of data from stably-expressing gene screen*

In the transient screen all cells were transfected with DNA encoding recombinant genes (both effector and Mab). Despite the levels of gene-encoding (and therefore transcriptionally active) vectors varying slightly due to differing dose levels this was minimal due to the Mab-encoding plasmid making up the majority of the vector load. However in the gene screen of stably-expressing cells, despite control cells being transfected with empty vector to control for the electroporation process and DNA cytotoxicity, with empty vector DNA not containing a gene it is much less transcriptionally active than encoding vectors.

As such the transcriptomic, translational and protein processing burden put upon cells transfected with effector gene will be greater than that put upon the control cells. To control for this increased transcriptional and metabolic burden in stably expressing cells transiently expressing effector genes, and to better make a comparison to non-effector control cells, titre and growth data underwent a mechanistic normalisation. This was performed by dividing each data point by the average of all data points (from all genes) transfected with the same level of DNA (i.e. titre from well 1 of cells transfected with 3.3% ERGIC53, replicate 1, are divided by the average of all cells transfected with a 3.3% dose of the effector gene).

This normalisation has the effect of bringing all data points, either above or below that of the control, slightly closer to the control. As such it can not be considered to artificially amplify any productivity or growth data but instead has more of a dampening effect upon the results. Data points for the controls did not undergo any normalisation as they were not transfected with transcriptionally active DNA. To calculate mechanistically normalised Qp data the original titre data was used alongside the mechanistically normalised IVCD data such that only one normalisation step was taken.

#### *4.2.3.2. Results of gene screen in CHO cells stably expressing a DTE Mab*

The results from the gene screen carried out in the MedI-CHO-DTE cell pool are shown in Figure 4-3. Data was processed as in the transient gene screen with regards to biological and technical replicates, with the exception of data undergoing a mechanistic normalisation to control for the increased transcriptional burden of cells expressing accessory genes. A Dunnett's test was used to compare the effect of accessory gene expression and dose level to that of control cells and determine statistical significance.

##### *4.2.3.2.1. Cell Growth*

As with the transiently-expressing cell line, many of the genes have no significant impact upon cell growth (figure 4-3 a). The average IVCD of control wells was 36.0 RFU day $mL^{-1}$. The only statistically significant result is that of XBP1s-CHO at a high gene dose, which appears to decrease growth. CERT at a high gene dose

also appears to reduce growth, albeit not significantly. Furthermore, there does not appear to be any obvious dose-response to any genes, though with only three gene levels tested this would be more difficult to fully elucidate than in the transient four-level screen.

#### *4.2.3.2.2. Titre*

A small selection of genes increased titre levels by a statistically significant amount (figure 4-3 b). The average titre across all control wells was 80.4 mg/L, a 20-fold increase compared to the transiently-expressed DTE screen (section 4.2.2). At the lowest gene dose level both Ero1-Lα and PDI increase productivity by 1.2- and 1.3-fold respectively, with the results being statistically significant. Both genes show a dose response with lower gene doses having a larger effect upon the overall titre than higher doses. At the highest level of gene dose PRDX4 displays a 1.2-fold titer increase, though PDIA4 does not have an impact. All of these genes are involved in disulphide bond formation. That expression of three of them increases DTE productivity suggests that for this DTE-Mab disulphide bond formation could be a potential production bottleneck.

Expression of the ER stress response regulator XBP1s-CHO had no obvious effect upon the productivity of a stably-expressing cell line. This is in contradiction to the transient gene screen where it saw a 2.5-fold increase in titre (figure 4-1 b) which reinforced previous literature studying the effect of XBP1s on a transiently-expressing reporter gene. Other transcription factors in the screen (CREB3L2, CRELD2, YY1 and TFE3) had no effect upon cell productivity levels, however a high dose of ATF6c did reduce overall titre. Of the genes directly involved in the secretory pathway, expression of Bet3 (vesicle tethering) and PREB (ERES formation) at high levels both resulted in 1.2-fold increases in productivity although neither were statistically significant. Whilst the former does not show a dose response the later does, with titre increasing as gene dose increases. With the exception of Bet3 genes involved in vesicle tethering and fusion (Stx5a, P115, Bet1, Rab1a, Rab11a) did not alter productivity.

SRP14 (protein import to ER) expression did not alter productivity. Neither did PRKD1 and SGMS1 though expression of CERT did reduce productivity despite all three working in Golgi lipid metabolism. Genes encoding protein chaperones (HSPA1a, FKBP11, Canx, Calr, Hyou1, HSP90B1, CypB) appear to have no significant impact upon titre levels, with only FKBP11 showing a different trend to that seen in transiently-expressing cells. The two ERGIC genes (ERGIC53, MCFD2) have differing effects upon titre levels. The former reduces productivity whilst the later has no effect. Whilst many genes showed little impact upon overall titre levels, the majority did see a slight increase, albeit not statistically significant.

#### *4.2.3.2.3. Specific productivity (Qp)*

When taking both growth and titre into account the majority of genes appear to improve cellular productivity (figure 4-3 c). However, variation in the data suggest it is less clear than the direct titre and growth measurement data sets. Average Qp across all control wells was 2.2 pg $cell^{-1}$ $day^{-1}$. In general expression of transcription factors derived from plasma cells appear to have no significant impact upon cell productivity (CREB3L2, CRELD2), although XBP1s-CHO and ATF6c at low dose levels and YY1 at a high dose level result in a 1.3-1.4-fold Qp increase, albeit this is not statistically significant.

The effect of genes involved in vesicle transport varies somewhat. Bet3 sees a 1.4-fold Qp increase at a high dose level, Rab11a a 1.3-fold increase at a low level and PREB a 1.3-fold increase at a medium level (though these changes are not significant), whilst other vesicle transport effectors have little impact on Qp levels. Expression of chaperones has varying effects on Qp. Calr and Canx have no effect, FKBP11, HSPA1a and Hyou1 all see an insignificant 1.3-fold Qp increase, whilst HSP90B1 sees a significant 1.4-fold increase. All these changes occur at differing dose levels. Genes linked to disulphide bond formation (ERO1Lα, PDI, PRDX4) nearly all see at least a 1.3-fold increase in Qp (with the exception of PDIA4) with lower levels of gene generally seeing the largest increase. PRKD1 is the only other gene (other than HSP90B1) to statistically increase Qp levels with a close-to 1.5-fold Qp increase seen at the medium dose level. However again there is no dose-response seen.

The difference in data between the transiently- and stably-expressing gene screens shows that in general the former can not inform the later. This is particularly prevalent when looking at the gene families. Transcription factors that increase transient productivity (XBP1s-CHO, CREB3L2, CRELD2) have little impact in stable cells, whilst chaperones and disulphide bond forming genes appear to have an enhanced effect upon productivity in stably-expressing cells. This is discussed further in section 4.5.

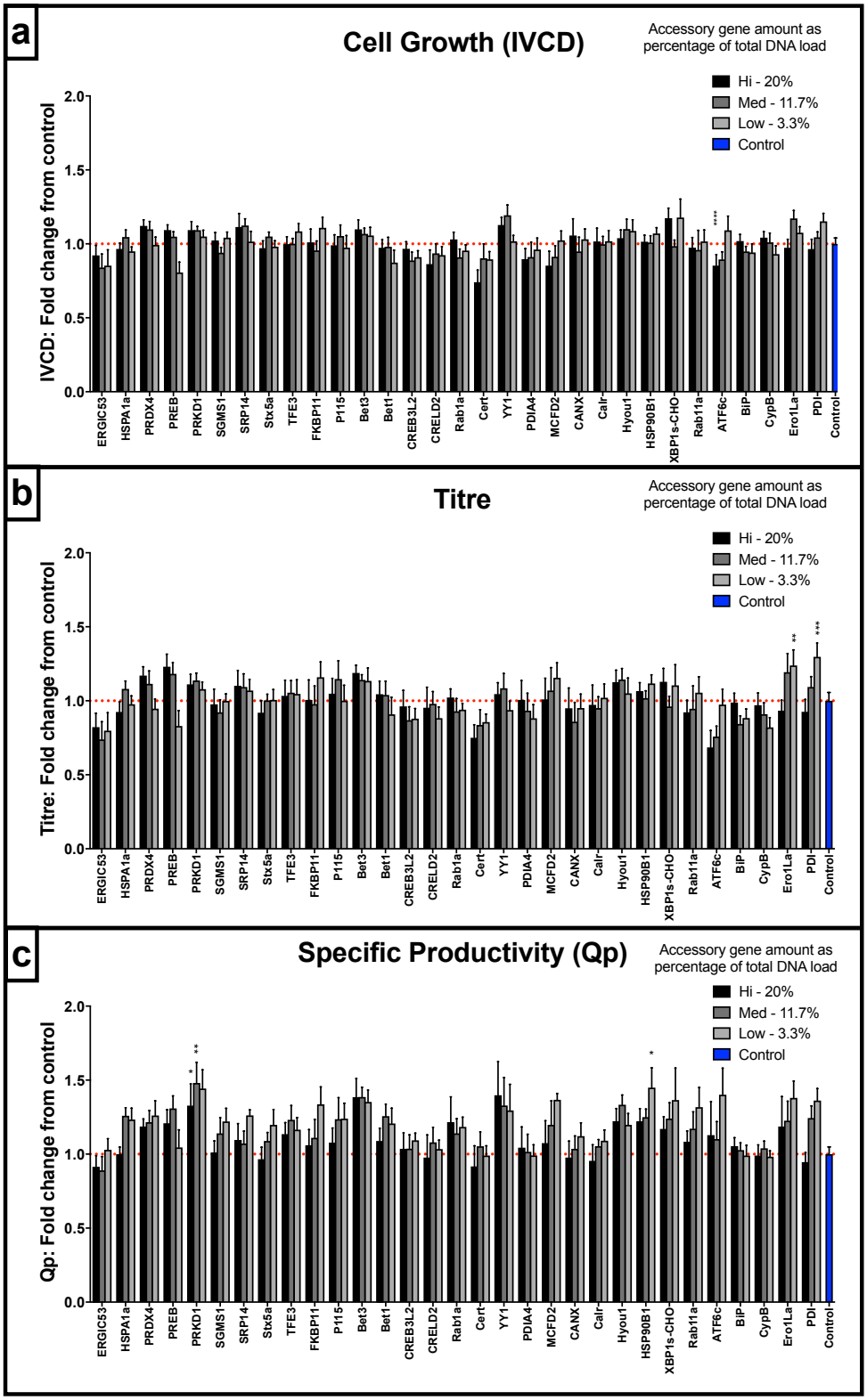


**Figure 4-3. Gene screen of transiently expressed effector genes in a CHO cell pool stably expressing a DTE-Mab.**

*.MedI-CHO-DTE cells (stably expressing DTE-Mab)were transiently transfected with a range of effector genes, each at 3 different dose levels. After 120 h incubation [a] growth, [b] titre and [c] specific productivity were assayed and displayed as a fold change of level of Mab-only expressing cells. Columns show mean of three technical and 3 biological replicates (n = 9; n = 27 for control cells). Error bars show SEM. Asterisks show statistical significance calculated by Dunnett's test. DNA loads explained in Table 4-1.*

Due to the large number and small volume of culture produced in the gene screen it was not possible to perform in-depth screens on the activity of transfected genes across all cultures. For this reason it was not possible to determine exactly the level at which the transfected genes were transcriptionally active using, for example, Chip-Seq or RNA-seq, or whether transfection resulted in an increase in protein levels with the cell by proteomic analysis. Testing of the gene screen platform with GFP showed that recombinant expression of transfected DNA was replicable across wells in the same plate, and across different plates (data not shown; Arnall et al. Manuscript in preparation). Furthermore, as all wells were treated the same, and with very similar levels of plasmid DNA, it can be assumed that all recombinant genes were equally active. Tightness of data across in-plate replicates and across replicate plates further adds to this.

## 4.3. Multigene engineering of CHO secretory functions

Single gene engineering of specific sections of the biosynthetic and secretory pathways can lead to stoichiometric imbalances. As such when increasing cellular levels of these genes there are three possible outcomes:

1) An increase in productivity occurs as the gene expressed was a limiting factor and rate-limiting step in protein processing/secretion.
2) There is no change in cell productivity. The expressed gene has no negative effects upon growth/productivity, but its function is dependent upon specific interactions with other molecules that have not been up-regulated. As such the engineered gene does not have the desired effect as the excess levels can not function.
3) Productivity decreases. A specific stoichiometric ratio of molecules is required to allow correct pathway functionality and increased levels of one molecule alters this balance to such an extent that the targeted pathway is blocked. This could be due to aggregation or oligomerisation of the excess protein inhibiting organelle/cell function, or multiple molecules binding one partner molecule, impinging upon its function.

Engineering of transcription factors was not expected to initiate these effects as they enhance the expression levels of functional gene groups with specific gene levels controlled by the cell, effecting a more global transcriptional control of an entire pathway (He et al. 2010). However, engineering of a single molecule involved within a pathway could result in these effects as, as far as the cell is concerned, no thought has been made to the pathway as a whole.

Whilst a direct single-gene engineering strategy can increase productivity – e.g. sole expression of CERT (Florin et al. 2009; Rahimpour et al. 2013) – and screening of single genes is useful to determine impact upon growth or potential

off-target effects, several previous CHO genetic engineering studies have shown that synergistic combinations of two or more genes can systematically increase CHO productivity. This is shown by engineering of the translocon and SRP complex, with expression of one component increasing productivity, but expression of multiple SRP subunits further increasing this effect (Le Fourn et al. 2014). It is also shown with engineering of SM family SNARE regulators, with Sly 1 and Munc18c both individually increasing CHO secretion levels, with levels increasing further upon their co-expression (Peng and Fussenegger 2009b). However, this is not always the case, with co-expression of interacting SNAREs resulting in no further productivity enhancement than that seen with singular SNARE expression (Peng et al. 2011).

This would suggest that in some molecular interactions only one component is rate limiting, in the case of SNARE interactions this being the presence of one specific SNARE and not necessarily the levels of all SNAREs in a bundle, with the rate at which SNARE proteins come together (controlled in part by tethers and SM proteins) also being important. That co-expression of BiP and PDI resulted in a decrease in Qp shows that expression of a protein that decreases productivity (BiP) can wipe-out the productivity improvements another conveys upon the cell (Borth et al. 2005).

### 4.3.1. Determining multigene combinations for CHO cell engineering

Both data-driven (from the single gene screen) and knowledge-driven (from secretory pathway mapping and previous literature) approaches were used to produce up-to three-gene combinations with which to transiently engineer the CHO biosynthetic and secretory pathways to increase CHO productivity of a DTE Mab. Despite separate vectors being used for each gene a maximum of three genes were selected for each combination to better mimic the three-gene MGEV platform with which the gene library will eventually be used. This limit also avoided excessive transcriptional burden upon the cell and the dilution of gene expression levels due to the presence of increasing numbers of genes. For both data- and knowledge-driven combinations the gene levels used were selected from the initial single-gene screen, with the level that had the most positive effect upon titre and/or Qp being chosen. In future studies gene expression levels from both single or multi-gene vectors can be tightly controlled by synthetic promoters based upon these single gene methods (Brown and James 2016).

Some genes used in the single-gene screen were initially chosen with the aim to work in combination targeting a specific pathway. As such they were not necessarily expected to work independently of each other in the single gene screen. Five gene combinations were selected and tested using a design of experiment (DoE) approach with each gene at two dose levels to ascertain their

functional effect both individually and in dual and triple combinations. A summary of each combination is shown in table 4-2.

#### *4.3.1.1. Lipid transport pathway multigene engineering.*

**PRKD1, SGMS1** and **CERT** will be co-expressed to try and enhance lipid transport to the Golgi and TGN and better catalyse the conversion of CERT to DAG (by SGMS1). The hypothesised outcome of this is an improvement and increase in vesicle fission at the TGN (initiated by PRKD1) with a view to increasing TGN capacity and budding events though the increase of PRKD1 and DAG levels at the TGN (Baron and Malhotra 2002; Fugmann et al. 2007; Hausser et al. 2005; Huitema et al. 2004; Liljedahl et al. 2001). Overexpression of an active mutant of CERT has been shown to improve CHO productivity, though this was not seen in our single-gene screen with expression of CERT, SGMS1 and PRKD1 all resulting in small titre and Qp reductions (Fig. 4-1; Florin et al. 2009; Rahimpour et al. 2013). This gene combination is intended to better balance the increased levels of ceramide transport brought about by CERT expression. Expression of SGMS1 should catalyse the conversion of ceramide into DAG (with sphingomyelin as a potentially advantageous by-product), halting the retrograde transport of ceramide back to the ER. The increased levels of PRKD1 will interact with the increased DAG levels in the TGN membrane to initiate vesicle budding and fission.

#### *4.3.1.2 Multigene engineering of the ERGIC.*

**ERGIC53** and **MCFD2** form a cargo receptor complex involved in ER-Golgi transport (section 1.5.1.2; Appenzeller-Herzog and Hauri 2006). The two are present in a 1:1 stoichiometry with mutations and miss-targeting of the components shown to inhibit the secretion of blood factors (Appenzeller et al. 1999; Baines and Zhang 2007; Nyfeler et al. 2006; Vollenweider et al. 1998; Zhang et al. 2005). Expression of both in the transient gene screen showed little impact upon productivity (figure 4-1) – expected due to their close synchronous functionality dependent upon a balanced stoichiometry – but in stable cells ERGIC53 did reduce productivity (figure 4-3). Overexpression of the two genes at the same level should enhance cargo transport between the ER and Golgi.

#### *4.3.1.3. Multigene engineering of Vesicle fusion: Group 1.*

**Bet3** is a vesicle-localised subunit of the TRAPPI tether complex. As well as targeting the TRAPPI complex to COPII-vesicles, thus aiding COPII vesicle docking at the Golgi, Bet3 acts as a GEF for vesicle-localised **Rab1** which in turn recruits the coiled-coil tether **P115** (Cai et al. 2007b; Sztul and Lupashin 2009). Over-expression of these three proteins together is expected to enhance targeting and tethering of COPII-vesicles at the *cis*-Golgi, thus increasing flux through the early secretory pathway.

#### *4.3.1.4. Multigene engineering of vesicle fusion: Group 2.*

**Bet1** is a v-SNARE that, alongside other SNAREs, interacts with the t-SNARE **Stx5** during vesicle fusion at the ERGIC and Golgi. The tether **P115** interacts with both of these SNAREs prior to vesicle docking, initiating the preliminary steps of vesicle fusion (Hong and Lev 2014; Sztul and Lupashin 2006; Sztul and Lupashin 2009). As with previous studies co-expression of Stx5 and Bet1 is hoped to increase vesicle fusion events at the Golgi (Peng et al. 2011), whilst co-expression of P115 is hoped to enhance vesicle tethering to further improve the rate of SNARE assembly and this vesicle docking (Peng and Fussenegger 2009b).

#### *4.3.1.5. Combining singularly effective transcription factors.*

**XBP1, CREB3L2** and **CRELD2** were the most effective targets of the single gene screen in transient cells. Expression of the three transcription factors together will be carried out to ascertain whether the trio have a compound effect upon cell productivity. The expression levels of all three are increased in plasma cells (section 3.2; table 3-1), with XBP1 and CREB3L2 both linked to directly enhancing secretion in many cell types through the initiation of secretory gene transcription (Fox et al. 2010; Shaffer et al. 2004). XBP1 has been expressed many times in CHO cells including alongside other effectors, in these cases seeing a further increase when compared to XBP1-only expression (Cain et al. 2013; Peng and Fussenegger 2009b).

With the exception of MCFD2, Rab1, CRELD2 and CREB3L2, all the above genes have been linked to increased productivity capacity in CHO cells (Harreither et al. 2015). All the genes selected also see an increase in transcript abundance in plasma cells. Gene combinations are summarised in table 4-2. The gene levels chosen were selected based upon the single-gene transient screen (figure 4-1) and are shown and reasons for selection described in table 4-1.

**Table 4-2. Gene combinations for multigene engineering of CHO cells.**

| Combination target | Genes expressed | Aim |
|---|---|---|
| **1. Lipid transport pathway** | CERT S132A<br>SGMS1<br>PRKD1 | Enhance transport of ceramide from ER to Golgi/TGN; Increased conversion of ceramide to DAG and sphingomyelin at TGN; Increased vesicle fission events at TGN. |
| **2. ERGIC** | ERGIC53<br>MCFD2 | Increase cargo reception and transport between ER and Golgi |
| **3. Vesicle fusion group 1** | Bet3<br>Rab1<br>P115 | Increasing levels of vesicle-localised tethers and tether-activators to enhance tethering events at ERGIC and *cis*-Golgi. |
| **4. Vesicle fusion group 2** | Stx5a<br>Bet1<br>P115 | Increasing levels of a tether that interacts with a v-SNARE and t-SNARE, to try and improve vesicle fusion at the *cis*-Golgi. |
| **5. Transcription Factors** | XBP1s-CHO<br>CREB3L2<br>CRELD2 | Increasing levels of three transcription factors shown to increase CHO productivity in single-gene screen and increased in plasma cells. See if there is a compounding of the effect of all three transcription factors. |

*The gene combinations for multigene engineering of the biosynthetic and secretory pathways are shown, along with rational for the genes and combinations selected.*

### 4.3.2. Multigene DoE of secretory gene combinations

MedI-TRAN cells were transiently transfected with DTE-Mab alongside all genes from the selected gene combinations. Each accessory gene was encoded by a separate vector, with gene dose levels shown in table 4-3. The transfection and growth was carried out as with the single gene screen, with growth and titre samples taken after 120 h shaken culture (sections 2.5.2, 2.1.2.2 and 4.3.2).

Five DoE two-level factorials were prepared to test all possible gene combinations within the five gene groups and to determine the effect each gene had upon cell growth, overall titre and specific productivity (section 2.9.5). This allowed the effect of each gene upon cell growth, overall titre and Qp to be analysed, both singularly and in combination. For each three-gene selection this resulted in three single-gene combinations (a, b, c), three two-gene combinations (ab, ac, bc) and one three-gene combination (abc). This set-up allowed all combinations and controls to be run on a single 96 well plate.

Each combination was repeated in triplicate within the plate to provide three biological replicates, with each plate containing six DTE-Mab-only control wells. The plate was repeated twice to provide two technical replicates (A third plate was prepared to provide a technical triplicate but was unable to be analysed due to contamination). Any contaminated wells or wells which did not grow were removed prior to data analysis. This resulted in a maximum of six data points for each gene combination and twelve data points for controls. A Dunnett's multiple comparison test was performed on all results data so as to compare them directly to the controls and determine any statistically significant results. Design-Expert® was also used in analysis to determine the most prevalent effectors within each group.

Results from the multigene engineering of lipid transport to the Golgi and TGN-PM vesicle transport are shown in figure 4-4 a. As with the initial transient single gene screen (figure 4-1) each of the single genes (PRKD1 at 10% of total DNA load, SGMS1 at 5% and CERT at 10%) have little impact upon cell growth and productivity, with a slight reduction in titre and/or productivity seen from all three, though none of these are statistically significant. All gene combinations have no obvious effect upon CHO growth levels with only small decreases seen in combinations 1.5, 1.7 and 1.8. Combinations 1.6, 1.7 and 1.8 result in slight decreases in overall titre and productivity. Combination 1.4 has little effect upon CHO productivity. However, none of these results are statistically significant.
Multigene engineering of the ERGIC has no obvious effect upon cell growth but does appear to reduce titre and Qp compared to the control (figure 4-4 b, combination 2.3). However, the single-gene data shows that MCFD2 at 10% decreases growth and productivity (combination 2.2), in contradiction to the initial single-gene screen where neither MCFD2 of ERGIC53 had a major impact upon titre and Qp (figure 4-1 b, c).

**Table 4-3. Gene dose levels for multigene combinations.**

| Gene | Optimum level (as percentage of total 800 ng DNA load) | Plasmid amount used (mass) | Reason (from single gene experiment) |
|---|---|---|---|
| **PRKD1** | 10% | 80 | Least negative effect on titre and growth. |
| **SGMS1** | 5% | 40 | Little impact on growth or titre at all levels, best Qp at 5% |
| **CERT S1323A** | 10% | 80 | Least effect on titre, no effect on IVCD, small decrease in Qp. |
| **ERGIC53** | 10% | 80 | Works 1:1 with MCFD2. Minimal effect upon titre/growth between different levels |
| **MCFD2** | 10% | 80 | No large effect on titre or IVCD, therefore none on Qp, but best of all gene dose levels. |
| **Bet3** | 5% | 40 | Slight titre increase, similar impact on growth across all levels; largest Qp increase. |
| **Rab1a** | 10% | 80 | No impact on titre, IVCD or Qp at all levels. Used highest dose for maximum effect. |
| **P115** | 2.5% | 20 | Increase in titre, growth decrease similar across all levels; best Qp increase |
| **Stx5a** | 10% | 80 | Least negative impact upon titre; impact upon growth similar across all levels. |
| **Bet1** | 5% | 40 | Slight titre increase; no difference in growth across all levels. |
| **CRELD2** | 5% | 40 | Increase in titre, no effect on IVCD. Qp increase. |
| **CREB3L2** | 10% | 80 | Increase in titre, minimal decrease in IVCD. Qp increase. |
| **XBP1s-CHO** | 10% | 80 | Large increase in titre, small decrease in growth, overall increase in titre |

*The gene dose levels for the multigene engineering screen are shown as both a percentage of total DNA load (800 ng) and overall DNA mass. Reasons for the selection of the DNA load are briefly described.*

As with other gene combinations, engineering of CHO cells with the vesicle binding genes Bet3 (at 5% of total DNA dose), Rab1a (10%), P115 (2.5%), Stx5a (10%) and Bet1 (5%) has no impact upon cell growth, either singularly or in combination (figure 4-5 a, b). Expression of these genes singularly appears to decrease overall titre and Qp (combinations 3.1, 3.2, 3.4 and 4.1) with the exception of Bet 1 that has no major impact (combination 4.3).

Compared to the initial single gene screen this is as expected for Stx5a and Bet1, but not for P115, Rab1a or Bet3. Most of the gene combinations appear to reduce overall titre and Qp, with combinations 3.3, 3.5, 3.6, 3.7 and 4.2 all seeing a reduction by approximately half compared to the control. Combinations 4.4, 4.5 and 4.6 still result in decreases in titre and productivity compared to the control, albeit the reductions are smaller than those seen in the previous combinations. However, none of the combination results are statistically significant. This is due in part to differences in titre between technical replicates and will be discussed in section 4.5.

Results for engineering of CHO cells with the transcription factors CRELD2, CREB3L2 and XBP1s-CHO are shown in figure 4-5 c. Single gene engineering does not match the results seen in the initial single gene screen (figure 4-1). Whilst all three single genes have little effect upon cell growth, titre and Qp were previously shown to increase upon singular expression of all three transcription factors. However here CRELD2 has little impact, CREB3L2 sees a small decrease and XBP1 only a small increase, none of which are statistically significant (combinations 5.1, 5.2 and 5.4).

When expressed in varying combinations, an increase in titre and productivity is seen when XBP1s is co-expressed alongside the other transcription factors. Combining CREB3L2 and CRELD2 results in a slight decrease on cell growth, productivity and titre (5.3). Co-expression of CRELD2 and XBP1 (5.5), CREB3L2 and XBP1 (5.6) and all three genes together (5.7) all result in a slight decrease in growth but an overall increase in titre and Qp. However these results are only statistically significant in the cases of CREB3L2 and XBP1 (titre and Qp) and all three genes (Qp and IVCD).

To determine the genes and gene combinations within the gene screen that had an effect upon CHO cell growth and productivity, analysis of the multigene experiment was carried out using Design-Expert® software. Analysis was performed on all gene screen combinations but, as the only significant positive results came from DoE 5 (transcription factor combinations; figure 4-5 c), only DoE analysis from this factorial is shown.

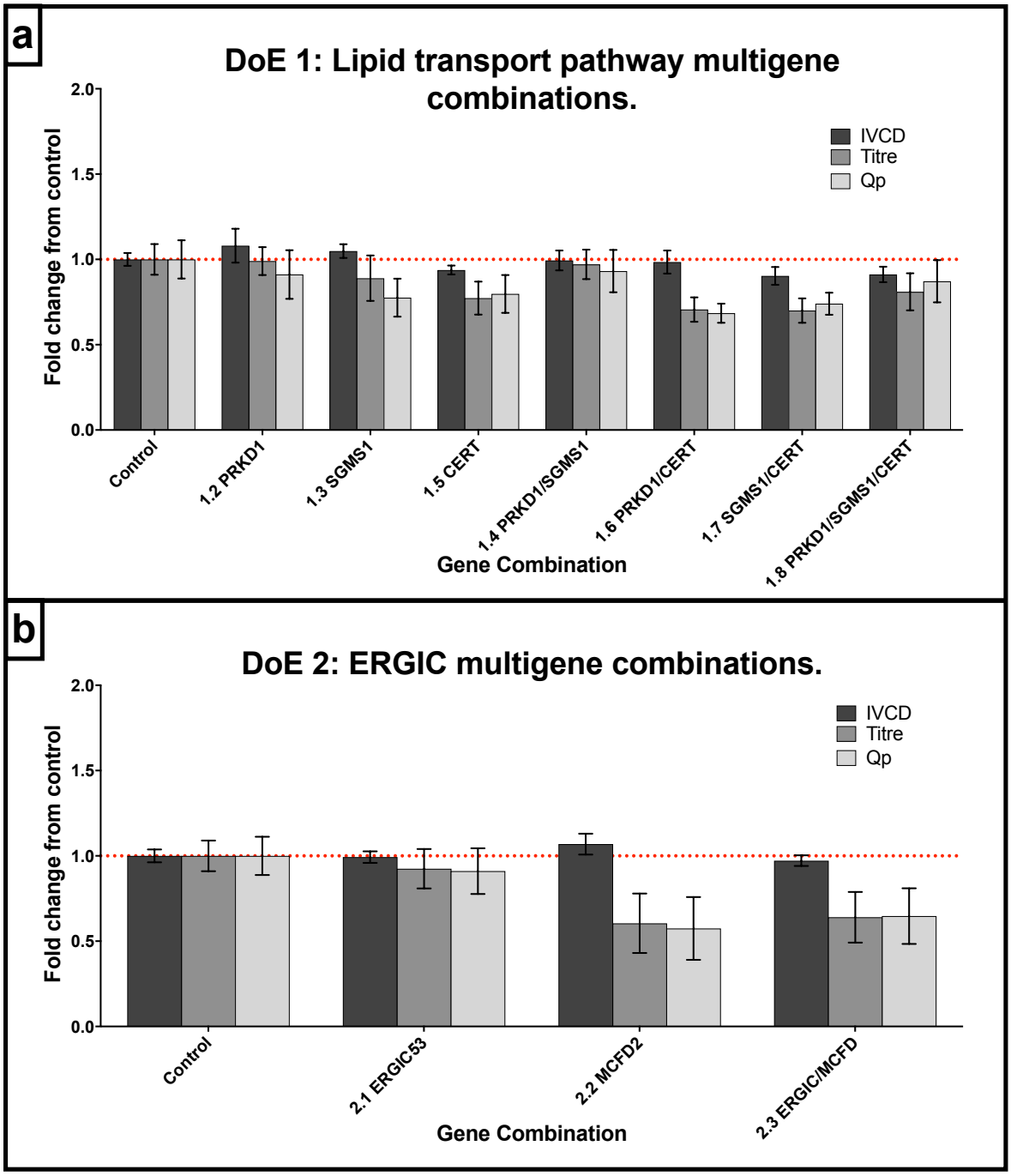


**Figure 4-3. Multigene engineering of Lipid transport and the ERGIC: Transient expression of DTE-Mab and up to three effector genes.**

*MedI-TRAN cells were transiently transfected with DTE-Mab alongside gene combinations designed to engineer the [a] lipid transport pathway, and [b] the ERGIC, both with a view to enhancing cell productivity. After five days' growth in 96 deep well plate culture growth and titre samples were analysed. Experimental data is shown as a fold-change compared to the control results. A Dunnett's test was performed to compare all results to that of the relevant control and determine any statistical significance. Columns show the mean of three biological and two technical replicates ( n =6) for DoE gene combinations, and an average of six biological and two technical replicates (n = 12) for DTE-Mab only control cells. Error bars show SEM.*

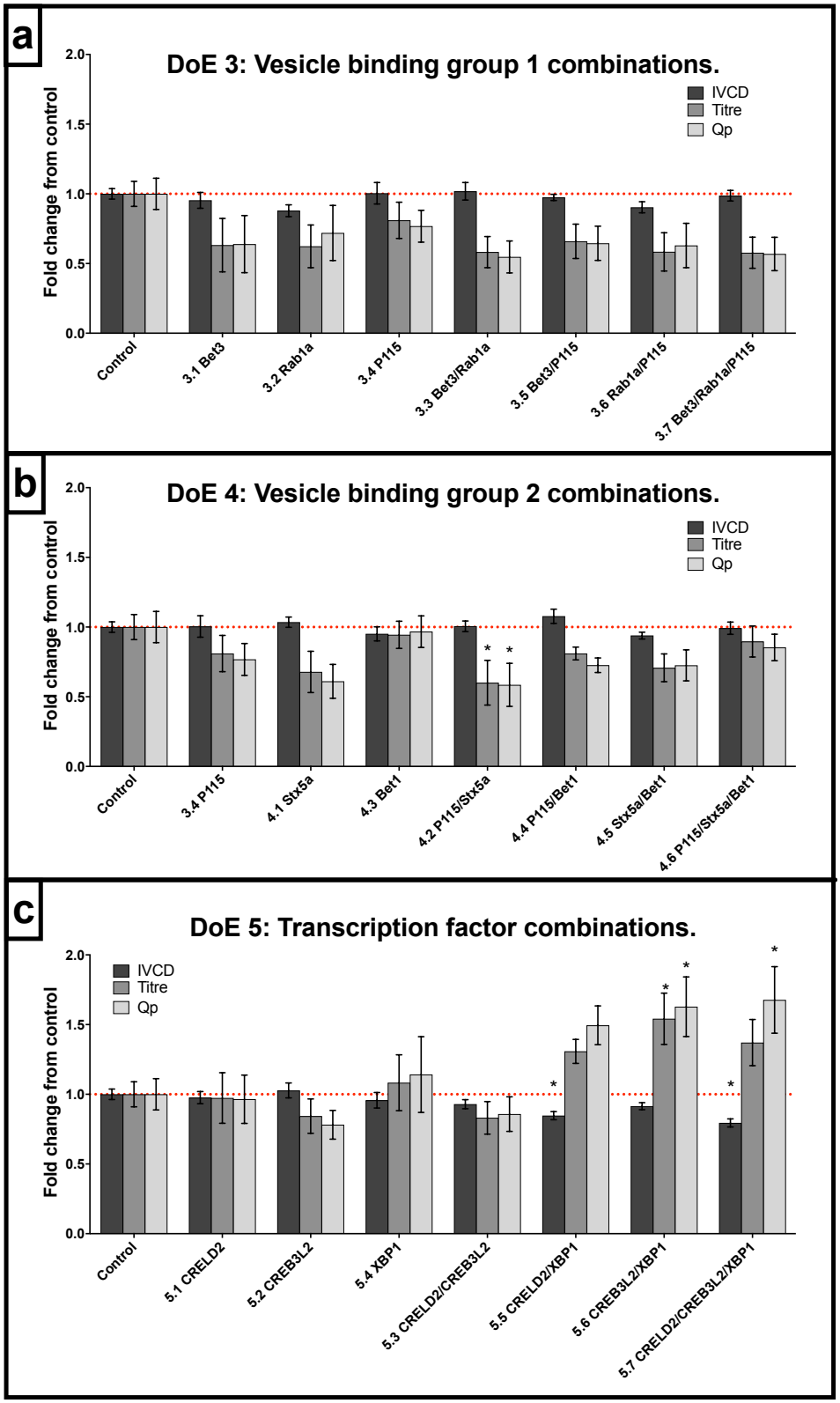


**Figure 4-4. Multigene engineering of vesicle binding and transcription factors: Transient expression of DTE-Mab and up to three effector genes.**

*CHO cells were transiently transfected with DTE-Mab alongside gene combinations designed to engineer [a, b] vesicle binding and tethering in the early secretory pathway, and [c] biosynthetic and secretory-controlling transcription factors. Experiment shown is as in Figure 4-4. Statistical significance was calculated using a Dunnett's test is shown by asterisks.*

For CHO cell growth DoE analysis shows that XBP1 on its own and XBP1 and CRELD2 in combination have a negative effect upon cell growth. Probability scores show that CRELD2 and XBP1 are both significant model terms (i.e. they have an effect upon the outcome being measured), with an adequate signal to noise ratio of 9.471 achieved (greater than 4 being perceived as adequate). ANOVA analysis produces a model F-value of 16.27, implying that the model produced is significant, in this case with 99.99% confidence (i.e. there is a 0.01% chance an F-value this large occurred due to noise; table 4-4). The final equation in terms of actual factors is shown in table 4-6. The half-normal plot shows that CRELD2 and XBP1 deviate from the normal and have a negative effect upon cell growth (figure 4-6 a).

With regards to overall titre, engineering with XBP1 individually and a XBP1/CREB3L2 combination both have an effect (figure 4-6 b). An F-value of 9.18 implies the model is significant with XBP1 a significant term, though individually CREB3L2 is not a significant term (table 4-4, 4-5). In terms of Qp the model is not significant (figure 4-6 c; tables 4-4, 4-5)

*Tables 4-4 and 4-5 (next page) show the DoE equation and model that came out of the DoE results shown in figures 4-4 and 4-5. The model is shown to be significant, with XBP1 on its own and XBP1 and CREB3L2 in combination shown to have a positive effect upon CHO productivity when stably expressing a DTE-Mab.*

**Table 4-4. DoE 5 Analysis of variance table.**

| Analysis of variance table [Partial sum of squares - Type III] | | | | | | |
|---|---|---|---|---|---|---|
| **IVCD/growth.** | | | | | | |
| **Source** | **Sum of Squares** | **df** | **Mean Square** | **F Value** | **p-value Prob > F** | **notes** |
| Block | 0.286021538 | 11 | 0.026001 | | | Significant |
| Model | 0.210649955 | 2 | 0.105325 | 16.27179 | 7.856 $E^{-06}$ | |
| A-CRELD2 | 0.077623323 | 1 | 0.077623 | 11.99213 | 0.001338 | |
| C-XBP1 | 0.133026631 | 1 | 0.133027 | 20.55145 | 5.636 $E^{-05}$ | |
| Residual | 0.245968585 | 38 | 0.006473 | | | |
| Cor Total | 0.742640079 | 51 | | | | |
| **Titre.** | | | | | | |
| Block | 2.954967844 | 11 | 0.268633 | | | Significant |
| Model | 2.300018008 | 3 | 0.766673 | 9.178685 | 0.000113 | |
| B-CREB3L2 | 0.004126053 | 1 | 0.004126 | 0.049398 | 0.825337 | |
| C-XBP1 | 1.847235317 | 1 | 1.847235 | 22.11529 | 3.521 $E^{-05}$ | |
| BC | 0.451842465 | 1 | 0.451842 | 5.409505 | 0.0256 | |
| Residual | 3.090517899 | 37 | 0.083528 | | | |
| Cor Total | 8.345503751 | 51 | | | | |
| **Qp.** | | | | | | |
| Block | 5.212522 | 11 | 0.473866 | | | Not significant |
| Model | 0 | 0 | | | | |
| Residual | 7.978766 | 40 | 0.99469 | | | |
| Cor total | 13.19129 | 51 | | | | |

**Table 4-5. DoE 5 final equations in terms of actual factors.**

| **Final equation in terms of actual factors:** | |
|---|---|
| **Growth =** | |
| 1.006821444 | |
| -0.010988512 | * CRELD2 |
| -0.007192536 | * XBP1 |
| **Titre =** | |
| 0.978236684 | |
| -0.012004106 | * CREB3L2 |
| 0.013530001 | * XBP1 |
| 0.001769655 | * CREB3L2 * XBP1 |
| **Qp =** | |
| 1.068771641 | |

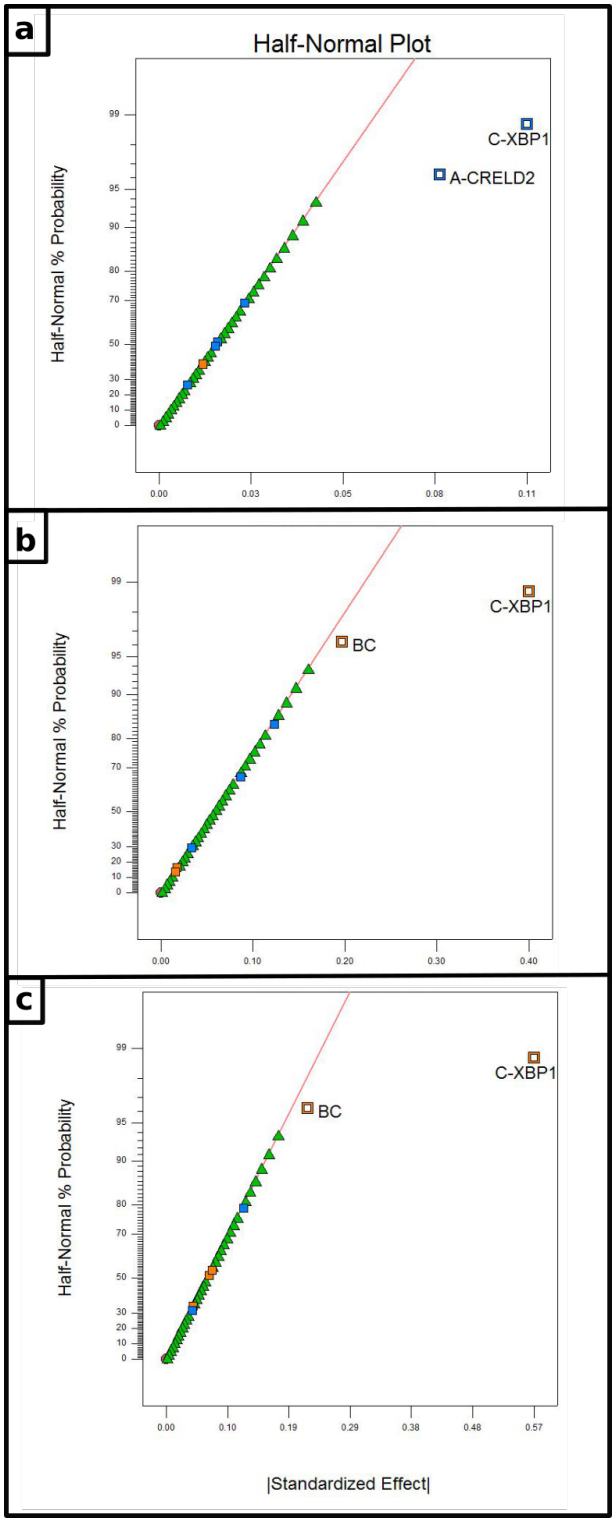


**Figure 4-5. Half normal plots highlighting effect of each gene on overall multigene effect of DoE 5.**

*Analysis of DoE 5 shows the effect engineering of CHO cells with XBP1, CRELD2 and CREB3L, both singularly and in combination, has upon [a] CHO cell growth (IVCD), [b] overall titre and [c] Qp. In plots gene effects are plotted along a normal distribution line, with points deviating from this line showing these factors have an effect upon cell growth, titre and/or Qp. Blue points show a negative effect with orange points a positive. A = CRELD2, B= CREB3L2 and C = XBP1, with combinatorial effects also shown by combining these pseudonyms.*

Overall this shows that whilst CRELD2 and XBP1s may have a slightly detrimental effect upon CHO cell growth, XBP1s and CREB3L2 have a combinatorial positive effect upon titre and Qp, though this is only statistically significant in the case of overall titre.

## 4.5. Discussion and conclusion

A gene screen of effector genes was performed in cells both transiently and stably expressing the same DTE Mab. This allowed the effectiveness of all the genes screened to be assayed in two model expression systems. Furthermore it allowed the comparison of gene effects between transiently and stably expressing cell lines to be compared, providing an overview of the effectiveness of genetic engineering of the CHO cell to enhance productivity in both screening and more production orientated processes.

The overall outcome of the transient gene screen was that many genes did not increase transient Mab expression. The exception to this was XBP1s which increased productivity up to 2.5-fold in both its human and CHO isoforms. In a transient system this replicates the XBP1s-mediated productivity enhancement seen in previous literature. Exact productivity increases can not be compared across this and previous experiments due to the different model products (including various Mabs, SEAP, SAMY and VEGF) , cell lines and protocols used - used (Rajendra et al. 2015; Tigges and Fussenegger 2006). Transient expression of a Mab in a CHO cell line stably expressing XBP1s has also previously resulted in increased productivity (Fig. 4-1; Cain et al. 2013).

However here the overexpression of XBP1s in a CHO cell line stably-producing a DTE-Mab did not result in the increase in productivity that has been seen previously and in the transient-expression system (Fig. 4-3; Becker et al. 2008). This lack of change suggests that, during the process of developing a stable pool cells may have developed a phenotype where ER stress response is dampened down. As such recombinantly expressed XBP1 may have no 'target' on which to act to improve cell productivity.

Continued overexpression of a DTE is likely to have put the cell under a level of chronic ER stress. It has previously been shown that this can result in suppression of the UPR response (Gomez and Rutkowski 2016). It has also been shown that under stress conditions ectopic expression of XBP1s (in CHO-S cells) can repress endogenous XBP1 splicing without also repressing the PERK pathway, suggesting the presence of a feedback mechanism in CHO-S cells to attenuate IRE1 activity in the presence of XBP1 (Chalmers et al. 2017). Of course the opposite could be happening, with continuous DTE expression and the ER stress this would bring resulting in XBP1s already being present and active at a

high level within the cell and as such further overexpression would not have much of an effect upon the cell. However this has not been seen in previous literature, with an increase in XBP1 not seen in CHO 'omic data in expressing cell lines.

The transcription factor and chromatin regulator YY1 has previously been shown to increase stable Mab titre by close to 4-fold in CHO cells, increasing both intra- and extra-cellular Mab levels. It was also shown to increase transient production by close to 3-fold. In both systems increasing YY1 DNA levels by too much resulted in a reduction in productivity enhancement (Tastanova et al. 2016). This suggests that excessive overexpression of a transcription factor can have less of an effect upon transient productivity and highlighting the need to perform dose-response tests with potential gene engineering targets. Here, unlike previous data, YY1 expression had no impact upon transient production levels. Whilst the titre of the stably-expressed Mab did not increase, the Qp saw a 1.4-fold increase, which somewhat matches the literature (figure 4-3 c).

Expression of genes involved in disulphide bond formation has previously been shown to increase productivity in CHO cell lines both transiently and stably expressing Mabs. In this screen, expression of Ero1-Lα transiently had no effect upon transient CHO productivity (figure 4-1), contradicting what has been seen previously (Cain et al. 2013). However, in this paper Ero1-Lα was stably expressed alongside XBP1s and whilst productivity did increase compared to an XBP1s-only cell line it is possible Ero1-Lα's function was enhanced by the activity of XBP1s. In this screen expression of PDIs generally resulted in a small increase in productivity in stably- but not transiently-producing cells (figure 4-1, 4-3). This matches trends seen previously in stably-expressing cell lines (Borth et al. 2005). However it does not match data from previous transiently-expressing studies, although in these it has also been shown that productivity is somewhat dependent upon the specific product (Johari et al. 2015; Pybus et al. 2014).

Both previous studies showing an increase in CHO productivity upon CERT overexpression were carried out in cell lines stably expressing both the product (Mab or tPa) and CERT (Florin et al. 2009; Rahimpour et al. 2013). As such transient CERT expression has not previously been reported. However, both studies show the opposite to what was seen in both the transient and stable gene screens where CERT expression resulted in a decrease in productivity (figure 4-1, 4-3). This could be due to the lower levels of CERT produced when expressed transiently, as it may prove that CERT must be expressed at a level above that possible in a transient system to have an effect upon ceramide transfer and Golgi lipid synthesis. However, that high levels of CERT were out-performed by low levels of CERT in the stable producer would suggest this is not the case.

Interestingly reduced ceramide synthesis, brought about by knockdown of the ceramide synthetase (CerS2), has been shown to alter the ceramide content of membranes and increased CHO cell IgG productivity (Pieper et al. 2017a), suggesting reduced, not increased, ceramide levels are linked to increased CHO cell productivity.

Whilst it is not possible to directly compare the vesicle tethering and fusion genes expressed in the gene screen to previous literature, previous expression of SNAREs (specifically SNAP23 and VAMP8) and SNARE effectors has been performed with both products and effectors transiently expressed. In both these cases increasing SNARE level or their fusion activity resulted in an increase in CHO productivity (Peng et al. 2011; Peng and Fussenegger 2009b). Whilst most of the vesicle docking proteins expressed increased CHO Qp in stably-expressing cells (Bet3, Bet1, Stx5a, P115, Rab11a, Rab1a), it is only by a small amount (up to 1.4-fold) and with little statistical significance and has less of an impact upon titre than Qp. All these genes also have little impact upon transient productivity levels.

Stable expression of SRP14 protein has been shown to increase stable Mab production (Le Fourn et al. 2014). This is contrary to what is seen in our screens where SRP expression resulted in decreased transient production and only a small increase in stable production. Increasing ATF6 activity has been shown to increase CHO transient productivity (Johari et al. 2015; Pieper et al. 2017b), but was not shown to increase the productivity of a stable product though again productivity was linked to the specific product used (Pybus et al. 2014). This matches data from this screen where expression in a stably-producing cell line resulted in a slight increase in Qp but a decrease in overall titre. Increasing Rab1 activity in CHO cells, achieved through knockdown of Tbc1D20, has been shown to increase CHO productivity, albeit in combination with CerS2 knockdown (Pieper et al. 2017a). Here Rab1 had no impact upon productivity in transient or stably producing cells, though a slight increase in stable Qp is seen once normalised for DNA load.

That the two main plasma-cell derived transcription factors CREB3L2 and CRELD2 both increase transient DTE-Mab titer by 1.4-fold somewhat upholds the hypothesis that transcription factors with a role in plasma cell differentiation can increase productivity of a transiently-expressed Mab. However, these transcription factors had no effect upon stable Mab production, wheras more specific targets did appear to have a slight impact upon stable Mab productivity. This mirrors the lack of effect on stable production of XBP1, with the overall effect of the three genes (and that of ATF6 on titer) suggesting that transcriptional engineering of the UPR in stably-producing cells is not a feasible strategy.

This would suggest that in a stably-producing system, especially when a DTE Mab is being produced, the cell population manages to overcome the UPR and as such it can not be further enhanced. This could be somewhat false and be due to either the selection process involved in stable pool production, or the natural selection of a cell line with enhanced ER capacity that does not require further UPR activation and therefore has an advantage over the rest of the cell population and thus becomes predominant. It is also possible that stable expression of a DTE-Mab induces chronic ER stress upon the cell, which can suppress the UPR through dampening expression of BiP mRNA (Gomez and Rutkowski 2016).

Of the genes that have not previously been ectopically expressed in CHO cells, many produced a slight increase in titre and productivity as discussed in section 4.3.3.2. Expression of a single SNARE or tether protein may slightly enhance cell productivity. However, as different SNAREs and tethers interact with other similar vesicle docking machinery proteins, the over-expression of a single gene is unlikely to bring about an increase in vesicle docking frequency.

Whilst chaperones can work independently of each other, it is possible that overexpression of one chaperone could only shift the protein processing bottleneck elsewhere within the ER, therefore not enhancing overall CHO productivity. Expression of HSPA1a has been shown to improve both titre and Qp of BHK-expressed factor VIII by approximately a third (Ishaque et al. 2007). Here it can be seen that in CHO cells HSPA1a expression resulted in no change in the titre of a stable Mab, but a small (25%) increase was seen in transient production levels at the 10% level of HSPA1a expression (figure 4-3 b, c).

The multi-gene engineering approach to CHO cell engineering did not result in an increase in transient CHO productivity in the four knowledge-driven gene combinations (lipid transporting, ERGIC and two vesicle docking combinations; Fig. 4-4 and Fig. 4-5 a, b). The data- and knowledge-driven combination (transcription factors; Fig. 4-5 c) did result in an increase in both Qp and overall titre, with XBP1 and CREB3L2 appearing to be the genes responsible for this effect according to DoE analysis (figure 4-6, tables 4-3 and 4-4).

However, the results of this multigene screen should be taken as implied and not as definitive as the DoE framework allows. Comparing the single-gene transfections from the DoE to those of the initial single-gene screen (figure 4-1) exposes quite marked differences between the same genes between different the screens. Whilst there was little difference in growth between the two separate experiments, in the multi-gene experiment titre and Qp levels appear to decrease more than seen in the initial single-gene screen. This is especially obvious in the

case of sole expression of XBP1s-CHO. A 2.5-fold increase in titre is seen in the single gene screen (figure 4-1 b), compared to a mere 1.2-fold increase in the multigene experiment (figure 4-5 c).

These differences could be explained by the presence of plate-to-plate, -and even intra-plate - variability during the multigene screen. Whilst this was present in the single-gene screen it was not to the same level as that seen in the multigene experiment. In the multigene experiment variability in titre and productivity was seen even between control replicates on the same plate, as well as between technical replicates of both test and control subjects on repeat plates. This variance is shown in the large error bars seen in the multigene experiment, with very few results shown to be statistically significant when compared to the control. It is also possible that the increased transcriptional burden of expressing up to four genes from separate plasmids produced too much of a transcriptional and translational burden upon the cell, impinging upon cell growth and productivity.

Across all the knowledge-driven gene combinations no positive effect upon CHO productivity was seen (figure 4-4 and 4-5 a, b). In these cases it is likely that even three genes is not enough to affect the enhancement of an entire pathway so as to increase capacity. Increasing the presence of only three genes may open a pathway up a bit more, but again it could shift the functional bottleneck to elsewhere within the pathway. For example with the lipid transport gene selection (PRKD1, SGMS1 and CERT), phosphatidylinositol kinases also play an important role in Golgi structure and function in this pathway and as such omitting this gene from the combination could limit its effectiveness (Hausser et al. 2005).

Excessive levels of DAG could well destroy a careful balance of lipids within the TGN membrane causing it to function incorrectly, though this has not been seen previously when CERT ahs been over-expressed in CHO cells. It is possible that over-engineering SGMS1 could favour sphingomyelin synthesis over that of DAG, whilst it has been suggested that not only DAG but protein kinase C may be required for PRKD1 activation (Wang 2006).

Whilst expression of a small selection of SNARE proteins has been shown to increase CHO productivity levels, even co-expressing proteins known to work together has previously resulted in no increase in productivity, highlighting that knowledge-driven approaches do not always work within the cell (Peng and Fussenegger 2009b). Furthermore the wide range of tether, SNARE and effector proteins that function in vesicle docking and fusion at their target membrane is too much to directly engineer. As is seen with XBP1 and CREB3L2, engineering the CHO cell with a global controller of a favourable pathway (in these cases the

UPR and secretory pathways respectively), allows the cell to make myriad changes at a transcriptomic level, enhancing the protein processing and secretion machinery to achieve an increased productivity phenotype.

With all genes treated the same it can be assumed that all data in both single- and multi-gene screens is comparable. However, it is possible that some genes are more easily transcribed than others, or their products more active due to post-translational modifications or more favourable stoichiometric ratios. Because of the large-scale high-throughput process used to test the genes, it was not possible to check the mRNA or proteomic levels of each gene, even at a low level (e.g. by RT-PCR or Western blotting/fluorescence measurement respectively), as has been performed in previous low-throughput gene expression studies to confirm increased levels of the accessory gene and/or its protein (Borth et al. 2005; Cain et al. 2013). This technique was also used to ensure there was no change in Mab expression levels between all test and control cells, allowing any increase to be attributed to the effect of the accessory gene(s) and not due to any possible increased Mab transcriptional levels (Peng and Fussenegger 2009b; Tastanova et al. 2016). Again this was not possible with the high-throughput system. However, as all accessory genes used were effectors of the UPR, protein folding and secretion, it can be considered that any effect they had upon cell productivity was due to their molecular role in protein folding, processing and/or secretion, not due to increasing the transcription rate of any product gene.

Expression of genes that increase overall cell growth or size could be a potential pathway with which to increase product yields. However, the Largen gene did not have any effect upon CHO cell productivity (figure 4-1). Furthermore, increasing the biomass of a culture past a certain point, through increasing either cell density or size, can increase sheer stress within bioreactors or increase burden upon the downstream processing of product, both of which can impinge upon ultimate final process yield. As such engineering within the cell to increase productivity without increasing biomass is still seen as the optimal process by which to improve overall product yields.

To better understand the potential value of direct multigene engineering of CHO cells in an industry-relevant scenario, the tested gene combinations would need to be stably expressed alongside a stably-expressed recombinant product. As seen in the single-gene screen, the effect of multiple genes is likely to differ in a cell line stably expressing a product compared to a transient expression system (section 4.3.3). Furthermore, for industry-relevant practices expression of recombinant product and accessory genes from up to four separate plasmids is not feasible. As such testing of accessory gene and product combinations expressed on the same MGEV including synthetic promoters is likely to provide a

simplified engineering solution, especially with regard to stable expression of multiple product and accessory genes from one transcriptionally active site in the genome. Use of synthetic promoters in this system can also ensure better transcriptional control of both the recombinant product but also the effector genes, allowing stoichiometry to be better managed (Patel & Brown et al., manuscript in preparation; Brown and James 2016).

From the results in this chapter it is clear that whilst single-gene engineering of a transiently-expressing cell line can increase CHO productivity, less of an effect is seen in a stably-expressing cell line. Furthermore, a transient gene screen would appear to not be the best way in which to screen genes that will be used in an attempt to enhance stable production. Multi-gene engineering of CHO cells can enhance the effect of single-gene engineering when gene combinations are data driven. Whilst it may be enticing to use knowledge-driven approaches to multi-gene CHO engineering data here shows that this results in no increase in productivity. This is likely due to there being a finite (and generally small) number of genes that can be expressed, taking little consideration in to the overall molecular balances and fine-tuning required for pathways such as the secretory pathway to function correctly.

# 5) Directed evolution of the CHO secretory pathway.

*Directed evolution was used as a method with which to effect large-scale genetic changes within the CHO cell with a view to creating a cell line with an enhanced productivity phenotype by evolving against a secretory blocking agent. Previous directed evolution strategies were reviewed and a review of secretory blocking agents was carried out to inform selection of suitable secretory-blocking agents against which to evolve the cell line. A parental CHO cell line was evolved against the selected secretory blockers. To quantify the effect the directed evolution strategy had upon CHO cell productivity levels, transient and stable transfections of the evolved cell lines were performed with easy- and difficult-to-express model monoclonal antibodies. Morphology of the evolved cell line was analysed to investigate possible molecular reasons for any phenotypic changes that occurred.*

## 5.1. Directed evolution and its use in mammalian cell engineering.

The inherent genetic instability within CHO cells has been leveraged by industry to help produce cell lines with phenotypes desirable for manufacturing and research purposes. Screening paired with single-cell sorting allows for the selection of a single cell with desirable growth and productivity characteristics. Over time this population would eventually diverge, becoming genetically, epigenetically and phenotypically diverse, as has been shown by growth of CHO clones over time resulting in a population's growth rate gradually increasing, even after the cells have been allowed to adapt to their growth conditions (Fernandez-Martell et al. 2017). As the cell culturing process progresses, faster-proliferating cells become more prevalent within the population, resulting in evolution of the population's growth rate. If basic long-term culturing of CHO cells can result in the enhancement of a desirable phenotype within the population, a more designed and specific long-term growth strategy, utilising an agent targeting a specific cellular pathway, could result in an enhanced phenotype specific to that pathway.

Directed evolution of cell lines can refer to two distinct processes by which new phenotypes are produced. The first and most common of these is based on a mutate-and-screen process whereby a mutant library is produced then screened for a desired phenotype, cells with this phenotype being selected. Mutations can be instigated by treatment with non-specific methods (e.g. some chemicals, UV light), resulting in mutations across the entire genome. More specific techniques

(epPCR) can be used to target specific genes/genomic areas, producing a more defined set of mutations (Yang et al. 2017). Libraries are then screened to select those with desirable phenotypes. Due to the slow growth characteristics of CHO, this screening process does not really improve upon current technologies and techniques used within cell line development (Wurm 2004). Furthermore, mutations produced will alter a specific target and, in the process of screening, more nuanced changes around this target may well be lost. However, the potential for many mutations to occur, affecting the whole cell, may result in global genetic changes and therefore more interesting phenotypic changes.

The second method involves imposing a harsh condition upon the cell population and allowing it to naturally overcome this over time. These conditions can be either global or specific to a certain problem/pathway within the cell. This method of directed evolution has been used to produce mammalian cell lines that are highly resilient to bioreactor stresses (e.g. increased ammonia/lactate concentrations, shear forces, nutrient deprivation) to improve cell growth within bioreactors (Matsumura et al. 1991; Prentice et al. 2007), as well as producing cold-adapted CHO cells, resulting in normal growth metrics despite a larger volume and increased productivity compared to the parental line (Syddall et al., manuscript in preparation). Continued exposure of CHO cells to the apoptosis inhibitor staurosporine resulted in progressively lower apoptosis induction thresholds, showing repeat exposure to a stress can bring about enhanced performance (Majors et al. 2012).

This method of exposure to a harsh environment is able to utilise any small variances within a cell population that is initially clonally derived and therefore has little genetic diversity. However, the targeting of specific pathways within the cell – e.g. secretion, amino acid transportation – with an aim to enhance their capacity and therefore cellular productivity, is a more recent development. As it has been shown that molecular inhibition of a target can sometimes be overcome by overexpression of that target (Claude et al. 1999; van der Linden et al. 2010), it stands to reason that a cell population may be able to naturally overcome an artificial blockage brought upon by a specific agent. This may result in an increase in molecular components of that pathway, or potentially enhancing neighbouring pathways (active or redundant) to overcome the imposition. It is likely that multiple fine changes at the molecular and genetic level will occur, resulting in a genetic engineering far more subtle and nuanced than would be possible through ectopic expression, selected based upon pathway mapping, of a small collection of genes at fairly uncontrolled levels.

The hypothesis has been developed that, if the evolutionary target and the agent with which it is inhibited are chosen well, a specific pathway - or section thereof - could be blocked. Allowing the population to naturally overcome this inhibition

could result in the development of a cellular phenotype with increased productivity, in this case due to cells overcoming a blockage of the cellular secretory pathway resulting in an enhanced secretory phenotype.

Mutation and selection has been used to good effect in both microbial and mammalian cells to evolve new desired phenotypes. However, due to the relatively long generation time of mammalian cells, applying mutation techniques to mammalian systems can be time consuming. As secretion is controlled by a complex pathway and not a single protein, a whole-cell directed evolution method, instead of random mutations of single proteins, was used. This global approach has been shown to produce a wide range of unique phenotypes (Majors et al. 2009).

By blocking the secretory system of CHO cells, it was intended that phenotypic variations across the cell population that were already present would be allowed to proliferate under the selective pressure. This was expected to result in the proliferation of cells that have overcome the secretory blockage due to an increased secretory phenotype. Stepped increases in the concentration blocking agent would allow greater evolution of the cell population and therefore greater enhancement of the CHO secretory pathway. When returned to normal conditions it was hoped that the evolved cell line would have an enhanced secretory pathway capacity, increasing secretion levels of recombinant protein products and therefore increasing production levels when compared to the wild type cell line. An example of a potential directed evolution process against an inhibitory agent is shown in Figure 5-1.

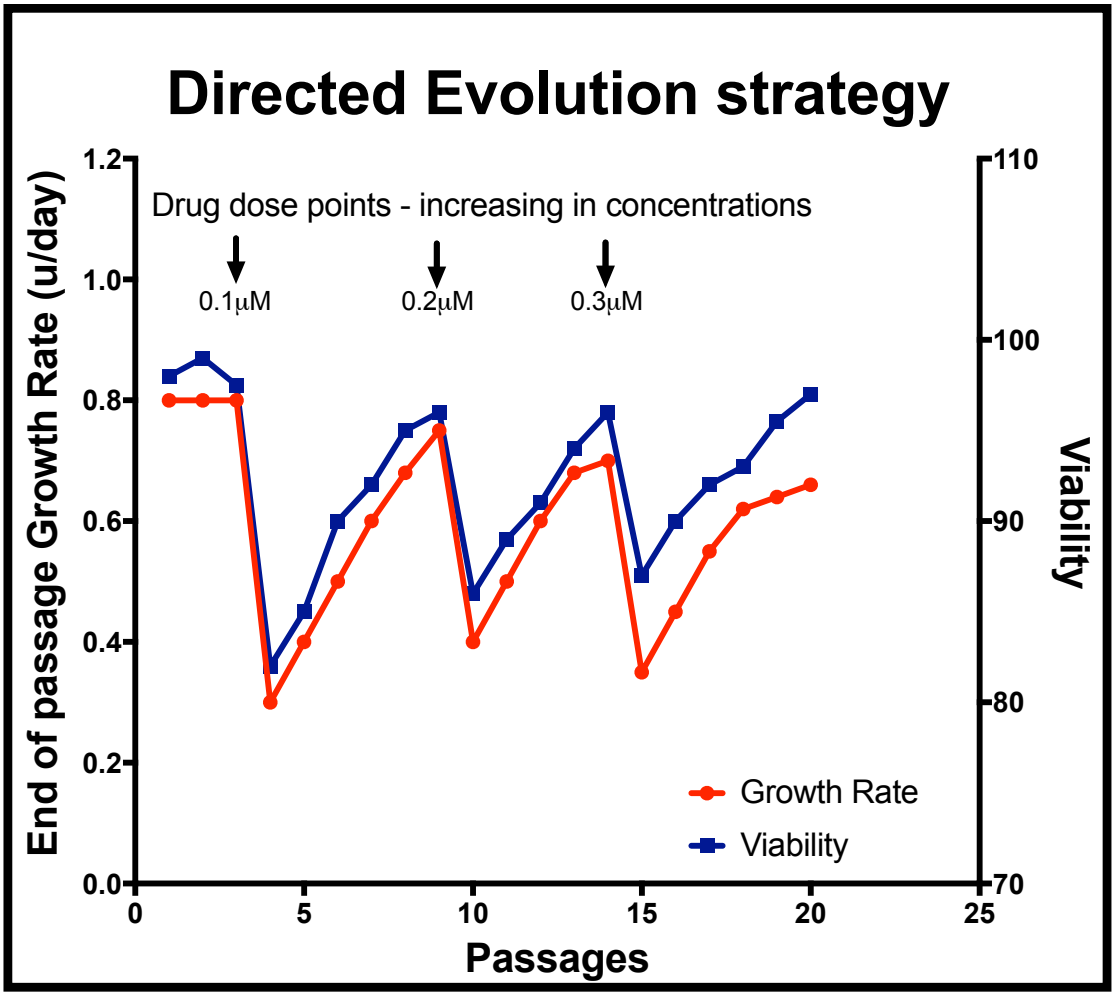


**Figure 5-1. Suggested strategy for directed evolution against an inhibitory agent.**

*After recovery from vial, cells would be treated with inhibitory agent at a lethal dose of between 50-75% (knocking growth down to 25-50% that of untreated cells), with compound added at each passage. After gradual recovery of cell growth and viability towards normal levels, concentration of inhibitory blocker would be increased, with the process being repeated upon cell growth recovery. The resulting cell line may have a slightly reduced level of growth, but when removed from the selective pressure this would be expected to return to normal levels.*

## 5.2. Selection of secretory blocking molecules for directed evolution

Many secretory-blocking chemicals have been used to dissect and therefore elucidate the mammalian cell secretory pathway. The majority of secretory-blocking compounds have been used to dissect the secretory pathway of many mammalian cell lines, but not specifically that of CHO cells. As such, compounds that shown to work within the literature may not necessarily have the desired effect within CHO cells.

Specific criteria were identified for the selection of secretory-blocking agents for use in a directed evolution experiment:

1. Targets a specific region of the mammalian secretory pathway, preferably a known molecular target.
2. Ideally targets a single molecule (though some pleiotrophy may be acceptable if affecting other regions of the biosynthetic and transport pathways).
3. Does not impact upon other essential pathways within the cell.
4. Does not impact upon recombinant protein production.
5. Soluble in water, ethanol or DMSO at a level at which these can be used in cell culture.
6. Commercially available.

Many secretory pathway proteins have family members present within other transport pathways (e.g. endocytosis). Due to the similarity of these proteins and their active domains, some off-pathway effects were expected, though if these were to greatly impinge upon cell growth and function, these cells would likely not survive the selection process. From previous secretory studies in mammalian cells, compounds targeting a range of secretory targets were considered. These are summarised in the following section.

### 5.2.1. Summary of compounds targeting the COPI vesicle transport pathway via Arf and Arf-related proteins

As discussed in section 1.5.1, COPI vesicles mediate both intra-Golgi and Golgi-ER retrograde transport, this latter process being a recycling step for COPII vesicle components. Blocking of COPI vesicle formation with Brefeldin A (BFA) halts ER-Golgi transport and therefore general protein trafficking in rat hepatocytes (Misumi et al. 1986). As such this is an ideal section of the secretory pathway to target for directed evolution as if CHO cells can overcome the blocking of COPI vesicle formation this will hopefully result in an increased secretory phenotype due to an increase in molecular components of the COPI pathway as well as potentially enhancing surrounding secretory pathways and components. Formation of COPI vesicles on the Golgi membrane is mediated by the ADP-

ribosylation factor 1 (Arf1) GTPase, which when activated recruits coat proteins and other effector molecules to the *cis*-Golgi membrane, creating a membrane region from which vesicle formation and budding can occur (Beck et al. 2009).

Arf1 activity is regulated by Arf-GEFs (generally activating) and Arf-GAPs (generally deactivating). The Golgi-specific Brefeldin A-resistance factor GEF 1 (GBF1) is a peripheral Arf-GEF associated with formation of COPI vesicles at the Golgi (Beck et al. 2009; Niu et al. 2005). The nucleotide exchange activity of Golgi-related Arf-GEFs is catalysed by a well conserved central Sec7 domain (Chardin et al. 1996; Jackson and Casanova 2000). Many drugs target this region which, when blocked by molecular binding, stops the GEF interacting with and activating its requisite Arf (Mansour et al. 1999; Saenz et al. 2009).

Despite generally being highly conserved across Arf-GEFs, the Sec7 domain differs slightly in GBF1 with an extra three residues extending a helix that makes up a face of the domain (Boal et al. 2010). Were a GBF1-specific compound to be used on CHO cells, it would impact upon COPI vesicle formation at the *cis*-Golgi without having off-target effects upon other Arf-GEFs, specifically the Golgi-localised Brefeldin A-inhibited GEFs 1 and 2 (BIG1 and BIG2), which are involved in endocytosis and endosomal and TGN transport (Anders and Juergens 2008).

Literature searches unearthed several compounds that could have a similar but more targeted effect to that of BFA. These could be used to treat CHO cells to modify the CHO secretory pathway with a view to enhancing CHO secretory capacity. A summary of these compounds can be found in table 5.1.

#### *5.2.1.1. Brefeldin A (BFA)*

BFA is a fungal lactone antibiotic discovered in 1968 that inhibits protein trafficking and induces morphological changes in secretory pathway organelles (Misumi et al. 1986; Tamura et al. 1968). The Sec7 domain has been shown to be BFA-sensitive and thus BFA blocks Arf1 activation, the formation of COPI vesicles at the Golgi, COPII-component recycling to the ER and intra-Golgi transport (Claude et al. 1999; Guillemain and Exton 1997; Mansour et al. 1999; Orci et al. 1991). Treatment of mammalian cells with BFA results in the merging of the Golgi and the ER due to the imbalance caused by on-going retrograde transport not being balanced by anterograde transport (Boal et al. 2010).

BFA-resistant CHO cell lines have previously been developed and shown to generally be genetically stable with their resistance not due to BFA degradation (Yan et al. 1994), suggesting resistance is due to a change in BFA's target or changes overcoming BFA's effect. At low concentrations (1 μg/mL) BFA does not have an effect upon protein production, although an effect is seen at 10 μg/mL (Misumi et al. 1986). The effective concentration of BFA is dependent upon the

cell species type, with BFA functional in reducing CHO cell growth (Orci et al. 1991; Torii et al. 1995), but ineffective within some mammalian cell lines such as MCDK canine cells. This is likely due to differences in GBF1's Arf-binding site in these cells (Boal et al. 2010; Hunziker et al. 1991; Prydz et al. 1992). Sec7-localised mutations have also been shown to alter the BFA resistance of yeast cells (Peyroche et al. 1999). Were this type of point mutation to be an outcome of the evolutionary process it would most likely result in a BFA-resistance phenotype but without the desired enhanced secretory or productivity phenotype.

In human cells BFA is pleiotropic, targeting and inhibiting at least 3 Arf-GEFs (GBF1, BIG1 and BIG2) that activate Arf1 proteins involved in vesicle formation and as such BFA treatment impacts upon endocytosis at the plasma membrane (Torii et al. 1995). That some BFA-resistant cell lines can maintain Golgi structure but are unable to grow in the presence of BFA further suggests that BFA has multiple organelle-specific targets (Yan et al. 1994), whilst overexpression of GBF1 has been shown to facilitate CHO cell growth in the presence of BFA (Claude et al. 1999). Treatment of mammalian cells with BFA blocked COPI vesicle transport but increased the formation of tubules between Golgi stacks, allowing transport of cargo to still occur by lateral diffusion (Orci et al. 1991).

#### *5.2.1.2. Exo1 [2-(4-Fluorobenzoylamino)-benzoic acid methyl ester]*

Exo1 targets the same section of the secretory pathway as BFA but via a different molecular mechanism that is not fully known (Ivanov 2014; Mishev et al. 2013). However it is likely to function through blocking Arf interactions with coat complexes, possibly acting as an Arf-GAP (Feng et al. 2003). Instead of directly inhibiting an Arf-GEF, it is believed to stimulate an Arf-GAP (Pauloin et al. 2016), deactivating Arf and therefore inducing the disassociation of Arf1 from Golgi membranes. Exo1 treatment inhibits traffic from the ER and induces a rapid collapse of the Golgi to the ER, but with a reduced effect upon TGN organisation compared to BFA. Unlike BFA, Exo1 does not directly interfere with Arf-GEFs or fatty acid exchange activity of other BFA-target genes (such as BIG1/2) so therefore would have fewer pleiotropic effects. Exo1 treatment of many cell types (including CHO) saw the blocking of VSVG-GFP exit from the ER, with treatment at 20 μM killing 50% of the cell population. Treatment with Exo1 is reversible, with Golgi markers reappearing as normal 30 minutes after its removal (Feng et al. 2003). However, it may act downstream of Arf1 (Pauloin et al. 2016) and not have a direct role in COPI vesicle recruitment at the Golgi.

#### *5.2.1.3. Exo2 (4-hydroxy-3-methyoxy-(5,6,7,8-tetrahydro[1]benzothieno[2,3-d]pyrimidin-4-yl)hydrazone benzaldehyde)*

Exo2 inhibits GBF1 function with minimal effect upon BIG1/2, therefore having little effect on endocytosis. Like Exo1 and BFA it blocks anterograde transport from the ER to the Golgi (Mishev et al. 2013) and completely disrupts the Golgi. Unlike BFA it has minimal effect on the TGN, retrograde transport and the endocytic pathway (Boal et al. 2010; Guetzoyan et al. 2010), although retrograde transport of toxin proteins in Exo2-treated cells varies depending upon the toxin (Feng et al. 2004). Exo2 may have an impact upon TGN morphology, if not its integrity (Spooner et al. 2008).

#### *5.2.1.4. Further COPI blocking compounds*

An engineered derivative of Exo2, **LG186** is a more selective Arf-GEF inhibitor, due to steric hindrance being antagonistic against GBF1 but not BIG1/2. It induces Golgi disassembly in cells unaffected by BFA (e.g. MDCK canine cells), whilst treatment of HeLa cells resulted in a decrease in Arf activation leading to a loss of COPI components from Golgi membranes, Golgi collapse and blockage of soluble and membrane-associated cargo transport, which remained in the ER despite COPII exit sites remaining present and relatively unaffected (Boal et al. 2010).

**Golgicide A (GCA)** is a Sec7 inhibitor, with GCA treatment of human cells inhibiting GBF1 function leading to disassociation of COPI coat proteins from the Golgi membrane and Golgi/TGN disassembly, halting protein secretion and causing protein aggregation at the ERGIC, with GBF1 relocating from cytoplasmic vesicles to the fragmented Golgi (Kraemer et al. 2013; Saenz et al. 2009). However, GCA treatment of BHK and canine cells did not match the response of human cells, suggesting that GBF1 in some cell lines is not affected by some drugs that target it, including in hamster-derived cell lines. As such, changes in cell morphology imparted by GCA are likely to come from its impact upon BIG1/2, as shown by retrograde transport and viral replication studies (Saenz et al. 2009; van der Linden et al. 2010).

**AG1478** inhibits GBF and has a lower cytotoxicity than BFA, selectively and reversibly targeting the *cis*-Golgi in human cells whilst not affecting endosomal compartments, the cytoskeleton, Golgi structure, internal cell pH or ATP levels and not impacting upon cell survival (Pan et al. 2008). **LM11** impairs targets the interaction between Arf1 and ARNO-GEF, producing a nonfunctional Arf1-GDP/ARNO complex, impairing Arf-dependent trafficking structures at the Golgi, including endosomal trafficking (Mishev et al. 2013; Viaud et al. 2007). **AMF-26** is an Arf1-ArfGef inhibitor, inducing reversible Golgi disruption but also apoptosis, inhibiting cell growth and disruption of recycling endosomes (Ivanov 2014; Mishev et al. 2013; Ohashi et al. 2012).

## 5.2.2. Summary of compounds targeting other regions of the mammalian secretory pathway

Disruption of intra-Golgi and Golgi-ER retrograde recycling transport is an obvious target for directed evolution due to the availability of many compounds that target it and the level of elucidation of this region of the secretory pathway. However, there are other less well-defined regions of the secretory pathway, the targeting of which should prove as fruitful during a directed evolution process.

#### *5.2.2.1. FLI-06*

FLI-06 is a dihydropyridine that affects the secretory transport of both transmembrane and secretory proteins in HeLa cells, inhibiting general secretion at an unknown step prior to cargo exiting the ER. In FLI-06-treated cells, recruitment of cargo to ERES is greatly reduced, but the continued presence of ERES markers suggests FLI-06 functions prior to the ERES and does not impact upon COPII vesicle budding, instead acting upon cargo recruitment. It also inhibits export from the TGN (Yonemura et al. 2016). FLI-06 does not interact with GBF1 but is shown to completely disrupt both the *cis*- and *trans*-Golgi whilst not affecting endocytosis or tubulating endosomes. Treatment changes ER morphology from a tubule to a sheet but does not result in ER-Golgi fusion or a large increase in Golgi stress, as is seen with BFA/GCA, although COPI dissociation takes longer than with BFA (Kraemer et al. 2013).

#### *5.2.2.2. Further non-COPI targeting secretory blocking compounds*

The antibiotics **Monensin** and **Nigericin** incorporate in to the *trans*-Golgi membrane, mediating $H^+/Na^+$ or $K^+/H^+$ exchange respectively, neutralising the Golgi lumen leading to Golgi stress and cisternael osmotic swelling This inhibited protein transport from the Golgi (with proteins accumulating in vesicles and storage bodies), the cell-surface expression of surface proteins and the reduction of effective protein glycosylation (Dinter and Berger 1998; Misumi et al. 1986; Miyata et al. 2013; Oku et al. 2011; Tartakoff 1983).

**Secramine B** inhibits activation and membrane-binding of the Rho GTPase Cdc42, which is essential for vesicle transport, halting Golgi-to-PM vesicle trafficking and perturbing Golgi-PM protein transport without an obvious impact upon retrograde transport (Mishev et al. 2013; Pelish et al. 2006). However, Cdc42 activation has many other cellular roles including cytoskeletal remodeling, salt transport and cell polarity determination and proliferation (Chi et al. 2013).

**Dispergo** induces reversible ER structure changes in HeLa and CHO cells, resulting in Golgi cisternae loss and tubulation, preventing ER-Golgi transport and cargo loading at the ER. (Ivanov 2014; Lu et al. 2013). **Okadaic acid** inhibits serine/threonine phosphatases, reducing amylase secretion by up to 90% in rat pancreatic cells. As well as impacting upon protein secretion and Golgi

morphology, it also inhibits protein synthesis even at low concentrations (Dinter and Berger 1998).

### 5.2.3. Summary and selection of compounds for directed evolution of the CHO secretory pathway

There are many drugs that target Arf-GEFs in modes similar to that of BFA. Whilst BFA is non-specific, targeting three Arf-GEFs (GBF1, BIG1/2), there are plenty of molecules that target only GBF1. However BFA may still prove a useful evolutionary agent as its mode of action is relatively well characterised and as such it could be used as a control compound. Like **BFA**, LM11 also impacts upon endocytosis so is not suitable. AMF-26 directly induces apoptosis and inhibits cell growth GCA and AG1478 have been shown not to work in rodent cell lines so are not suitable for work with CHO. As a specifically designed compound, LG186 is not available commercially. Therefore, to replicate BFA's actions whilst only targeting GBF1 (and therefore hopefully not impacting upon other pathways within the cell), **Exo1** and **Exo2** appear to be ideal evolutionary agents, targeting GBF1 via its ability to bind and process GTP respectively.

Of the compounds reviewed that do not target Arf1-mediated transport there is a range of secretory and ER stress pathway regions targeted. Monensin and Nigericin both induce Golgi stress through altering its $H^+$ profile and not through a specific section of the secretory pathway, so are not entirely suitable for specific secretory pathway targeting. Okadaic acid impinges upon protein secretion, but also protein synthesis and therefore is unsuitable. Dispergo and Secramine B are both potential evolutionary agents, targeting ER-Golgi and Golgi-PM transport respectively, but both do not appear to be commercially available.

**FLI-06** targets cargo-loading at the ERES and, whilst its exact mode of action is unknown, it has been shown to work in many different cell lines and targets a similar part of the secretory pathway as BFA/Exo1 and 2, but at a different point – COPII vesicle cargo loading rather than COPI component recycling – and will be tested in CHO cells as a potential evolutionary agent. A summary of all the reviewed secretory-blocking compounds can be found in Table 5.1.

*Table 5-1 (next page) summarises the secretory-blocking compounds discussed in section 5.2, their mode of action and commercial availability. Compounds highlighted in grey were taken forward for testing of their effectiveness in CHO cells.*

**Table 5-1 - Summary of secretory-blocking compounds reviewed.**

| Compound | Target(s): Molecular; *pathway* | Effect on cell; potential negative/off-target effects | Cell line; screen used | References | Commercial availability |
|---|---|---|---|---|---|
| **Brefeldin A (BFA)** | GBF1, BIG1, BIG2 Sec7 domain;<br>ER-Golgi transport, endocytosis | Blocks COPI vesicle formation. Golgi collapses into ER;<br>Targets endocytosis. | CHO;<br>*VSV.* | (Orci et al. 1991) | Sigma £129/5mg<br>MedChem €49/5mg |
| **Monensin/ Nigericin** | Ion transporters;<br>Golgi transporting of $H^+$ and $Na^+/K^+$ | Neutralisation of the Golgi lumen, leading to swelling and transport interruption;<br>Not a direct transport inhibitor. | CHO, Sycamore;<br>*VSV.* | (Dinter and Berger 1998), (Ono et al. 1985; Zhang et al. 1993) | Sigma £48.50/0.5 g |
| **Exo1** | Possibly an Arf-GAP;<br>ER-Golgi | Arf1 disassociation from Golgi. Collapse of the Golgi to ER, resulting in inhibition of traffic from the ER. Blocks ER exit. | CHO, HeLa, HEK293, NRK, BHK, MDCK;<br>*VSVG-GFP.* | (Feng et al. 2003) | Enzi £57/10 mg<br>Santa Cruz £45/10 mg |
| **Exo2** | GBF1 Sec7 domain;<br>ER-Golgi, intra-Golgi | Blocks anterograde transport from ER to Golgi. Disrupts Golgi integrity but not TGN | T84, Vero;<br>Cholera toxin uptake. | (Feng et al. 2004) | Sigma £61/5mg |
| **LG186** | GBF1;<br>ER-Golgi | More specific version of Exo2, targeting Arf1, stopping COPI vesicle formation. | HeLa;<br>GFP-tagged proteins. | (Boal et al. 2010) | Not available |
| **Golgicide A (GGA)** | GBF1 Sec7 domain;<br>ER-Golgi | Inhibits GBF1, inhibiting Arf1 function and therefore COPI vesicle formation. Transported proteins aggregate at ERGIC;<br>*Does not work in BHK-21 cells.* | Vero, BGN, HeLa, BHK-21;<br>Shiga toxin uptake;<br>Enteroviral replication. | (Saenz et al. 2009). (van der Linden et al. 2010) | Sigma £96.20/5mg<br>Select £79/5mg |
| **Okadaic acid** | Serine/Threonine phosphatases. | Exerts numerous effects on protein secretion and Golgi morphology;<br>Inhibits protein synthesis even at low concentrations. | CHO. | (Dinter and Berger 1998) | Not available |

| Compound | Target(s): Molecular; *pathway* | Effect on cell; potential negative/off-target effects | Cell line; screen used | References | Commercial availability |
|---|---|---|---|---|---|
| **FLI-06** | Unknown (cargo loading at ERES);<br>ER-Golgi transport, organelle disruption | Impacts cargo loading at ERES without impacting COPII release. Changes ER morphology from a tubule to a sheet. Disrupts cis/trans Golgi without impacting upon endosomes/endocytosis. Doesn't change Golgi stress or result in Golgi collapse into ER. | HeLa, U2OS, HEK293, zebrafish;<br>Notch signaling pathway receptor trafficking, Klotho protein. | (Kraemer et al. 2013; Yonemura et al. 2016) | Sigma £66.90/5 mg |
| **Tyrphostin AG1478** | GBF1;<br>ER-Golgi, intra-Golgi | Targets *cis*-Golgi whilst not affecting endosomal compartments, the cytoskeleton, internal cell pH or ATP levels. AG1478 reversibly disrupts the Golgi;<br>Does not work in rodent cells. | Human H4. | (Pan et al. 2008) | Sigma £117/5 mg<br>LCLabs £53/25 mg<br>Adooq £36/5mg |
| **Secramine B** | Cdc42 via RhoGDI1;<br>Golgi-PM transport | Inhibits Golgi to PM trafficking. Stops Cdc42's ability to bind the Golgi membrane and thus function correctly;<br>May impact upon salt transport. | Human T84;<br>cAMP-stimulated chloride secretion. | (Pelish et al. 2006) | Not available |
| **Dispergo** | Prevents access to ERES;<br>ER-Golgi | Loss of ER cisternae and extensive ER tubulation, preventing ER-Golgi transport which results in Golgi breakdown. | HeLa, CHO;<br>VSVG-GFP. | (Lu et al. 2013) | Not available |
| **LM11** | Arf1-GDP;<br>ER-Golgi, intra-Golgi | Targets the interaction between Arf1-GDP and its GEF ARNO. Causes Golgi and endosome dispersal | HeLa, MDCK. | (Viaud et al. 2007) | Not available |
| **AMF-26** | Arf-GEF Sec7 domain;<br>ER-Golgi, intra-Golgi | Induces Golgi disruption by targeting the interaction between Arf1-GDP and its GEF ARNO, binding the Sec7 domain;<br>Induces apoptosis, cell growth inhibitor. | Human cancer lines, HEK293T. | (Ohashi et al. 2012) | Abmole |

## 5.3. Testing of Secretory-Blocking Compounds in CHO cells

The compounds selected in section 5.2.3/Table 5.1 (BFA, Exo1, Exo2, FLI-06) were tested to determine their effect upon CHO cell growth and morphology. BFA was also tested both as a control and a potential evolutionary agent. Whilst it targets some regions of the endocytic pathway, its inhibition of ER-Golgi transport via blocking of COPII component recycling was likely to make it a potentially useful evolutionary agent. Secretory blocking compounds were produced as described in section 2.3.1 (Table 2.2), with lethal dose (LD) values for compounds determined (where required) as in section 2.3.2

### 5.3.1. Assaying effect of secretory blocking compounds on CHO cells

Kill curves of the selected secretory blocking compounds were carried out (as according to section 2.3.2) on parental MedI-CHO cells to test of the compounds' impact upon cell function, specifically looking at their impact upon cell growth and morphology. Where available, functional ranges for each of the compounds were taken from literature to inform the range of the kill curve. Although these values were not necessarily from related cell lines, mammalian cell line values were used where possible. Where previous data was not available, an initial kill curve at a wide concentration range (up to 1 mM) was carried out, informing proceeding kill curves resulting in a more accurate functional range. Cells were treated with each compound with two to three biological replicates at each concentration. Untreated and vehicle-only (at the volume of the highest concentration used) controls were used, with PB (static/shaken plate culture) or ViCell (larger shaken cultures) used to determine cell growth.

#### *5.3.2.1. Effect of Exo2 on CHO growth and morphology.*

Literature suggested that Exo2 impacts upon cell function and Golgi structure, through Arf inhibition, at a concentration of 50 μM. At this concentration Exo2 was shown to instigate Golgi collapse into the ER with little effect upon the endocytic pathway in monkey BSC1 cells (Feng et al. 2004). In HeLa cells, the same concentration resulted in statistically significant Arf inhibition, which was not seen at 10 μM (Boal et al. 2010; Spooner et al. 2008), whilst in *Arabidopsis* plant cells a higher concentration (200 μM) of both Exo2 and its derivatives did not result in ER or Golgi morphological changes (Sorieul et al. 2011). An initial Exo2 concentration range between 10-300 μM was used to determine Exo2's effect upon CHO cells, analysed by cell growth (static) and ER/Golgi morphology.

An initial kill curve of MedI-CHO cells between 10-100 μM showed Exo2 had little impact upon CHO cell growth at 100 μM Exo2 (data not shown). Exo2 concentration was increased, with a kill curve between 60 - 300 μM performed,

showing that, compared to controls, Exo2 had no effect upon CHO cell growth up to a concentration of 300 μM (figure 5.2 a).

The morphology of the Golgi (the organelle targeted by Exo2) was investigated by staining with antibodies against the Golgi markers Stx5 and GM130 (staining of AP1 was also performed but did not work as well), using the staining protocol described in section 2.7.1/Table 2.3. MedI-CHO cells were treated with 200 μM Exo2, DMSO at the same volume (negative control) and 10 μM BFA (positive control) at room temperature for 30 mins before being fixed, permeabilised and stained with anti-Stx5 and –GM130 primary antibodies and relevant secondary antibodies before imaging by microscope (figure. 5.2 b). Each drug treatment was repeated twice to provide biological duplicates, with three images of each slide taken to provide technical replicates.

In negative control DMSO-treated cells, staining of both Stx5 and GM130 matches the Golgi-staining pattern that was expected with one tight bundle of cisternae per cell being visible with both Stx5 and GM130 staining (although GM130 staining appears clearer and more precise than that of Stx5). Positive control treatment of cells with 10 μM BFA shows a dispersal of both Golgi markers, showing the Golgi is no longer maintaining its structure and suggesting impairment of its function. In cells treated with 200 μM Exo2, both GM130 and Stx5 staining show a slightly less well defined Golgi clump than that seen in the negative control, however the Golgi is still visible and there is no dispersal as is seen in the BFA-treated positive control.

With Exo2 at a concentration of 50 μM shown to have an upon cell growth and Golgi morphology in other mammalian cell lines (Boal et al. 2010; Feng et al. 2004; Spooner et al. 2008), given the lack of effect it has upon MedI-CHO cell growth and Golgi morphology at concentrations up to 300 μM, it can be considered that Exo2's Arf-inhibitory function is not replicated within CHO cells. That treatment of MedI-CHO with 200 μM Exo2 does not result in an obvious dispersal of the Golgi, as is seen in BFA-treated cells, shows that Exo2 does not alter Golgi morphology and structure as seen in other mammalian cells.

As such, it can be deduced that Exo2 is not effective at disrupting the Golgi within MedI-CHO in the same way as BFA and it therefore it not a suitable evolutionary agent with which to treat CHO cells. This may be due to a difference in the Sec7 binding pocket (that BFA binds) within CHO cells compared to human/monkey-derived cell lines. High concentrations of Exo2 were shown to not impact upon plant cell morphology and this may be for a similar reason (Sorieul et al. 2011). It was decided, having used Exo2 concentrations 6-fold higher than that seen to have a definitive effect upon CHO cell growth and morphology, that higher concentrations would not have an effect upon CHO cells. Also, pushing the

concentration much higher would have risked exceeding the suggested maximum levels of DMSO within CHO cell culture (0.5%).

#### *5.3.2.2. Effect of Exo1 on CHO growth and morphology*

There was less literature upon which to base Exo1 experimentation when compared to that for Exo2. Literature suggested that Exo1 inhibited mammalian cell function through inhibition of exocytosis with an $IC_{50}$ of approximately 20 μM, with 100 μM Exo1 completely halting ER-Golgi transport of VSVG-GFP in BSC1 monkey cells (Feng et al. 2003). As such, an initial kill curve between 20 - 100 μM Exo1 was set up to determine whether it has an impact upon MedI-CHO growth and morphology. The kill curve was carried out as previously described above for the Exo2 kill curve.

Increase in the concentration of Exo1 did not appear to impact upon growth of MedI-CHO. There is a slight decrease in growth at 40-60 μM, but at 100 μM growth recovers towards untreated levels and any impact upon growth seen is minimal, suggesting any decrease in growth is not due to the presence of Exo1 (figure 5.3). Staining of Golgi markers in Exo1-treated cells was inconclusive as to whether it had any impact upon Golgi morphology, but there was no obvious difference when compared to the negative controls (data not shown).

Due to the lack of effect Exo1 treatment has upon CHO cell growth at concentrations previously shown to have an effect upon mammalian cell growth and morphology, combined with the knowledge that Exo1 targets the same region of the secretory pathway as Exo2 which also did not impact upon cell growth, it was decided that Exo1 is likely to not block the CHO secretory pathway and is therefore unsuitable for use in a directed evolution process.

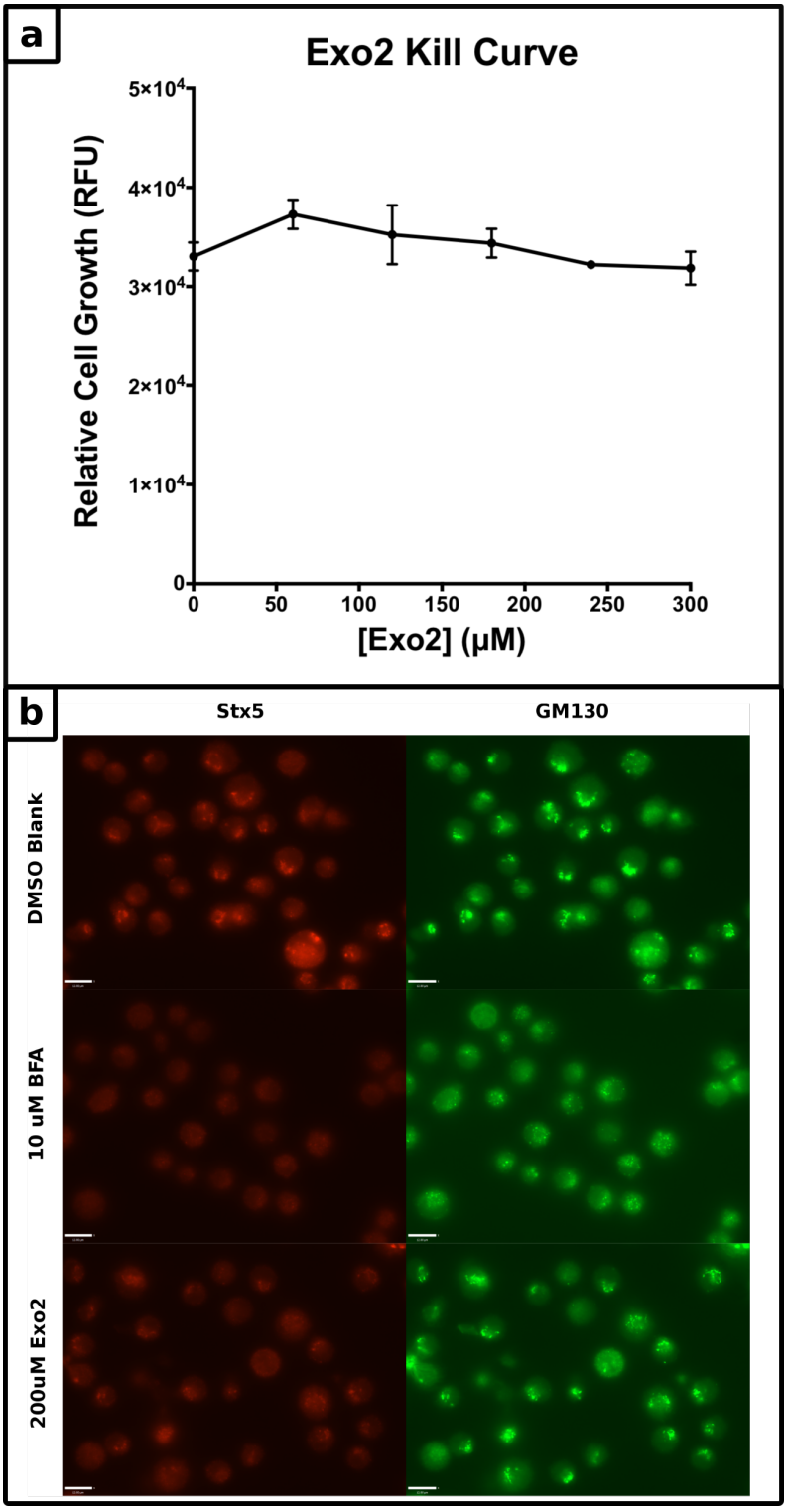


**Figure 5-2. Impact of Exo2 on CHO cell growth and morphology.**

*MedI-CHO was treated with the secretory blocking agent Exo2 at varying concentrations to determine its effect upon CHO cell growth and morphology. [a] Exo2 kill curve between 0-300 μM shows Exo2 has no impact upon CHO cell growth compared to the untreated control. [b] Staining of Exo2- and BFA-treated cells with anti-GM130 and anti-Stx5 (both Golgi markers) antibodies showed that treatment with Exo2 does not result in any change in Golgi morphology. Kill curve data points show average of three biological replicates. Microscope images are indicative of 2-3 images (technical replicates) of 2 biological replicates. Scale bar = 11μm.*

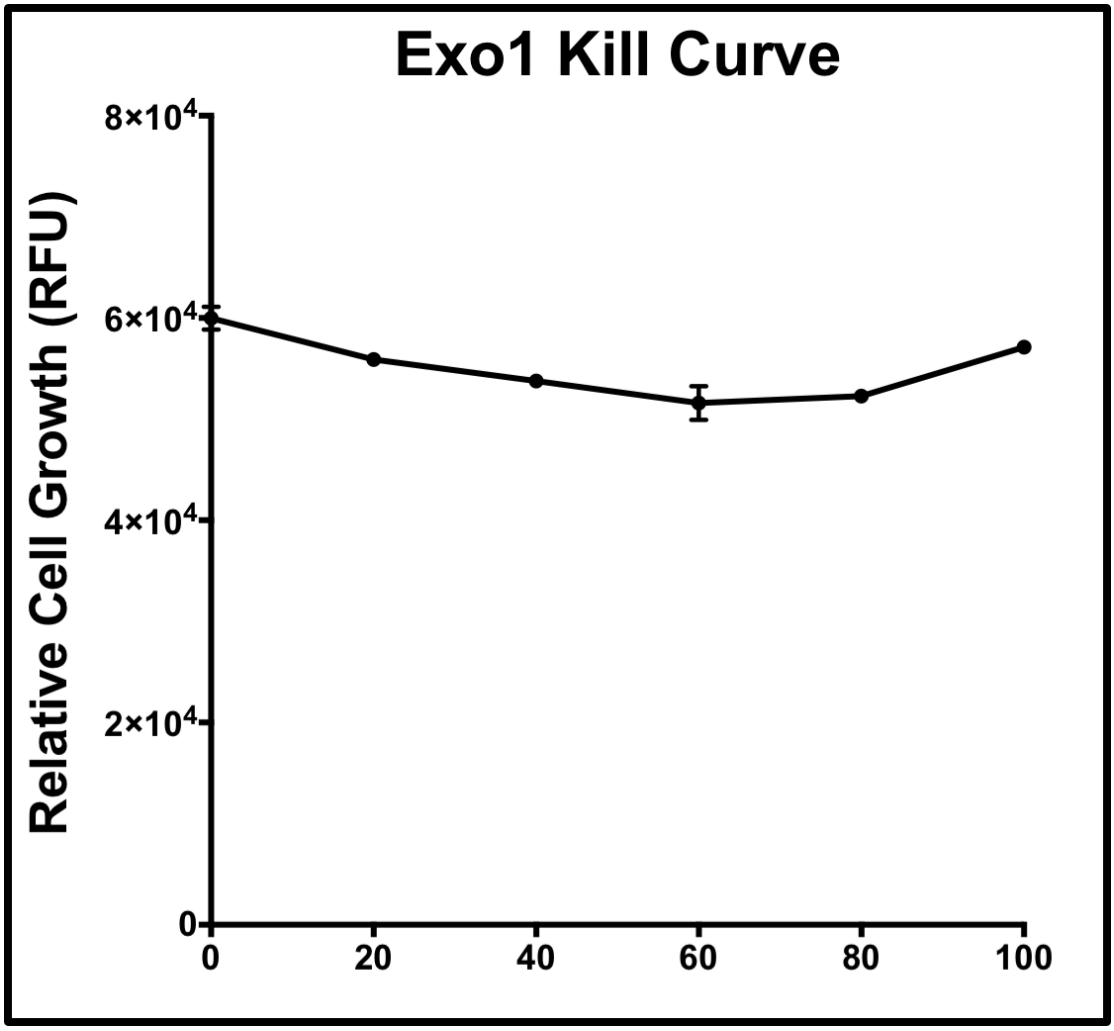


**Figure 5-3. Effect of Exo1 upon CHO cell growth.**

*MedI-CHO cells were seeded in 24-well static culture, treated with Exo1 at 20, 40, 60, 80 and 100 μM and grown for three days at which point cell growth was measured. Whilst there is a slight decrease in cell growth compared to the untreated control between 20-80 μM, this is not as severe as the wipe-out of cell growth seen with BFA treatment (figure 5-5 a), suggesting that Exo1 has little impact upon CHO cell growth.*

#### *5.3.2.3. Effect of FLI-06 on CHO cell growth and morphology*

As with Exo1, there is much less literature information regarding FLI-06 compared to Exo2 or BFA. FLI-06 was shown to function at 10 μM in HeLa cells, informing an initial kill curve between 0 - 20 μM (Kraemer et al. 2013; Yonemura et al. 2016). Even at the lowest concentration of 2.5 μM, cell growth was completely halted (data not shown) so proceeding kill curves were performed in shaken culture with a maximum concentration of 0.3 μM. Each FLI-06 concentration was tested with 2 biological replicates.

IVCD, cell viability and growth rate measurements show that FLI-06 has minimal effect upon CHO cell growth up to a concentration of 0.2 μM, but above this level even a small increase has a measureable effect, with cell growth being severely decreased at 0.25 μM and above (figure 5-4 a). This resulted in a sigmoidal kill curve and a small range of function for FLI-06 in CHO cells. To determine the LD values of FLI-06, the central section of the kill curve (in which CHO cells reacted to changing concentrations of FLI-06) was used. The change in IVCD over the course of the kill curve was used to determine LD values, as this produced the most linear reading of all the growth metrics measured. IVCD change was converted to a fold change compared to the growth of the untreated control and linear regression performed, with the formula produced being used to convert growth (as a percentage of the control's growth) to a LD value (figure 5-4, b; table 5-2).

Use of the central section of the kill curve resulted in a linear regression with an $r^2$ of 0.93, a value satisfactory for calculation of LD values. Use of all data points for linear regression resulted in an $r^2$ of 0.89 (data not shown). Growth curves of MedI-CHO in shaken culture showed that whilst addition of 0.25 μM FLI-06 at the start of culture completely halts cell growth, addition on day three of culture does not impact upon on-going cell growth of an established culture when compared to untreated cells (data not shown).

To determine the effect FLI-06 had upon CHO cell morphology, MedI-CHO cells were treated with 0.1 μM FLI-06 for 30 min before being fixed, permeabilised and stained with antibodies against GM130 and Calreticulin (ER marker) before imaging. Images show that FLI-06 treatment had little impact upon Golgi morphology but that ER morphology does appear to change upon FLI-06 treatment when compared to untreated control cells. Even at the relatively low concentration of 0.1 μM, the intensity of Calreticulin staining is greatly reduced, with the reticular staining pattern seen in the control cells not seen after FLI-06 treatment (figure 5-4 c).

The combination of kill curves and microscopy show that FLI-06 has an impact upon CHO cell growth as well as the morphology of the secretory pathway. FLI-06's functionality in CHO cells, corroborated by information from the literature

showing it functions to block cargo loading into ERES, suggest that FLI-06 is a suitable compound with which to treat CHO cells with a view to evolving a cell line with an enhanced secretory pathway (Yonemura et al. 2016).

#### *5.3.2.4. Effect of Brefeldin A on CHO cell growth and morphology*

From literature, approximately 0.18 μM BFA was known to reduce secretion capacity to below 20% of normal levels in COS-7 (monkey) cells, with 1.8 μM BFA having a similar effect in rat kidney cell secretion, as well as scattering β-COP from the Golgi in kidney and CHO cells (Torii et al. 1995). Knowing that CHO cells are susceptible to BFA treatment, a three-day kill curve with BFA concentrations up to 1.5 μM was performed to assay MedI-CHO's BFA susceptibility in shaken culture.

Growth rate (not shown), change in IVCD over the course of cell growth and viability measurements show that BFA has a large impact on CHO cell growth between 1 and 1.5 μM (figure 5-5 a). Fold change in IVCD values compared to the untreated cells were calculated and plotted so that linear regression could be carried out, the resulting line of best fit allowing LD values to be calculated. With the dose-response of MedI-CHO being quite linear between untreated and 1.5 μM BFA-treated cells, no data points had to be removed before linear regression was carried out so BFA LD values could be calculated as described previously (figure 5-5 b). Linear regression produced an $r^2$ value of 0.95, showing the linear relationship between BFA concentration (between 0 and 1.5 μM) and cell growth, whilst also showing that the linear regression was accurate enough to base LD values on (table 5-2).

MedI-CHO cells were treated with 1.0 μM BFA for 30 minutes at 37 °C before being fixed, permeabilised and stained with antibodies against GM130 and the BFA target molecule GBF1. Imaging shows that when treated with BFA the Golgi structure within CHO cells, is dispersed, with both GM130 and GBF1 being more scattered throughout the cell when compared to the tight Golgi structure seen in the untreated control cells (figure 5-5 c). That GBF1 is no longer Golgi localised suggests that COPI vesicle formation at the membranes of Golgi cisternae can not occur upon BFA treatment. This data replicates that seen previously in CHO-K1 cells (Torii et al. 1995).

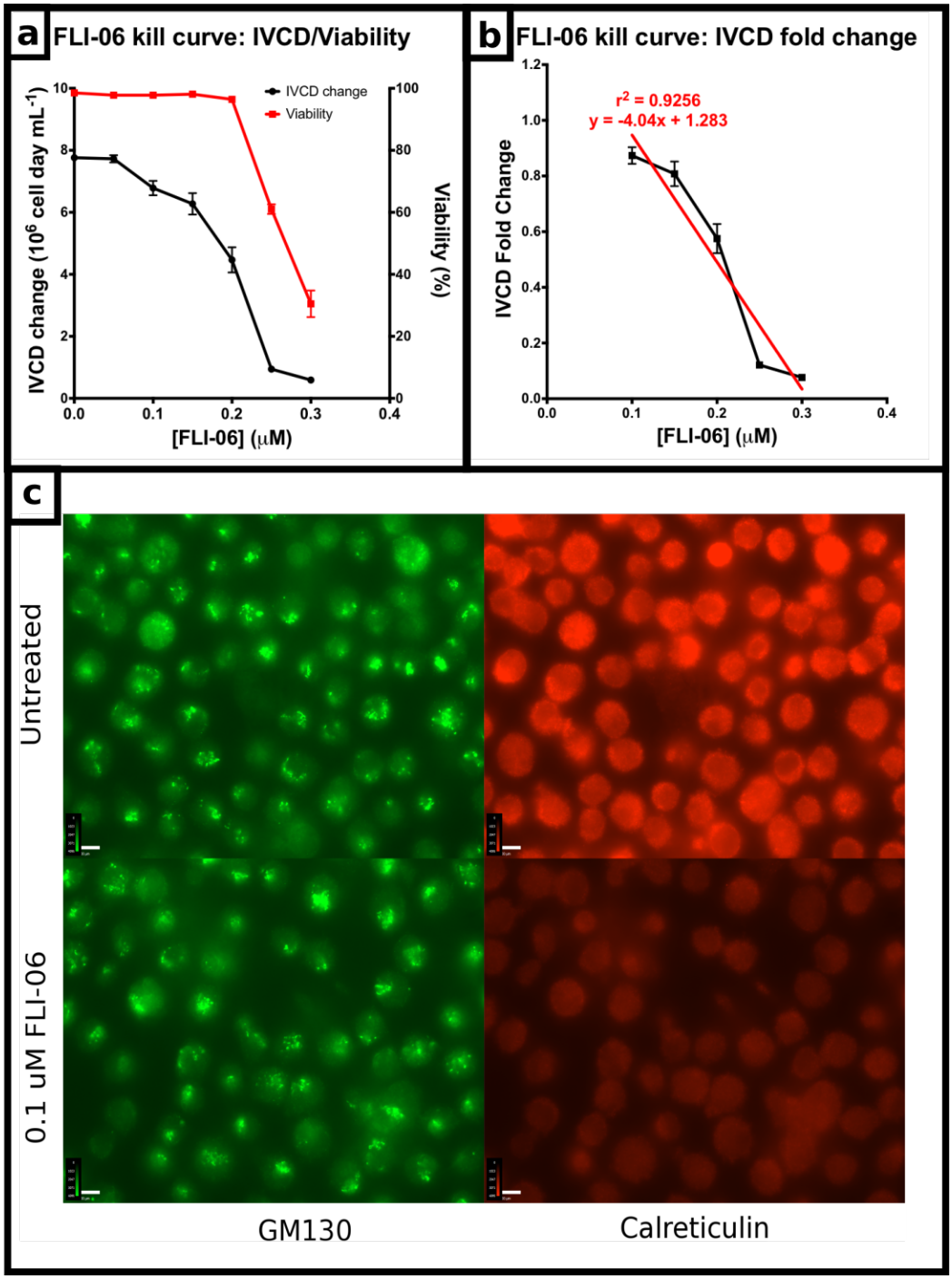


**Figure 5-4. FLI-06 impact on CHO growth and morphology.**

*MedI-CHO cells were treated with FLI-06 at varying concentrations to assess impact upon cell growth and morphology. [a] Kill curves show the impact FLI-06 has upon CHO cell growth and viability. [b]. Linear regression analysis of IVCD change kill curve fold change was used to determine FLI-06 LD levels. [c] MedI-CHO cells treated with FLI-06 were stained with anti-GM130 and anti-Calreticulin antibodies to better determine morphological effect of FLI-06. Kill curve data points show the mean average of three biological replicates, error bars show SEM. Each image is indicative of 2/3 images taken from 2 biological replicates. Scale bars: 11.00 μm.*

These growth and morphology data, alongside prior knowledge from literature, suggest that using BFA to direct evolution of the CHO secretory pathway should result in changes within the intracellular transport pathway in CHO cells. Whilst BFA targets are not limited to being involved in anterograde secretory transport (specifically the BIG1/2 Arf-GEFs involved in endocytosis), off-target effects should not impact too heavily upon cell growth and other regions of the biosynthetic pathway. As BFA is the most well researched of the previously listed secretory blocking agents (table 5-1), it will also provide a positive control with which to compare the effects of directed evolution against the other secretory blockers to be used.

**Table 5-2 – Lethal Dose concentrations of FLI-06 and BFA calculated from kill curves with MedI-CHO.**

| IVCD FC value | LD | [FLI-06] (µM) | [BFA] (µM) |
|---|---|---|---|
| 0.75 | 25 | 0.132 | 0.478 |
| 0.50 | 50 | 0.194 | 0.850 |
| 0.25 | 75 | 0.256 | 1.222 |
| 0.00 | 100 | 0.318 | 1.594 |

*The lethal dose (LD) in MedI-CHO cells for FLI-06 and BFA is shown. LD values were determined as the drug concentration at which 25, 50 and 75% of the cell population is killed. IVCD was used as a measure of this effect upon CHO growth. IVCD levels of the kill curve were compared to that of untreated control cells. If a concentration had a 25% effect on growth (LD25), it resulted in an IVCD that was 75% that of the control. Similarly an LD75 corresponded to a 75% reduction in IVCD compared to the control and an LD50 corresponded to a 50% reduction in IVCD compared to the control.*

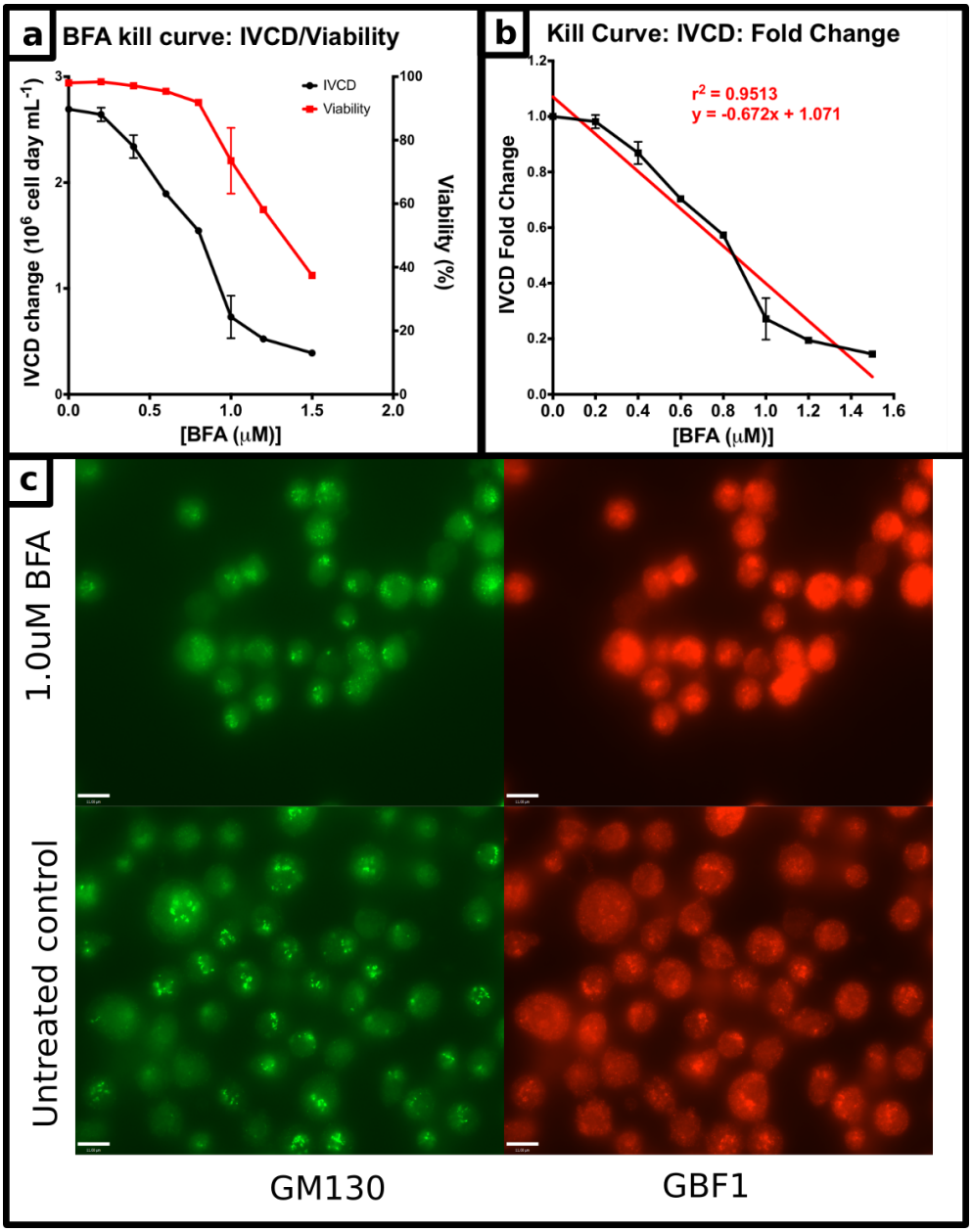


**Figure 5-5. Effect of BFA upon CHO cell growth and morphology.**

*MedI-CHO cells were treated with BFA at varying concentrations to assess impact upon cell growth and morphology. [a] Kill curve measuring BFA effect on cell growth and viability at increasing concentrations [b] Linear regression analysis of IVCD change fold change from untreated cells to determine BFA LD levels. [c] Golgi staining of BFA treated MedI-CHO cells with GM130 and GBF1 antibodies shows BFA treatment results in Golgi dispersal. Each data point shows mean average of biological duplicates, error bars show SEM. Microscopy images shown are representative of two separate images of three biological replicates. Scale bars: 11.00 μm.*

## 5.4. Directed evolution of CHO secretory pathway against secretory blocking compounds

Parental MedI-CHO cells (4 passages out of vial) were grown in the presence of the selected secretory blocking compounds. Cells were grown in 10 mL culture volumes in Cultiflasks (section 2.1.2.1). Concentration of compounds was determined from the initial kill curve data shown above. Cells were initially treated with a low concentration of the compound until growth rate and viability levels recovered to parental levels. After stabilisation at this concentration, compound concentration was increased with this process repeated until cell growth had recovered in the presence of 1.0 μM of the evolutionary compound (as described in figure 5-1). On occasions where cell growth was greatly impeded by compound presence, passage length was extended to 5-7 days and cell seeding density increased to 0.4 x $10^6$ cells $mL^{-1}$ where necessary. Post-evolution cell banks were produced which were then transfected (transient and stable) with model Mabs to ascertain the productivity capacity of the evolved cell lines. Cells staining and Western blotting was carried out to better define any morphological changes within the cell so as to better determine any biological changes the evolutionary process had upon the cell population.

### 5.4.1. Directed evolution strategy

The parental MedI-CHO cell line was used as a basis for all evolutionary strategies so that cell line productivity could be tested with different model Mabs. Evolution of an stably producing cell line would not allow this flexibility. Furthermore, as stable cell lines are selected (either clonally or through selection/amplification process) to be high producers, it can be considered that any stably producing cell line would already have a highly active biosynthetic and secretory pathway and as such would have little headroom in which to improve.

### 5.4.2. Effect of BFA evolution and selection on CHO cell productivity

Evolution against BFA was carried out by MedImmune (Granta Park, Cambridge). Over the course of 50 days, MedI-CHO cells were treated with increasing concentrations of BFA to a point where growth rate and viability had returned to normal parental levels whilst in the presence of 1.0 μM BFA, resulting in the MedI-BA BFA-evolved cell line. MCBs from one evolved strain were provided for testing in Sheffield, from which a WCB was produced after growth without BFA selective pressure for 3-4 passages. MedI-CHO cells was used as a control cell line. The effect of BFA evolution upon the productivity, growth and morphology of MedI-CHO cells was investigated to ascertain the viability of directed evolution as a method for improving cell factory productivity. All productivity and growth data shows values of MedI-BA cells as a fold change of the MedI-CHO values.

#### *5.4.2.1. Transient Mab expression of a BFA selected cell line*

MedI-CHO and MedI-BA were revived from WCB and grown for 4 passages, MedI-BA in the presence of 1 μM BFA after passage 1. Transient transfection (nucleofection) of both cell lines with GFP (pMAX, Lonza) was performed and cells grown for 4 days under standard cultiflask conditions (i.e. no BFA present in MedI-BA cells). Cell samples (1 mL) were taken for analysis by flow cytometry, with non-transfected cells used as a negative gating control. Analysis showed there was little difference in transfection efficiency between the two cell lines (both above 90%; data not shown), showing that the evolutionary process has not affect the cell lines' ability to be transfected.

To determine productivity levels of MedI-BA compared to MedI-CHO, both cell lines were transiently transfected (via nucleofection) with plasmid DNA encoding ETE and DTE model Mabs (section 2.5.1) before growth in 10 mL standard cultiflask culture for four days. Growth data was collected and supernatant samples taken for titre analysis, performed with the Valita™TITER HS assay plate (section 2.6). Each cell line was separately transfected three times to provide three biological replicates, with three supernatant samples from each cell line analysed in triplicate. Titre and Qp values show that, for both ETE and DTE Mabs, MedI-BA has significantly lower titre and specific productivity levels than MedI-CHO (figure 5-6 a, b, d, e), despite showing slightly improved growth (figure 5-6 c, f). When transiently expressing the ETE Mab, MedI-CHO line out-perform MedI-BA by 1.25- (overall titre) and 1.4-fold (Qp). This pattern is replicated with the DTE Mab, where MedI-CHO is 1.61- (overall titre) and 2.0-fold (Qp) more productive than MedI-BA.

When expressing both ETE and DTE Mabs, the growth rate of MedI-BA was significantly higher than that of MedI-CHO by 1.18-fold, with no difference in cell viability. It is likely that this is due to the removal of BFA from the MedI-BA culture conditions, which was seen to result in a slight increase in growth rate during normal cell culture as, despite MedI-BA being fully evolved to survive in the presence of BFA, the presence of BFA still has a small inhibitory effect upon growth compared to BFA-free culture.

#### *5.4.2.2. Stable transfection of BFA evolved cell line*

Transient expression of a recombinant Mab is a good indicator as to a cell line's production characteristics when stably expressing the same recombinant gene (Daramola et al. 2014). However, analysing productivity within a stable pool allows more robust conclusions to be made regarding a cell line's performance capabilities. The selection and amplification process required in the production of stable pools ensures all cells within the population are expressing the recombinant product and therefore the population is more likely to be

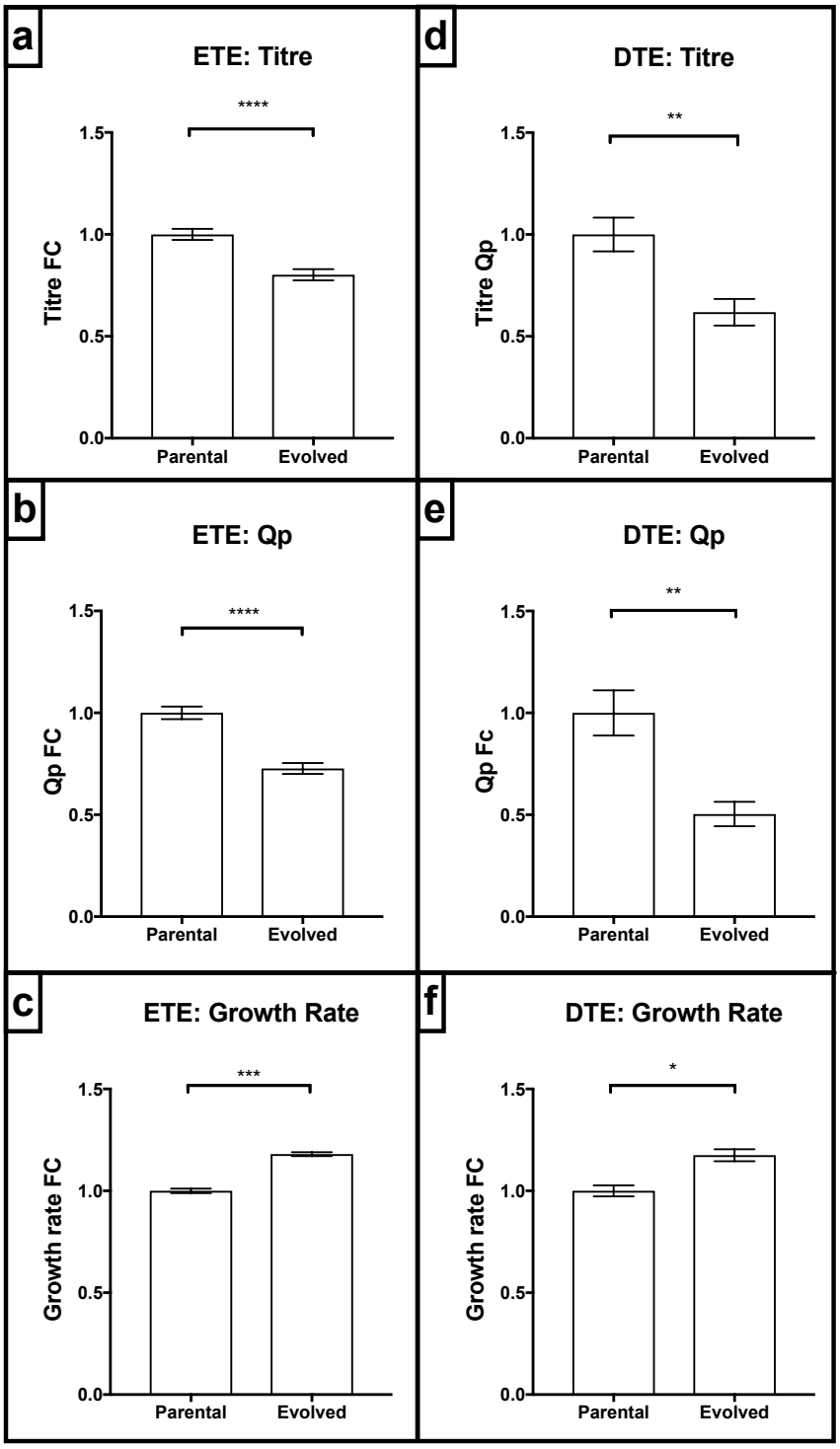


**Figure 5-6. Transient productivity of BFA-evolved CHO cells.**

*MedI-BA and MedI-CHO were transiently transfected with plasmid DNA encoding ETE and DTE Mabs. After 4 days' growth supernatant and cell growth samples were taken and titre determined by Valita™TITER HS assay. [a, b; d, e] Productivity of the parental cell line was greater than that of the evolved cell line when comparing both titre and Qp of both DTE and ETE Mabs. [c, f] This is despite MedI-BA growing quicker than MedI-CHO post-transfection. Titre and Qp data shows the mean of three technical replicates (repeats of Valita™TITER analysis) of each of the three biological replicates (3 separate transfections; n = 9). Growth data shows the mean of three biological replicates. Error bars show SEM. All charts show fold change of the evolved cell line value compared to the mean of the parental cell lines.*

normally distributed with regards to Mab production. The higher level of production seen in stable pools allows smaller differences to be magnified. That the cells are always producing the recombinant gene also allows long-term culture (both batch or fed-batch) to be carried out to better understand the productivity profile of a cell over an entire culture cycle. The use of a stable pool, as opposed to a clonally-derived population, allows for much quicker production of the stably-producing cell line, as well as allowing data on an entire population of cells to be collected, not just data from a stably-expressing cell line derived from the highest-producing single-cell clone of each pool. This would therefore allow a better measure of the directed evolutionary process's true effect.

#### *5.4.2.3. Production of stable pools from standard expression vector*

Stably-expressing pools were created from MedI-CHO and MedI-BA cell lines, expressing ETE and DTE Mabs from linearised standard plasmid DNA (that is, DNA encoding just the Mab and a GS selection marker; section 2.4). Cell lines were revived and passaged alongside each other for comparison purposes. Plasmid DNA was linearised by restriction digest with PVUI endonuclease and purified by phenol-chloroform extraction and ethanol precipitation (section 2.4.4). Linearised DNA was introduced in to cells by cuvette nucleofection (section 2.5.3). Transfected cells were resuspended in 40 mL warmed CD-CHO which was split across 8 x 5 mL cultures, each in a separate T25 flask which was incubated at 37 °C for 24 h, at which point MSX selective pressure was added to a final concentration of 50 μM in 6 mL CD-CHO. Cells were maintained in static culture until reaching confluency, at which point they were transferred to shaken culture in 30 mL volumes. After recovery of cell growth cell banks of the stably expressing lines were produced. Each combination of cell line and model Mab was replicated four times to provide biological replicates.

Stably-producing cell lines were revived in shaken culture in the presence of MSX and, for evolved cell lines, 1 μM BFA from the start of passage 2. Cell lines were passaged as normal, with samples taken for cell growth calculations to ensure cell lines were growing as expected and similarly to each other, such that they could be compared to each other. Supernatant samples for productivity analysis were taken when cultures reached 30 (±2) generations (G30) of shaken culture growth, a point at which a pool is considered stable, its population normally distributed. Samples were taken from cells in exponential growth phase (either day 3 or 4 of passage) so once growth data was accounted for productivity values could be compared directly. Once cells had reached 30 generations of growth out of the vial, cultures were left for 10 days so an overgrow study (cells cultured until all nutrients have been taken up and all cells have died) could be performed, with growth and productivity samples taken on day 10 of culture so productivities of pools' entire growth cycles could be compared. Supernatant samples were analysed by Valita™TITER Mab assay plate (section 2.6). Each sample was

repeated three times on the same Valita™TITER plate so as to provide three technical replicates of each biological replicate. Growth and specific productivity values were calculated using the titre values and growth data collected (for G30 samples only, as these cultures have only been in a growth phase whereas overgrow cultures have passed through exponential, stationary and death phases and so cell growth measurements can not be used to calculate productivity). All titre and productivity data was normalised to that of the MedI-CHO control cells to show fold-change in productivity levels compared to the parental line.

#### *5.4.2.3.1. Stable ETE expression levels in BFA-selected cell lines*

In both MedI-CHO and MedI-BA derived stable pools, cell growth rate and viability values recovered at a similar rate over the initial 10 generations in after culture. This shows there was no initial growth difference between the two cell lines (figure 5-7 a, b). However, after the selective pressure of BFA was returned to the MedI-BA pool (passage 5), both the growth rate and viability of MedI-BA were reduced when compared to that of MedI-CHO, with both growth rate and viability taking until approximately the 30$^{th}$ generation to recover to the level seen in MedI-CHO. Whilst MedI-BA has evolved to grow as normal in the presence of BFA, this drop in growth rate may be due to genetic drift that has occurred during the process of stable pool recovery.

The mean average titre from MedI-CHO G30 ETE stable pools (from cultures in exponential phase) was in the range of 25 mg/L. In MedI-BA cells, this dropped to around 10 mg/L, a 0.4-fold change in titre when compared to the parental cell line (figure 5-7 c). This result is replicated in the Qp data, with the parental line producing an average of 1.5 pg cell$^{-1}$ day$^{-1}$ compared to 0.5 pg cell$^{-1}$ day$^{-1}$ from the evolved cell line, a 0.31-fold change (figure 5-7 d). Finally, over a 10-day overgrow, MedI-CHO produced an average of approximately 320 mg/L of ETE Mab, compared to approximately 130 mg/L from MedI-BA, a 0.42-fold change in productivity (figure 5-7 e).

#### *5.4.2.3.2. Stable DTE expression levels in BFA-selected cell lines*

As in the ETE-expressing stable pools, the growth of pools derived from both MedI-CHO and MedI-BA recover at a similar rate to each other over the initial 10 generations of shaken culture (figure 5-8 a, b), with growth rate and viability levels of MedI-BA again reducing slightly upon re-introduction of BFA selective pressure at passage 5. This disparity disappears within approximately 7 generations, whereas in ETE-expressing pools closer to 15 generations were required to close this gap. After 30 generations in shaken culture, the mean growth-phase titre of MedI-CHO reaches 26.9 mg/L, whilst MedI-BA averages 6.2 mg/L, a 0.27-fold change in titre compared to that of the parental cell line (figure 5-8 c).

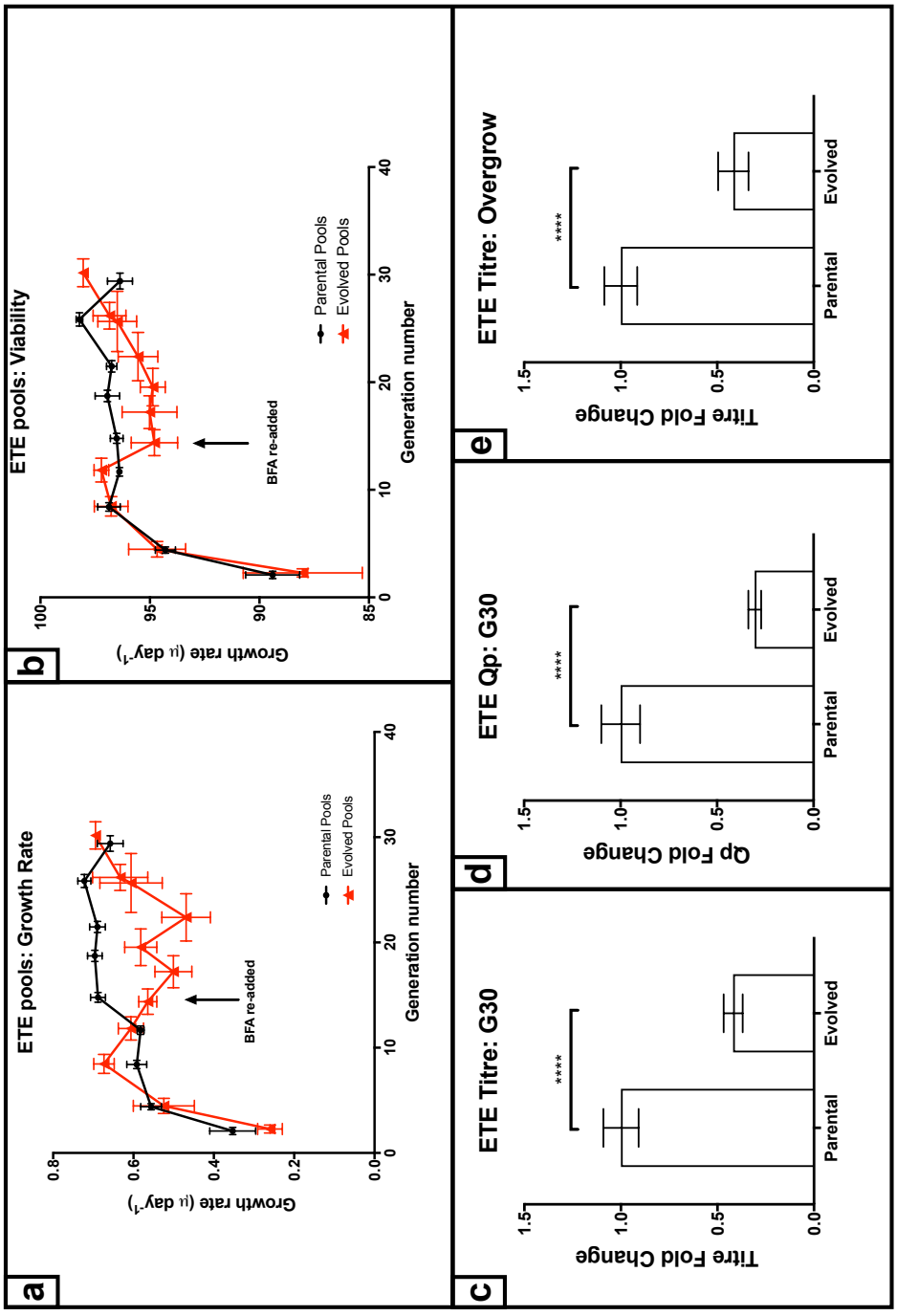


**Figure 5-7. Stable Productivity of an ETE Mab and growth of BFA-evolved CHO cells.**

*BFA-evolved and non-evolved CHO cells were stably transfected with an ETE model Mab and cell growth recovered in shaken culture for 30 generations. [a, b] Stable pool growth was tracked to compare cell recovery between different pools. [c, d] Titre samples taken at G30 showed that MedI-CHO cell line productivity out-performed that of MedI-BA. [e] At G30 10-day batch-overgrow was performed, with MedI-CHO titres higher than those seen from MedI-BA. All titre values were from Valita™TITER analysis. Growth data points show mean average of 4 biological replicates. Productivity data points show mean of 3 technical replicates of 4 biological replicates (n = 12). All error bars show SEM.*

Due to MedI-BA slightly out-growing MedI-CHO, as cells reached 30 generations the 0.18 Qp fold-change of MedI-BA compared to MedI-CHO (0.23 compared to 1.23 pg $cell^{-1}$ $day^{-1}$) is greater than that seen from titre measurements only (figure 5-8 d). This pattern is repeated in 10-day overgrow cultures, where a mean MedI-BA yield of 11 mg/L reflects a 0.11-fold change compared MedI-CHO's mean yield of 104 mg/L (figure 5-8 e).

As expected there appears to be little difference in cell growth between stably-expressing pools derived from either parental or evolved lines, as evolved cells have recovered their growth ability in the presence of BFA selective pressure. However, due to the evolved cells having been exposed to, and developing resistance to, an ER stress inducing compound (aside from its secretory blocking properties), it could be expected that these cells would react better to stable expression of a recombinant Mab compared to the parental cell line. However, this is mitigated by the knowledge that, as MedI-CHO is an industrially derived cell line selected specifically for its ability to produce a wide range of model recombinant proteins, its cell growth when faced with the extra metabolic burden of Mab production is unlikely to be greatly impacted.

In both ETE and DTE-expressing pools, productivity data shows a statistically significant decrease in Mab production levels – both titre and specific productivity – in MedI-BA when compared to the MedI-CHO. However, this data does come with some caveats. The production of stable pools using standard plasmids allows for random plasmid insertion into the CHO cell genome. This insertion could be into a region with reduced genetic activity (e.g. a region of heterochromatin) where regions of the transgene can be modified (e.g. methylation of CMV promoter), reducing productivity (Yang et al. 2010).

Whilst the MSX/GS knockout-driven selection and amplification process should provide stable pools with the gene inserted in an active section of the genome, the lack of knowledge as to the activity of the recombinant gene results in unknown variability between the replicate pools. Whilst use of biological replicates should control for this to some extent, that the titres of biological replicates varied widely from each other suggests that Mab genes may well be being expressed at different levels across different pools, resulting in large error bars on the productivity graphs. As such, due to unstable and variable expression of the transgenes within the CHO cell, I can not be sure that in these stable pools we are actually testing the secretory capacity of the cell (as planned) and are in fact measuring a difference in transgene transcription levels across different pools.

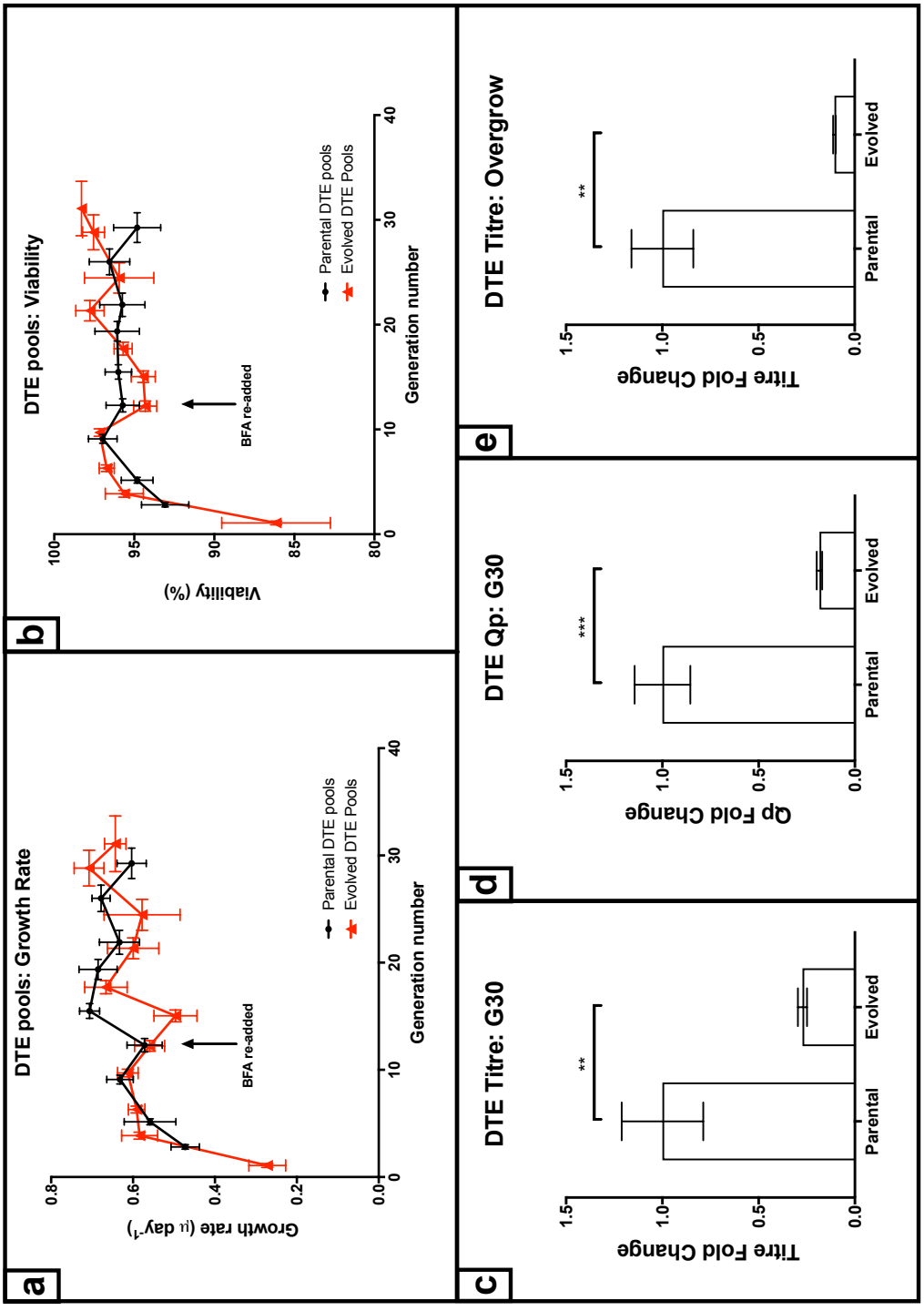


**Figure 5-8. Stable productivity of a DTE Mab and growth of BFA-evolved CHO cells.**

*BFA-evolved and non-evolved CHO cells were stably transfected with a DTE model Mab and cell growth recovered in shaken culture. [a, b] Stable pool growth was tracked to compare recovery across different pools. [c, d] Titre samples taken at G30 showed that MedI-CHO out-performed MedI-BA in terms of both titre and Qp. [e] At G30 a 10-day batch-overgrow showed that over the course of a full culture life cycle, MedI-CHO cells out-produce MedI-BA. All growth data points show the mean value of 4 biological replicates. Productivity data points show the mean of 3 technical replicates of 4 biological replicates (n = 12). All error bars show SEM.*

#### *5.4.2.4. Production of stable pools with comparable transgene transcriptional levels*

To try and better control for the activity level of the recombinant gene within different CHO cell lines it needs to be known that the recombinant gene is stably inserted into an area of genome for which the gene's activity can be assumed to be at the same level across all replicates and cell lines. Whilst qPCR could be used to assay the transcriptional activity of the recombinant gene across all pools, using this information would only normalise the data for transcriptional activity, with no consideration made for other factors (e.g. translational, post-translational) that would impact the Mab productivity levels of a population pool. However, by producing stable pools in which the transcription level can be assumed to be very similar (if not the same) across all replicates and cell lines, all results are naturally controlled for in terms of transcriptional activity, allowing data to be analysed with much less doubt than retrospective normalising to results from an entirely separate experiment such as qPCR.

There are two main ways of ensuring recombinant genes are inserted into a region of the genome with the same/similar levels of transcription activity. Use of a site specific targeted integration (TI) system allows a recombinant gene to be inserted, using a 'landing pad', into a known and highly active region of the genome. Use of a TI also ensures only one copy of the gene is present within each cell, resulting in a more homogenous population (Lee et al. 2015). With a TI system not available for use, stable pools were produced using a plasmid containing a transcription enhancing element (TEE). A TEE ensures the stable transcriptional activity of the transgene upon its insertion into the genome, reducing transgene silencing and allowing for stable, consistent and high-level gene expression (Saunders et al. 2015). Plasmid DNA is still randomly integrated, with copy numbers varying between cells, but the transcription levels of each insert are the same across a population in which the transcript levels of the recombinant product can be assumed to be somewhat heterogeneous, but normally distributed.

MedI-CHO and MedI-BA cell lines were both transfected with two separate linearised plasmids, one encoding an ETE model Mab, the other a DTE model Mab, with both plasmids contained the aforementioned TEE alongside the Mab and GS genes. The model Mabs used were the same as those used to create previous stable pools. Stable pool production was carried out as described previously (sections 2.4.4; 2.5; 5.4.2.3). For each cell line/vector combination tested, three separate transfections were carried out to provide three biological replicates. Transfections with normal plasmids were also performed out to control for differences in productivity and growth levels produced by the different types of plasmid vector. Initial productivity data was taken at the end of passage one so as to screen all pools to decide which were to be taken forward. This data showed that at this early stage, TEE-containing stable pools had higher production levels

than those produced with normal vectors (figure 5-7, 5-8), with TEE ETE titres exceeding 200 mg/L compared to approximately 30 mg/L in normal plasmid pools (figure 5-9).

Use of a TEE also provides more reproducible data across biological replicates, with minimal difference seen across TEE pools compared to as much as a 5-10-fold difference between top and bottom producing non-TEE pools (data not shown). Whilst this is due in part to more of the TEE pools recovering fully, resulting in more data points being available with which to calculate spread about the mean, the titre and Qp values of biological replicates were much more reproducible in TEE pools compared to non-TEE pools, the latter often having one pool which far out-performed its replicates, likely an artifact of random genomic integration. Taking approximately two weeks, recovery of TEE pools to static culture confluency was quicker and more standardised across replicates than that of non-TEE pools, a trait again due to the open genetic conformation of the TEE-containing genome-inserted plasmid allowing better production of the GS selection gene, and therefore enhanced survival in MSX-containing growth media.

#### *5.4.2.4.1. Stable pool production levels.*

After pool recovery in shaken culture over 3-4 passages, TEE pools were banked. All pools were revived concurrently such that direct comparisons could be made across different test and biological replicates. MedI-BA-derived stable pools were supplemented with 1 μM BFA to maintain selective pressure. Pools were grown until 30 generations (± 2) were reached (taking 9-11 passages), such that Mab expression could be considered stable. By G30 the growth rate, viability and cell diameter of comparable pools (e.g. parental ETE vs. evolved ETE; parental DTE vs. evolved DTE) at the end of each 3-4 day passage had settled to comparable levels between cell lines and biological replicates (ETE: Figure 5-10 a, b, c; DTE: Figure 5-11 a, b, c). With cultures being the same age and growing similarly at G30 , a 12-day batch-overgrow was performed in 50 mL shaken cultures such that a direct comparison of growth and productivity of MedI-BA and MedI-CHOs could be made. For the productivity curves BFA selective pressure was removed from MedI-BA cell lines, whilst MSX selective pressure was maintained across all cultures. Cell growth and productivity samples were taken every 2 - 3 days throughout the 12-day period. Productivity was assayed by Valita™TITER Mab assay.

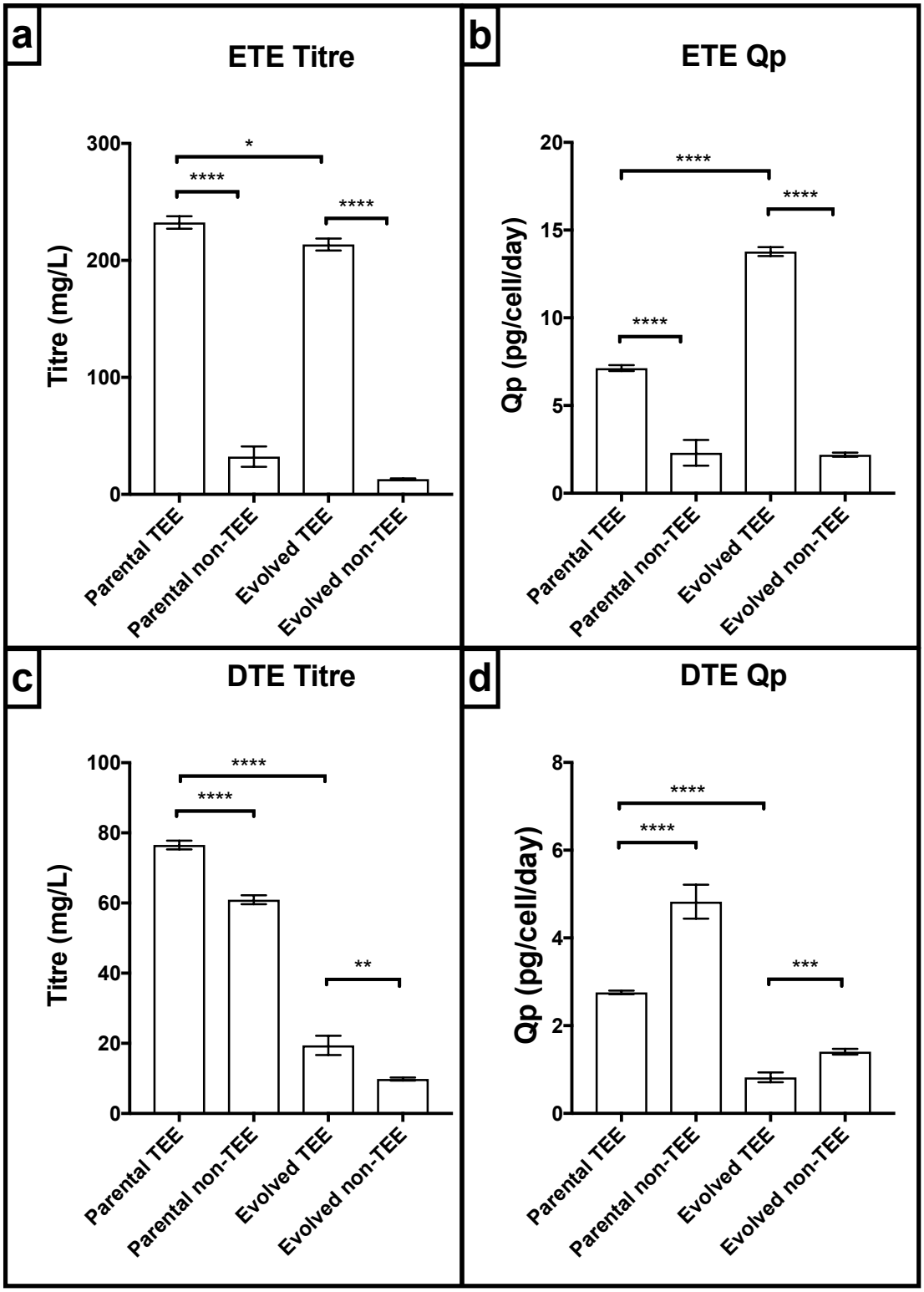


**Figure 5-9. Comparison of stable pools produced using TEE and non-TEE vectors.**

*MedI-CHO and MedI-BA were stably transfected with ETE and DTE Mab-encoding plasmids, with or without a TEE to better control transgene transcription levels. Data is taken from pool screening samples taken at the end of passage 1 of shaken culture. [a, c] Stable expression of both ETE and DTE from a TEE-containing plasmid resulted in consistently higher titres than from non-TEE containing plasmids. [b] Qp levels also follow this pattern in an ETE-expressing cells. [d] However, in DTE-expressing cells Qp was higher in cells transfected with non-TEE vectors than those with TEEs. Data points show the mean of 3 (TEE) or 2 (non-TEE) biological replicates. Error bars show SEM.*

During the first five days of culture growth in ETE-expressing cells the cumulative IVCD of both MedI-BA and MedI-CHO cell lines are comparative. However, once beyond this point the growth of the two cell lines diverge, with MedI-BA reaching a higher cumulative IVCD than MedI-CHO at the end of culture (figure 5.10 d). Likewise, MedI-BA maintains healthy viability levels (above 95%) to day 8 before the culture enters its death phase, whilst a small drop in MedI-CHO viability is seen as early as day 5 before the death phase sets in around days 7 and 8 of culture (figure 5-10 e). These differences in growth are likely due to the removal of the BFA selective pressure as, despite MedI-BA having adapted to grow in the presence of BFA, its removal will still decrease an some inhibition ( albeit small) that the cell population can exploit to grow quicker.

Despite this enhanced level of growth, the overall ETE titre level of MedI-BA is much less than that of MedI-CHO. MedI-BA reaches a peak titre of 650 mg/L on day 7, compared to MedI-CHO's 1085 mg/L at the same time-point, which increases to a maximum of 1167 mg/L on day 10 (figure 5-8 d). The difference in both growth and productivity results in MedI-BA cells having a large deficit in terms of Qp compared to MedI-CHO cells during exponential growth (data not shown).

As was seen in the ETE stable pools, in DTE-expressing pools the cumulative IVCD of both cell lines are the same to day five, at which point the MedI-BA cell line out-grew MedI-CHO, with cell again viability staying higher for slightly longer and little difference in cell diameter after 9 generations of shaken culture (figure 5-11 a, b, c). Comparing cumulative IVCD of both ETE- and DTE-expressing cell lines, the expression of a different Mab does not heavily impact upon the growth of either MedI-CHO (with a day 12 DTE cumulative IVCD of 40.6 x $10^6$ cell day $mL^{-1}$, compared to ETE cumulative IVCD of 43.5 x $10^6$ cell day $mL^{-1}$, $p$ = 0.48) or MedI-BA (with a 12-day DTE cumulative IVCD of 56.4 x $10^6$ cell day $mL^{-1}$, compared to an ETE cumulative IVCD of 55.0 x $10^6$ cell day $mL^{-1}$, $p$ = 0.39).

As with the ETE-expressing stable pools there is a decrease in overall titre levels in MedI-BA compared to MedI-CHO, with DTE expression levels appearing to be completely wiped out in MedI-BA (maximum level of 32.7 mg/L on day 10) compared to MedI-CHO (maximum 359.9 mg/L on day 10).

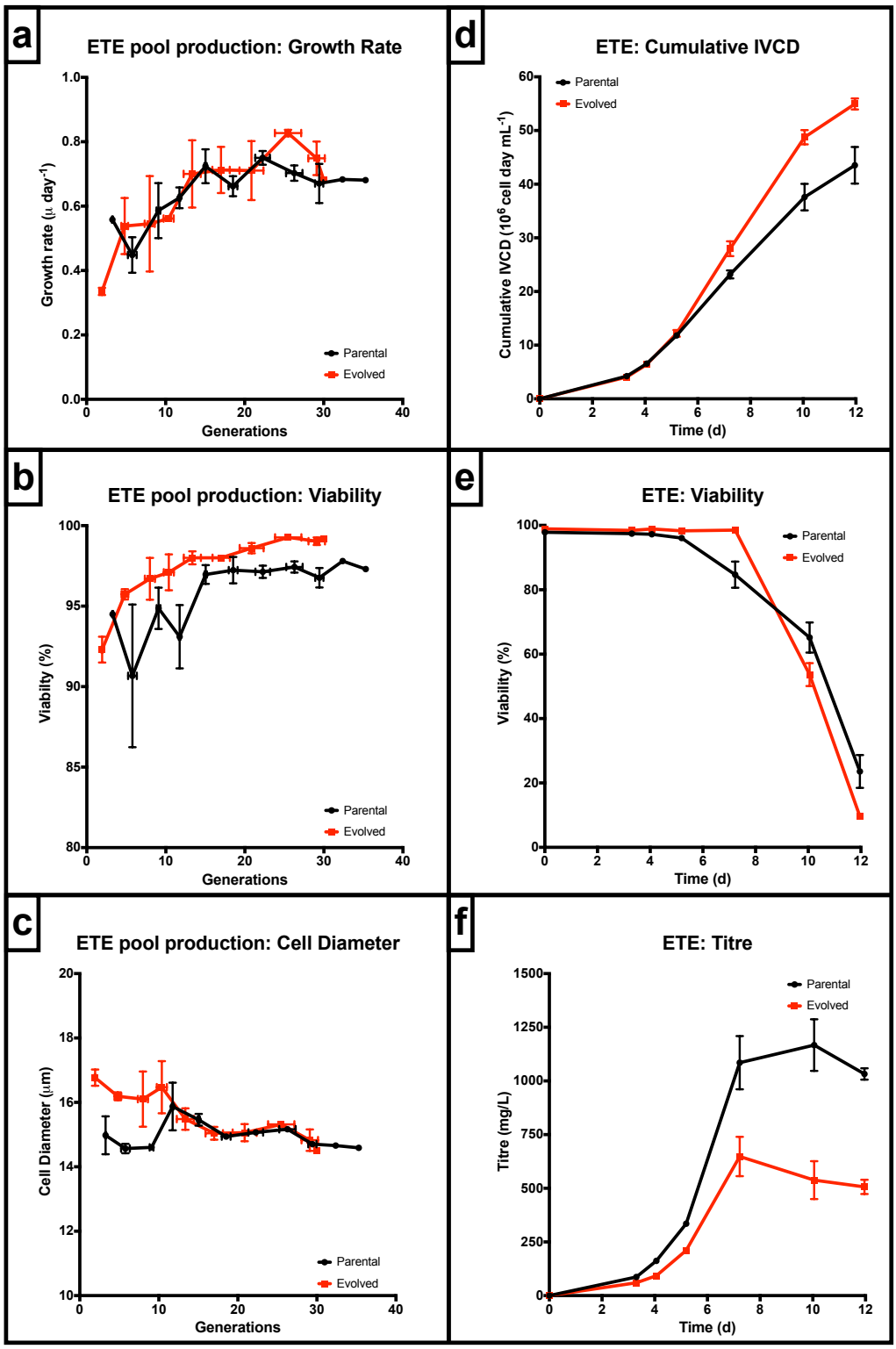


**Figure 5-10. TEE-driven ETE Mab production in BFA-evolved CHO cells.**

*MedI-CHO and MedI-BA pools stably expressing an ETE Mab alongside a TEE underwent a 12-day growth and productivity curve. [a, b, c] Pools were recovered from static culture for 30 generations in shaken culture until the growth characteristics were comparative. [d, e, f] 12-day batch overgrow cultures were set up so that the cumulative cell growth and productivity of each cell line could be compared over the course of an entire culture life cycle. All growth data points show the mean of three biological replicates. All titre data points show the average of three technical replicates of three biological replicates (n = 9). All error bars show SEM.*

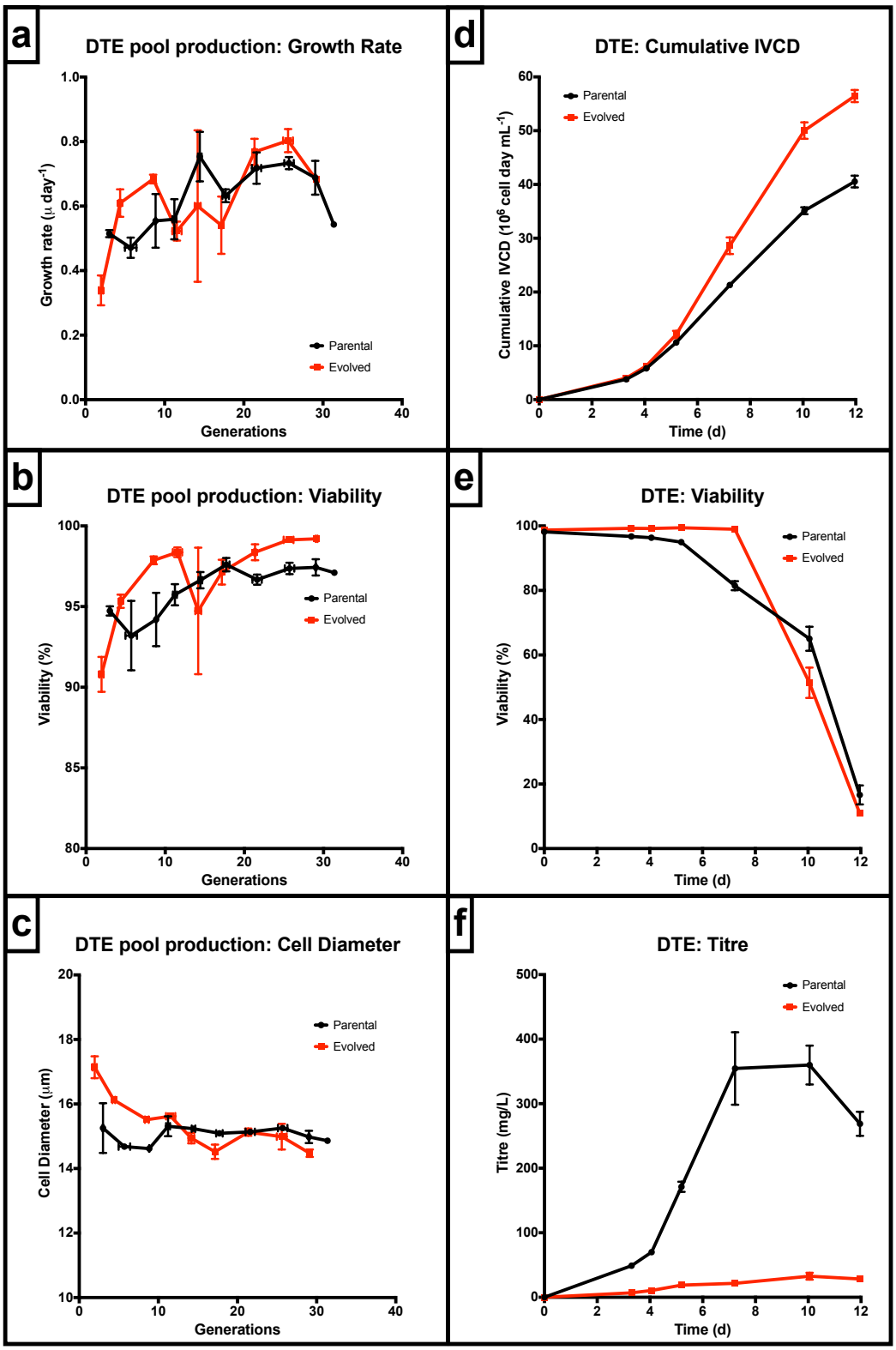


**Figure 5-11. TEE-driven Mab production in BFA-evolved CHO cells.**

*A model DTE Mab was stably expressed in MedI-CHO and MedI-BA. [a, b, c] Stable pools were recovered from static culture for 30 generations in shaken culture until their growth characteristics were comparative. [d, e, f] 12-day batch overgrow cultures were performed to compare the cumulative cell growth and productivity of both cell lines over the course of an entire culture life cycle. All titre data points show the average of three technical replicates of three biological replicates (n = 9). All error bars show SEM.*

#### *5.4.2.5. Homogeneity of Mab heavy and light chains in stable pools.*

Expression levels of the heavy and light chains of Mabs has been shown to affect the productivity levels of CHO cells, with light chain DNA normally being provided in excess of heavy chain DNA to aid with post-translational formation of the Mab molecule (Pybus et al. 2014). To ensure that any differences in stable pool productivity levels were not artifacts of Mab HC or LC levels, either due to lack of expression of the HC and/or LC genes or the obvious overexpression of one of the HC/LC genes compared to the other, stably-expressing pools were stained with anti-HC and anti-LC antibodies before analysis by flow cytometry to elucidate intracellular HC and LC levels across the pool. Staining also allowed the distribution of HC/LC expression across the entire population to be determined, highlighting whether Mab production across a pool was relatively homogenous (normally distributed), or whether the pool was not normally distributed, being composed of sub-populations with different productivity levels resulting in a heterogeneous pool, resulting in an inaccurate indication as to a cell line's actual productivity capacity.

TEE containing stably-expressing pools were harvested in exponential phase (day 4) and methanol fixed (section 2.7.1.2) before being stained with anti-HC and anti-LC fluorescently-tagged antibodies (section 2.7.2.1). Unstained and stained but non-transfected CHO cells were used as negative controls to ensure fluorescence reads were not due to natural cell fluorescence or non-specific antibody binding. Fluorescence levels of stained cells were analysed by flow cytometry. Cell debris and cell doublets were removed from the data by gating based on cells' SSC:FSC ratio. Data on cell fluorescence was collected from 20,000 cells within this gate. For both HC and LC staining negative controls were used to set up a gate for cells not containing any antibody, with 99% of the cells within this gate presumed to be not expressing recombinant Mab. All other cells with fluorescence levels above this gating threshold were assumed to be positive and containing intracellular HC and LC (figure 5-12, 5-13). The presence of HC was measured using (using a Red detection laser/filter combination), with LC measured (using a FITC/green laser/filter combination). Both TEE-containing and non-TEE pools were fixed and stained such that the HC and LC expression levels from the two different kinds of plasmid could be compared.

##### *5.4.2.5.1. Homogeneity of ETE Mab expression levels.*

Negative controls showed the background level of staining for non-stained cells. That non-transfected CHO cells stained with the anti-HC and anti–LC antibodies showed a similar low level of fluorescence shows that these tagged antibodies do not have any off-target effects and therefore any fluorescence seen is due to expression of the recombinant Mab (figure 5-12 a, 5-13 a). In all flow cytometry data the LC fluorescence data is shown on the top, with HC shown on the bottom.

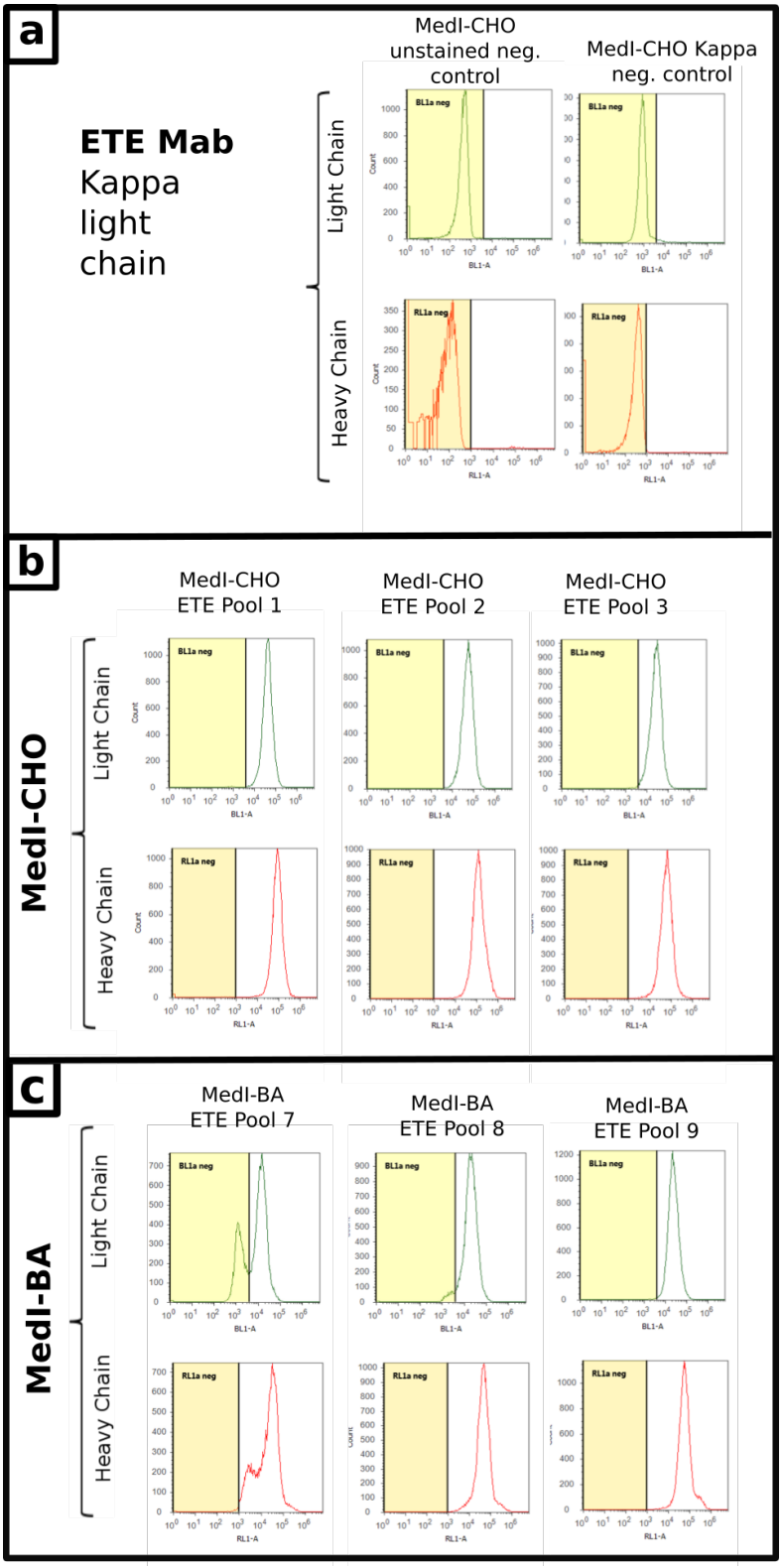


**Figure 5-12. Intracellular HC/LC abundance of ETE-expressing stable pools.**

*ETE-expressing stable pools (with a TEE) were MeOH fixed before staining with anti-human HC and anti human kappa LC antibodies. Stained cells were analysed by flow cytometry to elucidate intracellular HC and LC levels across the stable pool population. [a] negative controls confirmed that intracellular Mab HC and LC was present in [b] MedI-CHO and [c] MedI-BA stable pools. X axis show the level of fluorescence of each cell (proportional to HC/LC levels); y axis shows the number of cells at that level of fluorescence. Upper charts show LC abundance. Bottom charts show HC abundance.*

All three MedI-CHO ETE-expressing pools (figure 5-12, b) show a normal distribution of HC and LC expression across their entire population. The fluorescence levels and the number of cells at each fluorescence level are similar across all pools suggesting that the biological triplicates are quite close in terms of reproducibility. In the three MedI-BA ETE stably-expressing pools (figure 5-12 c), pool 7 shows a double peak for both HC and LC presence, with the subsidiary LC peak shown to be within the negative control range. LC distribution suggests a small amount of the pool 8 population are not LC positive, although the data for pool 8 HC abundance and pool 9 HC and LC abundance show a fairly normal distribution.

#### *5.4.2.5.1. Homogeneity of DTE Mab expression levels.*

All DTE-expressing stable pools show some level of non-normal distribution with regards to intracellular HC and LC content (figure 5-13 b, c). However, all distributions are positive for both Mab chains, showing that all cells are expressing Mab HC and LC, albeit that expression levels are not fully normally distributed. There is a small double-peak in the LC profile of pool 12 which is matched slightly by the concurrent HC fluorescence profile.

The slightly non-normal HC and LC distributions, alongside the slightly lower fluorescence levels seen in the MedI-BA ETE stable pools when compared to MedI-CHO cells may partly explain the difference in production levels between the two cell lines, suggesting that, there is a slightly reduced level of Mab HC and LC present within the MedI-BA cells. In DTE-expressing stable pools there is non-normal distribution of both HC and LC abundance present in both parental and evolved cell lines, with the exception of pool 5. That there is little difference between ETE HC/LC levels in different cell lines on day 4 of culture is unsurprising, given that at this stage the titre of both cell lines is comparable, with the main difference in production levels between the two cell lines coming after this point in growth (figure 5-10 f). However, in DTE-expressing cells at day four of culture there is already a big difference in the titre levels of parental and evolved cell lines.

The non-normal distribution of DTE HC/LC abundance levels seen in both MedI-CHO and MedI-BA cell lines, alongside the fairly similar distribution levels seen in parental and evolved cell lines expressing ETE Mab – both Mabs showing a difference in productivity levels between the two cell lines despite similar HC/LC abundance profiles - suggests that the difference in Mab productivity levels between the parental and evolved cell lines is not due to differing levels of Mab HC and LCs being produced. It may be possible that the expression plasmid copy number in the parental cells was higher than that in evolved cells, but biological replicates have controlled for this occurrence.

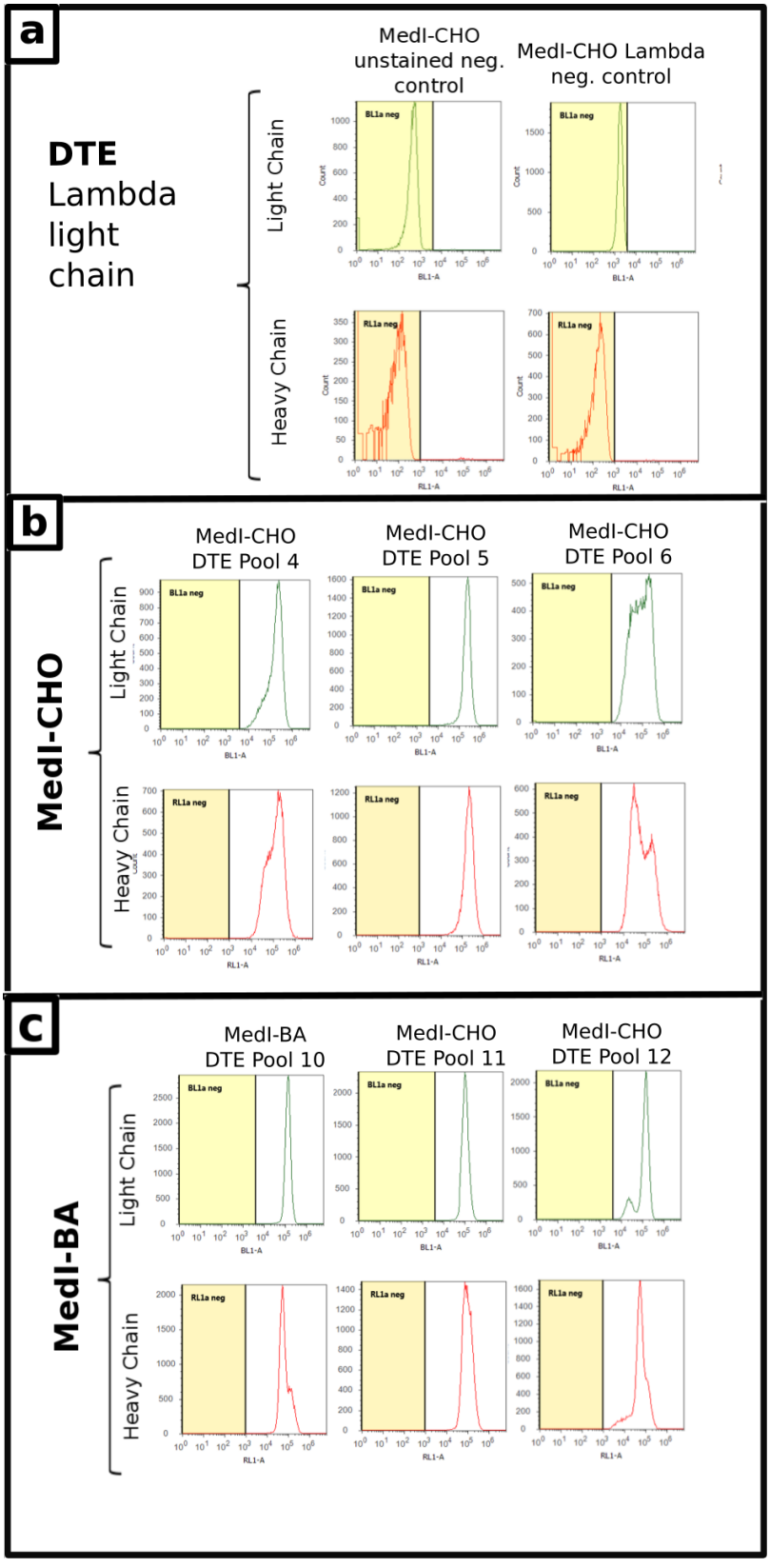


**Figure 5-13. Intracellular HC/LC abundance in DTE-expressing stable pools.**

*ETE-expressing stable pools (with a TEE) were MeOH fixed before staining with anti-human HC and anti human lambda LC antibodies. [a] Flow cytometry analysis was carried out to elucidate intracellular HC and LC levels across the [b] MedI-CHO and [c] MedI-BA stable pool populations. Charts to be inferred as Figure 5-12.*

A decrease in productivity in BFA-evolved cells compared to parental cells disproves the hypothesis that evolution of CHO cells against BFA would allow the cell population to overcome the secretory blockage imparted by BFA, resulting in an population with an enhanced secretory and productivity phenotype. This is especially surprising given that MedI-BA stable pools out-grew MedI-CHO stable pools, although as discussed this growth difference may be due to a positive reaction to the removal of BFA selective pressure from MedI-BA cells. That the productivity decrease was seen in both ETE- and DTE- expressing stable pools confirms data seen in transient pools (figures 5-5, 5-6), but is the opposite of what was seen in the initial early screening of stable pools at the end of one passage in shaken culture (figure 5-7). This suggests that it takes time for stable pools to stabilise with regards to Mab production, but also raises the question as to the efficiency of screening cells/cell populations for phenotypic differences early in their life cycle.

#### *5.4.2.6. Organelle morphology of BFA-evolved CHO cells.*

Previous antibody staining of BFA-treated MedI-CHO cells showed that BFA treatment results in the dispersal of the Golgi in CHO cells (figure 5-5). ETE- and DTE- expressing TEE pools (both parental and evolved) were sampled in the exponential growth phase of shaken culture at G30. Cells were PFA fixed and stained with anti-GM130, anti-GBF1 and anti-Calreticulin antibodies (section 2.7.1) to determine whether stable Mab expression has any impact upon ER and Golgi morphology or the cellular localisation of the BFA target protein GBF1. Pools were also stained with an anti-BiP (HSPA5/GRP78) antibody to determine the UPR status of the cells. Non-transfected MedI-CHO cells were stained as a non-Mab producing control, with all cell imaged by IF (figure 5-14).

All staining images show positive staining when compared to secondary-antibody only negative controls (images not shown), confirming that fluorescence is not due to background binding. That staining with all antibodies results in similar signal intensity to that seen in the non-expressing negative control confirms that none of the antibodies used cross-reacted with the recombinant Mab. Were this the case signals would be very intense (based on previous staining of a Mab stably-expressing cell line stained with anti-human antibody showing high, often saturating fluorescence levels; data not shown). The lack of large deviation in signal between test and control subjects in these images suggests that staining in Mab-expressing cells is real and not an artifact of cross-reaction between primary antibodies and recombinant Mab.

The molecular chaperone BiP is a downstream activator of XBP1s and its levels are known to increase upon UPR activation, with BiP levels correlating to the level of UPR activation (Johari et al. 2015). The abundance levels of BiP do not appear to differ from that of the negative control (figure 5-14), whether cells are ETE- or

DTE-expressing, MedI-CHO or MedI-BA derived, suggesting that none of these different factors has an obvious impact upon the UPR status of the cell.. Golgi morphology, shown by GM130 staining, does not appear to differ in ETE-expressing MedI-CHO and MedI-BA compared to the non-expressing control, with concentrated perinuclear cisternal bundles seen in all cells types. However, there does appear to be slight disruption in Golgi morphology in some DTE-expressing MedI-BA cells, suggesting a potential reason for the reduction in productivity seen in MedI-BA cells.

Across all cell lines ER morphology, shown by Calreticulin staining, does not appear to differ. However, when taking in to account the difference in image exposure required to ensure image saturation did not occur, it can be suggested that Calreticulin staining in DTE-expressing cell lines is brighter than that seen in non- and ETE-expressing cell lines, suggesting a possible increase in Calreticulin levels, and therefore possibly ER volume, within DTE-expressing cells. GBF1 staining in expressing cells is slightly more intense than that seen in the non-expressing control. For expression of both model Mabs, once differing exposure times have been taken in to account (specifically the lower exposure required for DTE-MedI-BA), it can be suggested that there GBF1 levels in MedI-BA cells are higher than in MedI-CHO cells, suggesting that evolution against BFA has resulted in an increase in GBF1 levels and therefore potentially in COPI vesicle activity.

Overall the microscopy data shows that there is little difference in ER and Golgi morphology between ETE-, DTE- and non-expressing cells, nor between the two different cell lines. There may be some difference in image intensity across the differing cell and Mab lines, suggesting possible differences in BiP and GBF1 expression levels. However, IF provides only a qualitative measurement of expression levels that can not be used to accurately quantify expression levels. Differences in staining intensity could also be due to artifacts arising from staining (e.g. cell density across slide, cell number per slide differing, interaction between primary and secondary antibodies) and imaging (exposure settings, bleaching of samples during preparation/imaging). For more accurate data on BiP and GBF1 levels within the different cell lines, a more quantitative technique, such as quantitative western blotting, would be required.

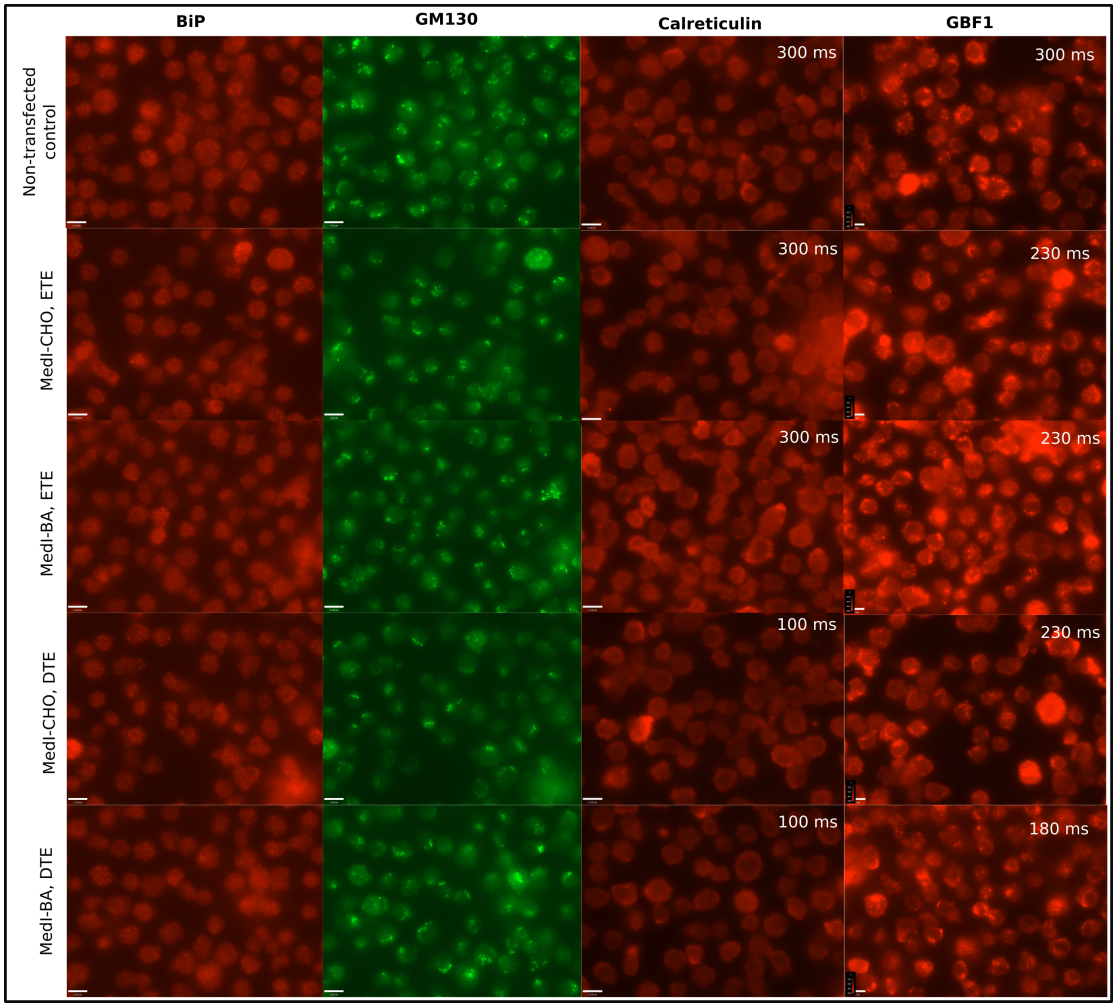


**Figure 5-14. ER/Golgi morphology of ETE and DTE stably-expressing CHO cells.**

*ETE- and DTE- stably-expressing MedI-CHO and MedI-BA cells were harvested in exponential growth phase, PFA fixed and stained with antibodies against the UPR marker BiP, the Golgi marker GM130, the ER marker Calreticulin and the Golgi-located BFA target protein GBF1. Samples stained with the same antibodies were imaged for the same duration (ms) where possible. Where this was not possible the different exposure times are shown so better comparison between images can be made (longer exposure time will lead to greater colour intensity. Gain and offset settings remained the same across all images of each antibody). Images shown are indicative of two technical replicates of three biological replicates. Scale bars: 11.00 µm*

### 5.4.3. Transcriptomic and proteomic analysis of BFA-evolved CHO cells.

To determine whether the process of directed evolution had resulted in a change in the transcriptomic and/or proteomic nature of the CHO cell, the levels of genes/proteins that are either targeted by BFA or involved in the UPR were measured. Western blotting was used to determine protein levels, with RT-PCR used to determine transcript levels. With mutation of BFA's target likely to result in BFA resistance without increasing cell productivity, further to the transcriptomic analysis sequencing of the secretory blocking agent's specific target was performed do see whether a point mutation(s) had indeed occurred.

#### *5.4.3.1. Determining UPR status of evolved CHO cells.*

MedI-CHO and MedI-BA cell lines were grown in biological triplicate for ten passages before sampling during exponential growth. Whole cell protein samples were produced by RIPA extraction (section 2.7.3.1), with all RIPA samples normalised for cell number. Protein samples were separated by SDS-PAGE (section 2.7.3.2) before transfer to a membrane for blotting with primary antibodies raised against GBF1 (a target of BFA) or the UPR marker BiP. β-actin was used as a housekeeping protein (sections 2.7.3.3, 2.7.3.4; table 2-4). After secondary antibody incubation membranes were imaged by Li-Cor and quantitatively analysed. Whilst initial samples were normalised for cell number, quantitative data from the test antibodies (GBF1/BiP) were normalised to the level of the housekeeping control to further ensure samples could be better compared. Samples were taken from three biological replicates of both cell lines, with Western blotting repeated twice to provide two technical replicates. Non-producing cell lines were used to produce samples for Western blot analysis as intracellular Mab had been shown to cross-react with some test antibodies.

Initial chemiluminescent Western blotting of GBF1 confirmed the functionality of the GBF1 antibody, suggesting there may be an increase in GBF1 levels in BFA-evolved cells, however further experiments showed this was not replicable. During this experiment GBF1 staining (with both Li-Cor and chemiluminescent secondary antibodies) was not successful and as such no GBF1 data is shown.

Quantitative Western blotting of BFA-evolved and parental cell samples with an anti-BiP antibody implies that there is a small increase in BiP levels in MedI-BA cells, although this is not statistically significant (figure 5-15). This suggests that the BFA-evolutionary process has no significant effect upon BiP levels and as such it can be assumed that the evolutionary process has had no impact upon the UPR status of the CHO cell.

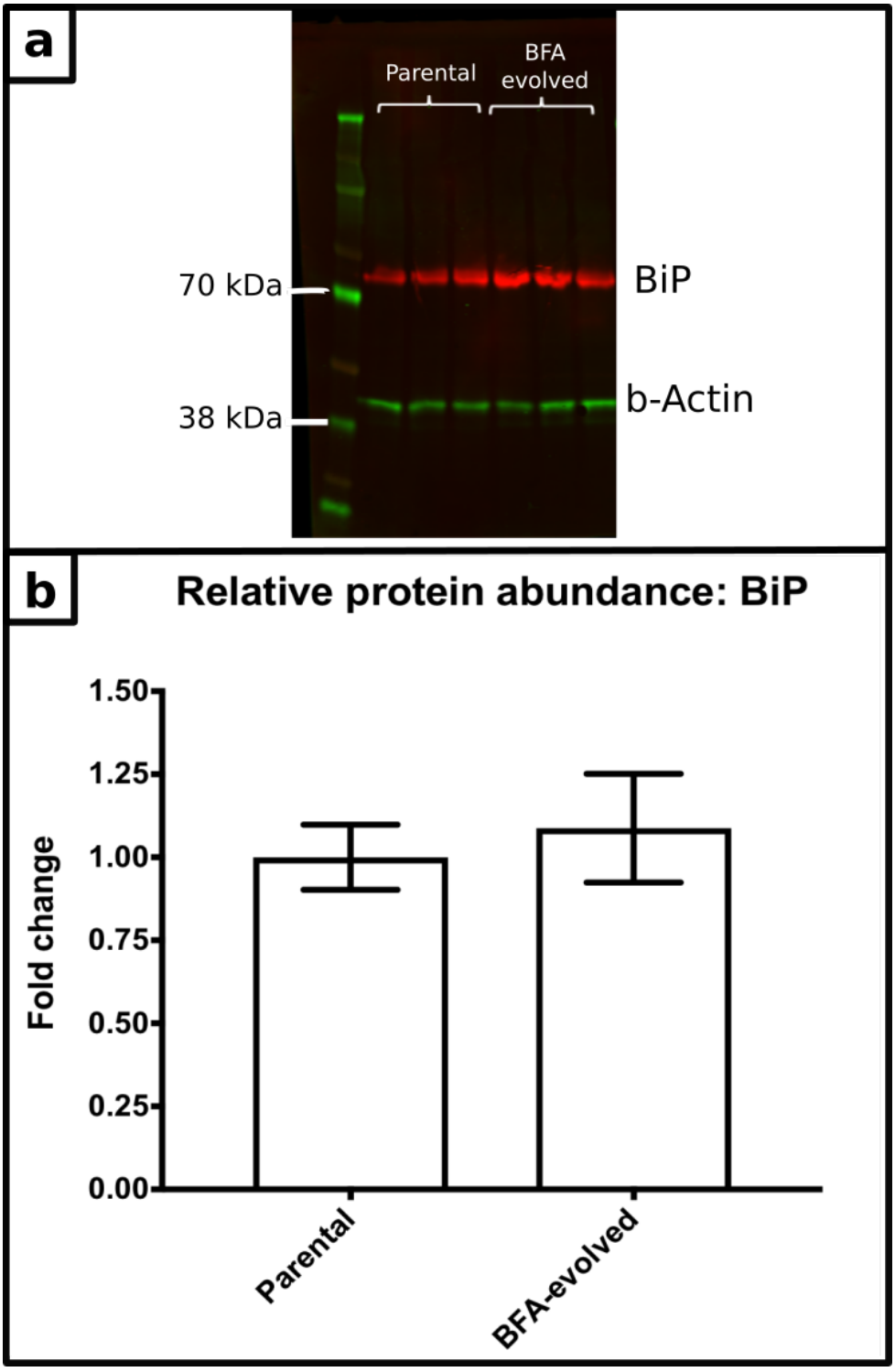


**Figure 5-15. Western blot analysis of BiP levels in evolved CHO cells.**

*The levels of BiP proteins in MedI-BA and MedI-CHO cells were analysed by Western blot and Li-Cor quantification. β-actin was used as a housekeeping control. [a] BiP-representative bands were visible at the expected 72.4 kDa and β-actin at the representative 41.7 kDa. Image is representative of two technical replicates. [b] The intensity of bands from the Western blot image were quantified using Image Studio software. Intensity of the BiP bands were first normalised against the housekeeping β-actin before fold-change compared to the parental control was calculated. Columns show the mean of two technical replicates of three biological replicates (n = 6). Error bars show SEM. No statistical significance was found.*

#### *5.4.3.2. Determining transcription levels of UPR response genes and GBF1 in BFA-evolved cells.*

The effect of the BFA-evolution process upon the transcriptional activity of BFA targets and UPR-related genes was investigated. Total RNA of three biological replicates of MedI-CHO and MedI-BA was extracted and converted to cDNA by reverse transcription (sections 2.8.1 , 2.8.2). Forward and reverse primers were designed to determine the transcript levels of three known BFA targets and four UPR-related genes, with two housekeeping genes used as controls. To analyse the effect BFA evolution had on specific BFA targets, primers were designed to analyse transcript levels of GBF1, ArfGef1 and ArfGef2. Whilst BIG1/2 had been determined as GBF1 targets their gene sequences were not available in the CHO genome. Alignment of BIG/ArfGef protein sequences from other species resulted in 100% alignment (including coverage of the Sec7 domain) and so ArfGef1/2 gene sequences were substituted for BIG1/2 gene sequences. To determine the effect BFA evolution had upon the UPR, primers were designed to amplify segments of BiP, CREB3L2, ATF6 and XBP1 genes. The MMADHC and FKBP1a genes had previously been determined from MedI-CHO RNA-seq data as housekeeping genes (Brown et al. 2018). Primer sequences are shown in table 2-5, with primers synthesised by IDT.

Primer pairs were tested across a cDNA serial dilution to determine efficiency and cDNA levels required. Primer efficiencies between 95-105% were deemed acceptable for RT-PCR (table 5-3; section 2.8.4.1). Analysis of transcript levels was performed by RT-PCR. Each biological replicate underwent two analyses with each primer set to provide technical duplicates. Negative controls comprised of RT-PCR samples with DNA originating from no reverse transcriptase and no cDNA samples for each primer pair (section 2.8.4). RT-PCR Ct values were converted into a transcript abundance fold-change value using the double-delta calculation, comparing test transcript levels in the parental and test cell lines to those of the two housekeeping genes (section 2.8.4, equation 8).

Transcript analysis of MedI-BA cells shows that there is a 2.68-fold increase in GBF1 transcript levels in MedI-BA cells compared to the parental cells, with this difference being statistically significant. A 1.52-fold increase in ArfGef2 levels is seen in MedI-BA, alongside slight increases in ATF6, BiP, XBP1, although a slight decrease is seen in ArfGef1. However, none of these are statistically significant. However, there does appear to be some variance in the housekeeping genes, with MMADHC seeing a 15% decrease and FKBP1a a 14% increase in transcript levels when compared to the parental cell line, though this difference is not statistically significant (figure 5-16). With initial cDNA levels controlled for in the double-delta calculation, this variance from the expected value (of 1) suggests that low-level transcript abundance changes should not necessarily be fully attributed to the evolution process. The relative transcript abundance of CREB3L2 proved too low

to differentiate between the test subject and the negative controls and as such the data is not presented.

**Table 5-3. RT-PCR primer efficiencies.**

| Primer pair | Efficiency |
|---|---|
| GBF1 | 102.73% |
| ArfGef1 | 95.34% |
| ArfGef2 | 101.48% |
| ATF6 | 99.89% |
| BiP | 102.34% |
| CREB3L2 | 102.02% |
| XBP1 | 96.81% |

*The primer pairs designed for RT-PCR were tested for efficiency. Efficiency values - calculated by the double-delta calculation - between 95-105% were deemed acceptable for RT-PCR.*

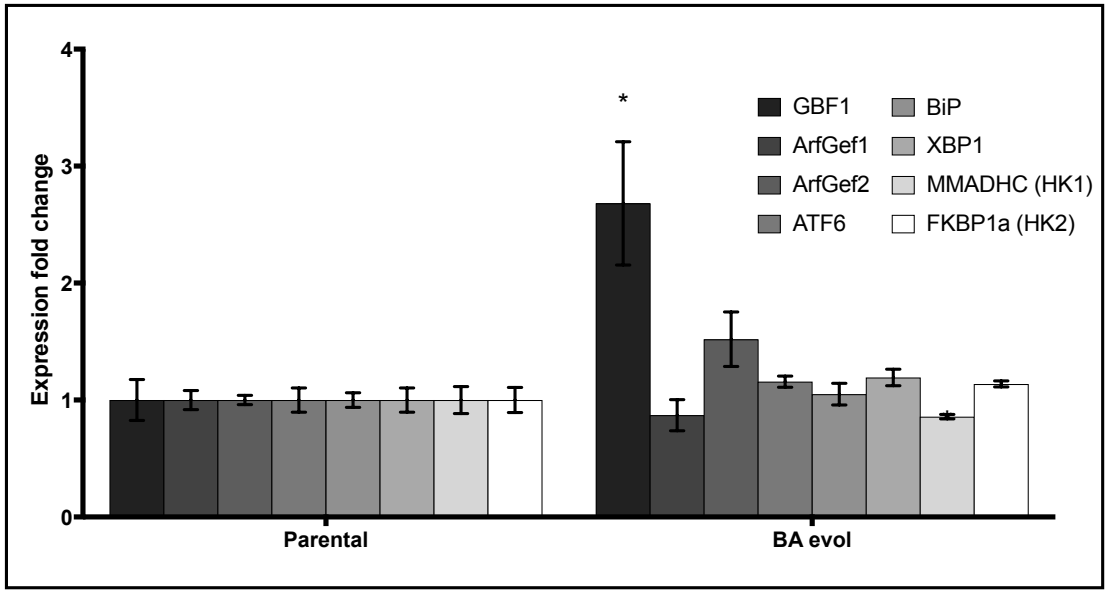


**Figure 5-16. RT-PCR analysis of secretory and UPR gene levels in BFA-evolved cells**

*The transcript levels of UPR-related genes and BFA targets were analysed by RT-PCR. RNA was extracted from MedI-CHO and MedI-BA cells and reverse transcribed to cDNA. Primers were designed to analyse transcript levels of GBF1, ArfGef1, ArfGef2, ATF6, BiP and XBP1 genes, with MMADHC and FKBP1a used as housekeeping genes and positive controls. Non-cDNA negative controls showed no transcript levels (data not shown). RT-PCR Ct values were converted to an expression fold change using the double-delta calculation, comparing test gene and housekeeping gene transcript levels in both parental and evolved CHO cells. Columns show the mean average of three technical replicates of three biological replicates (n = 9). Error bars show SEM. Statistical significance was calculated by a Dunnett's multiple comparisons test.*

#### *5.4.3.3. Determining presence of mutations in GBF1 Sec7 domain*

BFA is known to interact with the Sec7 domain of the GBF1, with Sec7 inhibition by BFA inhibiting Arf-GEF function (Claude et al. 1999; Mansour et al. 1999; Peyroche et al. 1999). It was planned to sequence the Sec7 domain of endogenous GBF1 so as to determine whether BFA resistance in MedI-BA cells had come about due to a mutation within the Sec7 domain, making GBF1 impervious to BFA treatment. Had this been the case the cell population would have developed a BFA resistance phenotype without the hypothesised increased secretory capacity. RNA samples were taken from MedI-BA and MedI-CHO cells and reverse transcribed to produce cDNA. From this the Sec7 domain of GBF1 was amplified by PCR using forward and reverse primers also used for RT-PCR (sections 2.8.3, 2.8.5; tables 2-4, 2-7, 2-8). However between the PCR and PCR product purification processes there was not enough Sec7 sequence to allow sequencing to be performed. Repetition of this experiment was planned but could not be preformed due to time constraints.

### 5.4.4. Effect of FLI-06 evolution/selection on productivity and morphology of CHO cells.

The FLI-06 compound is known to halt constitutive secretion in mammalian cells through the blocking of cargo loading into COPII vesicles at the ERES, though the mode of action through which it achieves this is not known (section 5.2.2.2). It is shown to have a large impact upon CHO cell growth, with addition of 0.25 μM at the start of culture being shown to stop cell growth whilst possibly having an impact upon ER morphology (figure 5-4 a). As with BFA, MedI-CHO cells were evolved against increasing concentrations of FLI-06 with the aim of producing a CHO cell with an enhanced productivity phenotype through the improved capacity of the secretory pathway. Stable and transient expression with model Mabs was carried out to assay the impact directed evolution had upon CHO cell productivity. Morphology testing was performed to investigate potential molecular reasons for any phenotypic changes.

#### *5.4.4.1. Evolution of CHO cells against FLI-06.*

Evolution against FLI-06 was carried out in Sheffield. Four passages post-revival, MedI-CHO cells were treated with FLI-06 upon seeding. Untreated cells were maintained as a control for general genetic drift, with DMSO only treated cells (at the same volume as FLI-06 solution added to the evolution cells) maintained as a vehicle control. Each permutation (test, untreated control and vehicle control) was repeated in triplicate to provide biological replicates. Evolved cells were initially treated with 0.1 μM FLI-06. Upon recovery and stabilisation of cell growth, FLI-06 concentration was increased to 0.15 μM (at 37 generations) then 0.2 μM (at 50 generations). Intermediate kill curves were performed to determine the evolved cell line's resistance to FLI-06 to inform the concentration to which

FLI-06 should be raised with which to have an effect upon CHO cells. Kill curves at G103 showed that evolved cells could grow as normal in 0.2 μM FLI-06 and as such FLI-06 levels were increased further and quicker than previously to 0.6 μM (data not shown), then 1.0 μM at 175 generations. Cells were cultured in the presence of 1.0 μM FLI-06 until 270 generations at which point cell banks were produced. Henceforth, FLI-06-evolved cells will be referred to as MedI-FLI and DMSO vehicle controls as MedI-DMSO.

Growth metrics of MedI-CHO, MedI-DMSO and MedI-FLI over the course of evolution are shown in figure 5-17. Across the three different evolution conditions viability and growth rate are generally comparable, with growth of MedI-FLI being slightly below that of the two control conditions. Wide error bars and sudden drops in viability (and, to some extent, growth rate) can often be explained by contamination resulting in a passage having to be taken from a previous culture closer to stationary than exponential phase. Across all cultures, growth rate is seen to gradually increase across the course of 270 generations (figure 5-17 a; from an initial 0.6 towards 0.8 $\mu$ $day^{-1}$) whilst cell volume (figure 5-17 c; as a function of cell diameter, from an initial 15 μm to 14.5 μm) is seen to decrease, both being trends that have been previously seen in long culture of CHO cells (Fernandez-Martell et al. 2017). The long gaps between increases in secretory blocker concentration, due to other time pressures, were not ideal as the evolution process took longer than ideal (nearly nine months in duration).

Whilst the selective pressure of the evolutionary agent remained, the lack of regular stimulus change within the growth environment may have allowed sub-populations to develop within the cell population as a whole. Cells may have become more stable in the presence of FLI-06 such that when an increase in concentration was made more cells within the population were able to withstand it, so diluting the evolutionary process. That an increase in FLI-06 concentration only resulted in a small decrease in growth rate and viability at the next passage, and that recovery from this is often quick (cell growth/viability levels recovering within a couple of passages), would suggest that this is indeed the case (figure 5-17 a, b. Points highlighted by arrows). A FLI-06 kill curve (up to 1 μM FLI-06) showed that MedI-FLI cells see little decrease in growth rate or viability at 1 μM compared to 0 μM FLI-06, whist treatment with 0.75 μM FLI-06 halted growth and reduced viability to ~40% in both MedI-CHO and MedI-DMSO control lines (data not shown).

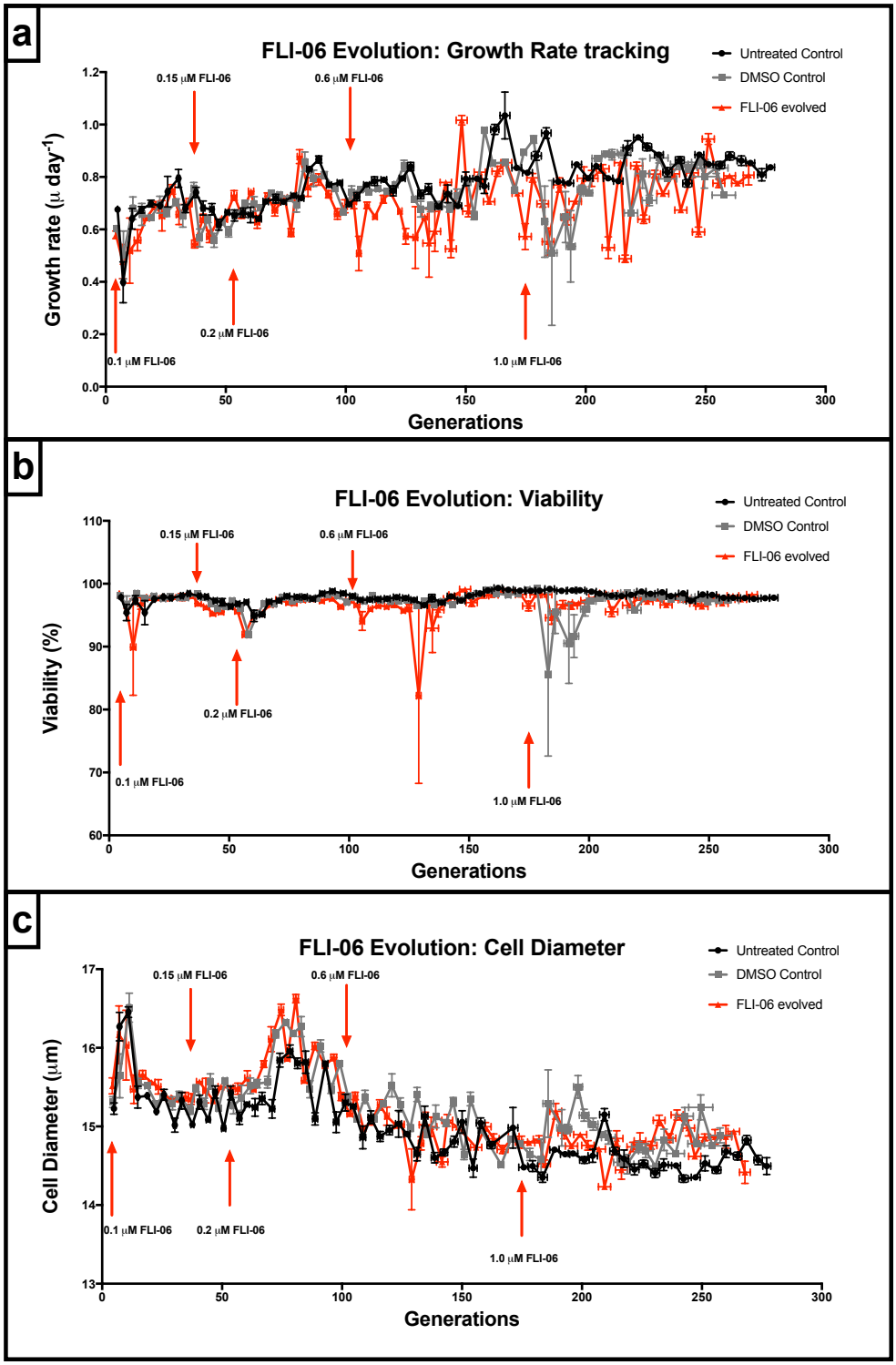


**Figure 5-17. Tracking of FLI-06 CHO cell evolutionary process.**

*Over 270 generations CHO cells were evolved against the secretory blocking agent FLI-06. [a] Growth rate, [b] cell viability and [c] cell diameter metrics were measured to allow the comparison of the progression of control (black), DMSO vehicle control (grey) and FLI-06 evolved (red) cells. Arrows highlight points at which FLI-06 concentration was increased and the level to which it was increased to. Each data point is the mean of 3 biological replicates. Both horizontal and vertical error bars show SEM.*

#### *5.4.4.2. Transient Mab productivity of FLI-06-selected CHO cells.*

Evolved and control cell lines were transiently transfected (by nucleofection, as section 2.5.1) with plasmid DNA expressing an ETE or DTE model Mab. Transfected cells were seeded in 10 mL cultiflask culture and grown for four days, at which point cell growth and supernatant samples were taken. Mab titre levels were assayed by Valita™TITER HS Mab assay (section 2.6). No evolutionary compounds were included in growth medium post-transfection. Growth and productivity of transiently-transfected cells is shown in figure 5-18.

##### *5.4.4.2.1. Transient ETE expression levels of FLI-06-selected CHO cells.*

Parental and vehicle control cell lines show similar levels of ETE expression, with the MedI-FLI cell line seeing a 1.7-fold increase in titre compared to the parental control. This is despite the reduced growth-rate of transiently transfected MedI-FLI cells, which is 20% of that seen in the MedI-CHO. Once this reduced growth-rate is taken in to account, the Qp of MedI-FLI ETE productivity shows a 3.7-fold increase when compared to MedI-CHO, with little difference between the Qp values between the MedI-CHO and MedI-DMSO control cells. The increase in both titre and Qp between MedI-CHO and MedI-FLI are both statistically significant (figure 5-18 a).

##### *5.4.4.2.2. Transient DTE expression levels of FLI-06-selected CHO cells.*

As with ETE-expression, transient transfection of MedI-FLI has with a DTE has a large impact upon cell growth, MedI-FLI's growth rate being only 4% of that of MedI-CHO, with little cell growth occurring. When expressing a DTE Mab, there is a slight difference between expression levels in the MedI-CHO and MedI-DMSO vehicle control cells. This may be due to the lower DTE productivity levels being in a noisier region of the Valita™TITER HS assay standard curve, resulting in a slight reduction in titre measurement accuracy. The titre levels of MedI-FLI are approximately half that of parental cell lines and two thirds that of the vehicle control. However, once the greatly reduced growth rate of MedI-FLI is taken in to account it sees a statistically significant 2.15-fold increase in DTE Qp when compared to MedI-CHO, although the smaller increase compared to the MedI-DMSO control is not statistically significant. This suggests that the long-term presence of DMSO may have some effect upon CHO productivity, although variance in data may result in this being noise (figure 5-18 b).

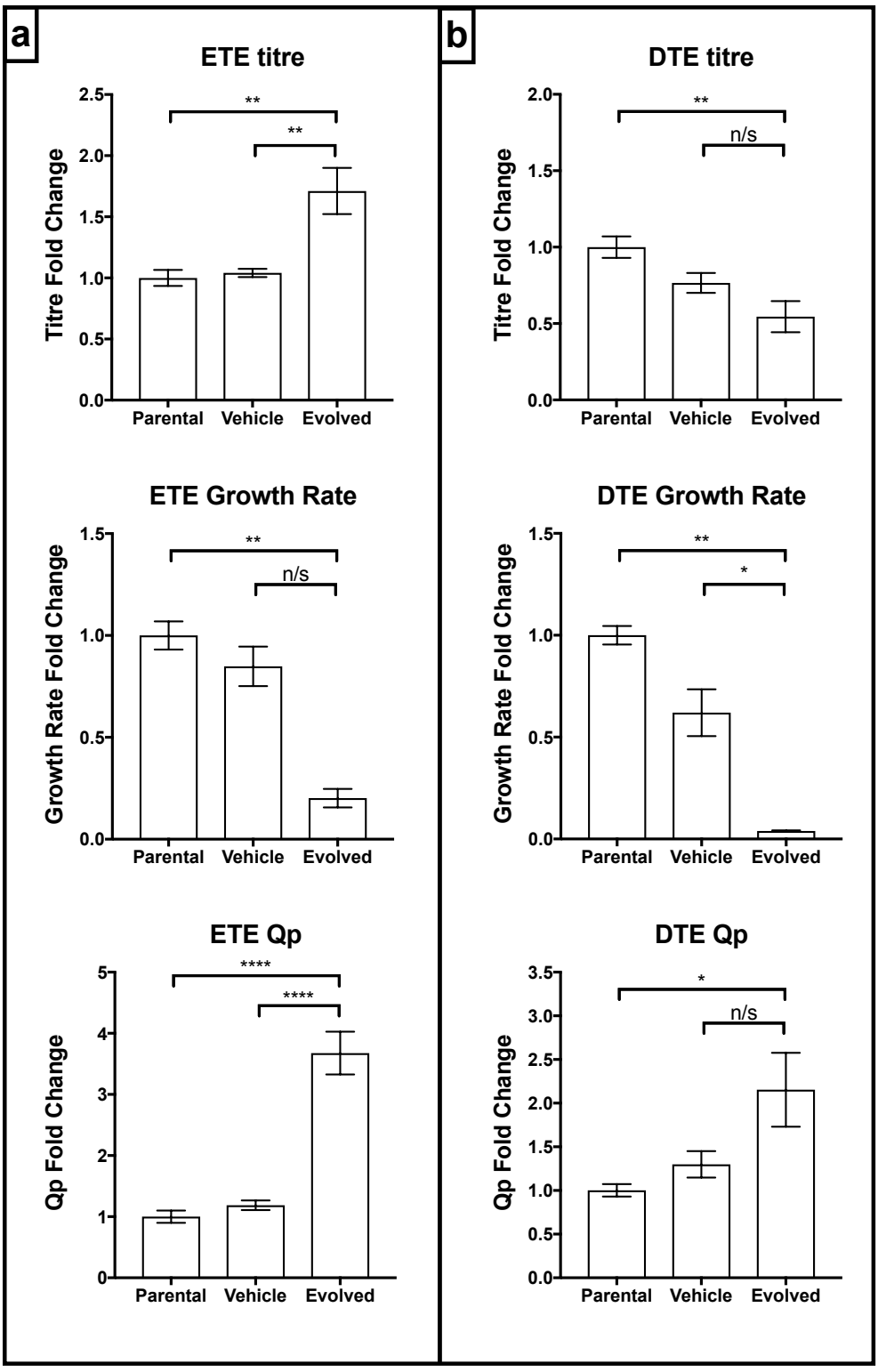


**Figure 5-18. ETE and DTE expression levels of FLI-06 evolved CHO cells.**

*Parental, vehicle control and FLI-06-evolved CHO cells were transiently transfected with ETE and DTE Mab which was expressed over 4 days. [a] Productivity and growth data for ETE Mab-expressing cells. [b] Productivity and growth data for DTE Mab-expressing cells. Data is shown as a fold-change compared to the average value of the parental cell line. Bars show mean of three technical replicates of 3 biological replicates (n = 9), with the exception of ETE vehicle control which shows 3 technical replicates of 2 biological replicates due to contamination (n = 6). Error bars show SEM.*

When taking all titre and growth data into account it appears that the production of both ETE and DTE Mabs in an FLI-06-evolved cell line is greater than that seen in the parental cell line. However, the Qp increase seen of both model Mabs is due mainly to the decrease in cell growth of transiently transfected MedI-FLI compared to MedI-CHO, with growth rate over 4 days being only 20% (ETE) and 4% (DTE) of that seen in the parental cell lines (figure 5-18). That both evolved cell lines saw a decrease in growth rate, with cell viability also decreasing and cell diameter reducing slightly (data not shown) suggests that the process of transfection may have an impact upon growth of MedI-FLI cells that is not seen in MedI-CHO. Therefore the FLI-06 evolutionary process may have altered the cell population such that it is less conducive to transfection than the parental population. It is possible that the age of the cells (close to 270 generations/100 passages over the course of the evolution procedure) has had an impact upon their susceptibility to transfection, but this is controlled by the MedI-CHO control cell lines that out-grew MedI-FLI post-transfection. As such, the transfection efficiency of both cell lines was investigated by transient transfection with a GFP reporter plasmid.

#### *5.4.4.3. Confirming transfectability of FLI-06-evolved cell line.*

Comparative MedI-CHO, MedI-DMSO and MedI-FLI cells were transiently transfected with pMAXGFP as per the protocol used for Mab expression (sections 5.4.3.2 and 2.5.1). A non-transfected MedI-CHO cell line was used as a negative control for both growth and transfection efficiency. Four days post-transfection cell samples were taken for cell growth analysis and transfection efficiency analysis by measurement of positive GFP expression by flow cytometry. Cells with fluorescence above that of 99% of the non-transfected population were deemed to have been positively transfected with GFP. Furthermore, had MedI-FLI's decrease in growth rate post-transfection been due to the physical process of transfection, a similar decrease in cell growth would be seen upon transfection with pMAXGFP.

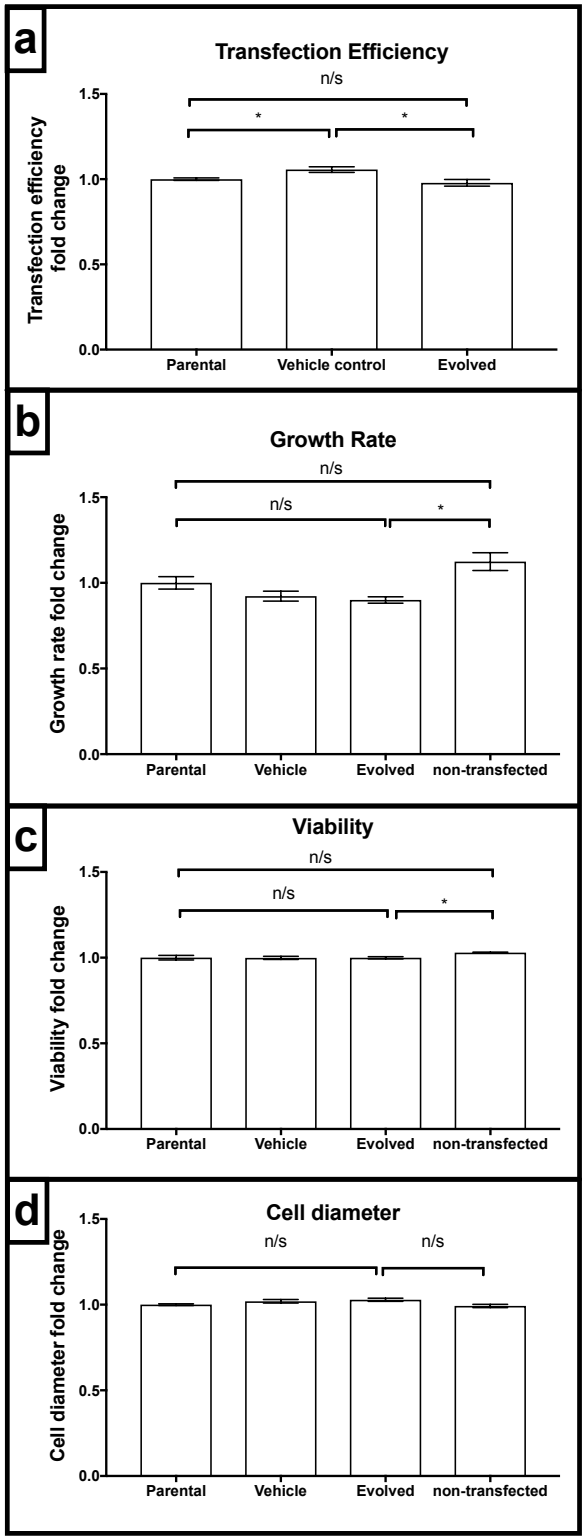


**Figure 5-19. Transfection efficiency of FLI-06 evolved CHO cells.**

*MedI-FLI and control cell lines were transiently transfected with GFP to assay the effect the evolutionary process has had upon CHO cell transfectability. A non-transfected cell line was used as a growth control. [a] Flow cytometry of GFP-transfected cells was used to determine transfection efficiency. [b, c, d] Cell growth data was collected to compare how different cell lines reacted to transfection. All bars show fold change from the mean of parental cell line data. For growth data, each bar shows one data point from three biological replicates (n = 3). For transfection efficiency, bars show the mean of two technical replicates of three biological replicates (n = 6). All error bars show SEM.*

There is no statistically significant difference in transfection efficiency between MedI-FLI and MedI-CHO, with mean transfection efficiencies of 85.0% and 83.2% respectively. However, the MedI-DMSO vehicle control cell lines, with an average transfection efficiency of 89.8%, are significantly higher (with 95% confidence) than both MedI-CHO and MedI-FLI (figure 5-19, a). There is a slight decrease in MedI-FLI's growth rate (90% of that of Med-CHO; figure 5-19 b), although this is not statistically significant and is also a much smaller decrease than that seen previously upon transfection with ETE and DTE Mabs (figure 5-18, section 5.4.3.2). There is no statistical difference in viability or cell diameter between MedI-CHO and MedI-FLI (figure 5-19), again different to what was seen in Mab-transfected cells (figure 5-18). As expected, due to undergoing electroporation MedI-CHO transfected cells grew slower (at 89.2% the rate of non-transfected cells) and with a slightly lower (2.9%) overall viability than non-transfected MedI-CHO cells.

Transfection efficiency data suggests that the evolution process has not resulted in a cell line that does not grow well post-transfection. After transfection with pMAX-GFP, whilst MedI-FLI cells saw a slight decrease in growth rate compared to MedI-CHO, this decrease (to 90%) was not statistically significant and was a much smaller decrease than that seen when both cell lines were transfected with Mab-encoding DNA (figure 5-18), where a drop to 20% (ETE) and 4% (DTE) of the MedI-CHO growth rate was seen in MedI-FLI, with viability and cell diameter data also showing a reduction (data not shown). This therefore suggests that the decrease in MedI-FLI growth-rate post-transfection is due to the increased stress upon the cell due to Mab production.

That the growth rate decrease was much larger in DTE-expressing cells further enhances this hypothesis (figure 5-18 a vs. b), with the cell unable to grow and produce DTE Mab at a high level at the same time. However, it is also possible that the difference in size between the two plasmids (ETE and DTE Mabs being approximately three times larger than that of pMAX-GFP) with larger DNA constructs know to have a slightly toxic effect upon CHO cells, though all cells were transfected with the same mass of plasmid DNA. This experiment was performed to control for the effect Mab production has upon different CHO lines as well as the physical process of electroporation opening pores in the cell membrane, not necessarily against any toxic effects of DNA. Were the experiment to be repeated, a control transfection with a non-coding vector of a similar size to the Mab vectors could be used to control for plasmid size and the number of plasmids each cell was transfected with.

The high specific productivity of MedI-FLI cells is due in part to their slow growth rate and, whilst an increase in Qp is a positive outcome, a large decrease in cell growth (such that the overall titre sees a decrease) is not. As such, FLI-06

evolution appears to improve the overall performance of both culture and individual cells transiently expressing the ETE model Mab, but whilst greatly enhancing the Qp of DTE the overall titre is not enhanced. This would suggest that FLI-06 evolved cells have a greater capacity to produce recombinant Mab, but this impinges (especially in the case of DTE production) upon their ability to produce biomass and grow.

#### *5.4.4.5. Stable transfection of FLI-06-evolved CHO cells.*

With a transient expression process only lasting four days, it does not allow cells time to recover from the physical process of transfection, as well as only allowing a short time-frame in which to produce Mab. Stable pools expressing DTE Mab were produced such that transfected cells could be allowed to recover over a longer period of time and still produce recombinant Mab such that cell line productivity could be better elucidated. All evolved lines (three each of MedI-CHO, MedI-DMSO and MedI-FLI) were stably transfected (nucleofection) with linearised plasmid DNA encoding a DTE Mab expressed from a standard expression vector as TEE-containing vectors were not available for use (protocol described in sections 2.4.4 and 2.5.3). After recovery from static to shaken culture in 50 μM MSX, stable pools were grown for 4 shaken passages before being frozen down. All banks were revived and passaged concurrently for 6 passages. Batch overgrow cultures (12 day duration) for each cell line (providing three biological replicates) were performed. Cell growth and supernatant samples were taken every 2-3 days to track the growth and productivity of the stable pools across an entire culture life-cycle. FLI-06 selective pressure was removed from MedI-FLI cells for the growth curve. Samples for intracellular HC/LC abundance analysis (section 2.7.2.1) were taken on day 4.

MedI-FLI grows slightly better than the parental and vehicle control cell lines, with cumulative IVCD reaching a higher level between days 10-12 and cell viability staying higher for slightly longer after day 8 of culture when compared to the control cell lines (figure 5-20 a). This follows the pattern of growth characteristics seen in MedI-BA stably-expressing pools (section 5.2.4.4, figure 5-10). As then, this growth increase may be due to removal of FLI-06 from the growth media used during the productivity curve experiment, resulting in a small increase in overall cell growth.

The DTE titre data collected varied greatly across all tested pools, including across biological replicates. This resulted in titre data from which no conclusions can be made due to large error bars. To investigate the reasons for this wide difference in cell productivity the homogeneity of the cell pools with regards to intracellular abundance of HC and LC was tested (section 2.7.2.1). It was hypothesised that heterogeneity of heavy and light chain production levels within the cell pools

resulted in the production differences between cell pools, and not the secretory capacity of the separate cell lines themselves.

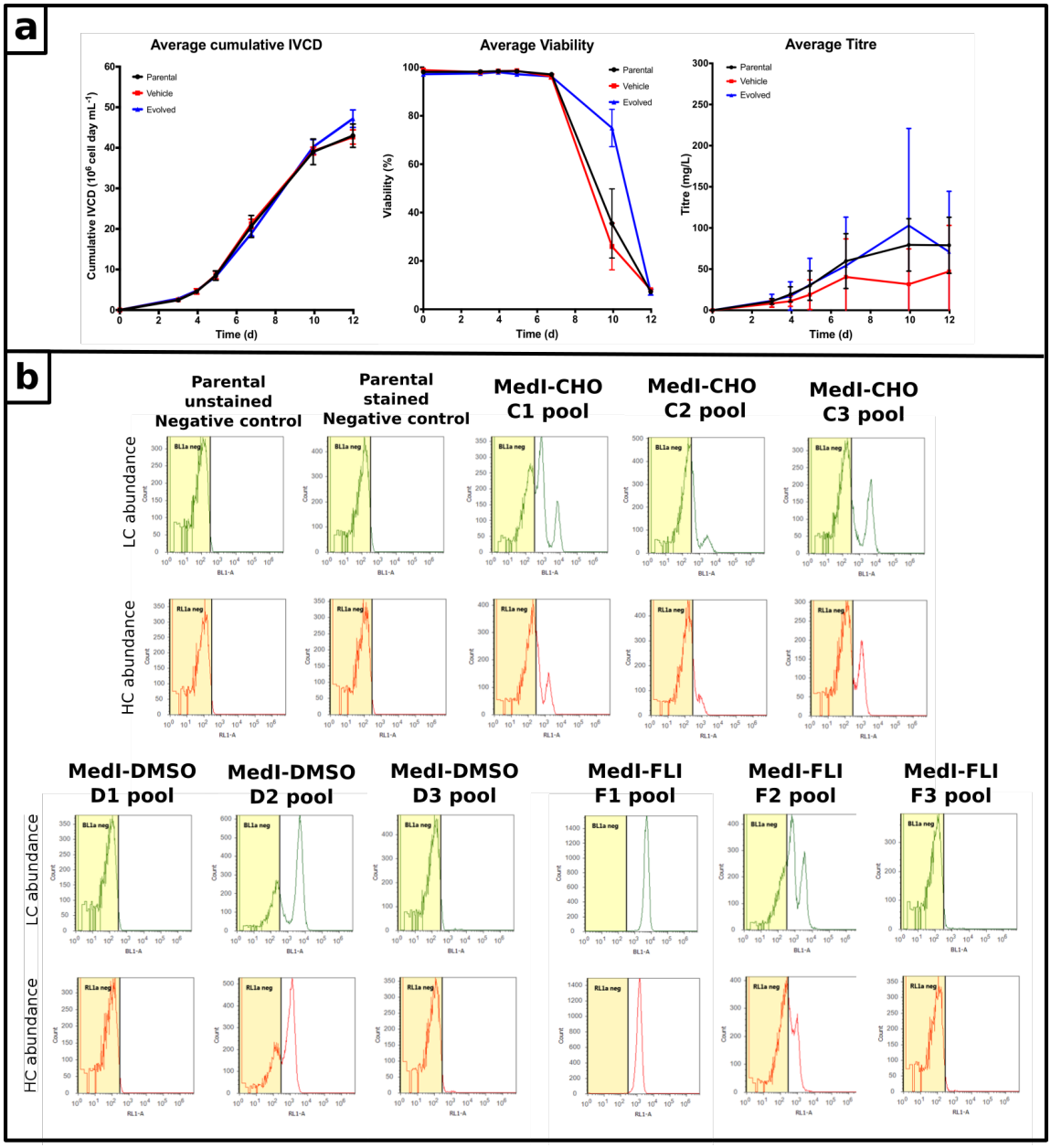


**Figure 5-20. Production of stable pools expressing FLI-06 evolved CHO cells.**

*Stable pools expressing DTE Mab were produced using MedI-FLI (and control cells) so as to assay their level of Mab productivity. [a] Growth and titre data over the course of a 12-day batch overgrow was collected to compare the three different cell lines. All data points are the average of three biological replicates. Error bars show SEM. [b] Flow cytometry of HC and LC abundance within each stable pools and two negative controls (un-stained and non-expressing but stained). Top rows (green) show LC abundance, bottom rows (red) HC abundance. Cell gating process explained in section 5.2.4.5 and figure 5-10 legend.*

HC and LC abundance across all pools shows that in many of the pools, levels of HC and LC present are heterogeneous (figure 5-20, b). In several pools (D1, D3, F3) all cells match both negative controls for both HC and LC levels, indicating that neither Mab gene is being expressed within these pools. In many pools (C1, C2, C3, D2, F2) the cell population is split in to two distinct peaks showing that the population is heterogeneous with regards to HC/LC expression levels. In all these pools, for both HC and LC abundance, one of the peaks is mostly within the negative control fluorescence gate, with the other peak showing that Mab HC and LC are both present in a section of the population. Only one pool (F1) shows a positive, normal distribution for both HC and LC abundance that matches what was expected (as seen in TEE stable pools, figures 5-12, 5-13).

This HC/LC abundance data indicates the reason for the wide difference between the different cell line, being due to their expression of HC and LC across the pools. Comparison of individual productivity data of each cell line (data not shown) further supports the importance HC/LC level population homogeneity. The F1 pool was the highest producer of all pools, with D1, D3 and F3 producing the lowest titres. The remaining pools all produce titre levels between these two distinct groups, the exact level correlating to the amount of cells positive for HC and LC abundance.

That positive HC and LC signal intensities are similar within the same pool suggests that the cause for low Mab productivity levels is not necessarily due to a stoichiometric imbalance between HC and LC levels, showing instead that there is a potential problem with the expression of genes encoded by the Mab plasmid. That non-expressing populations of all pools show similar profiles to the negative controls would suggest that a non-expressing sub-population has developed and thrived within culture. It is likely that these populations have derived from cells with high endogenous GS levels that can survive the MSX growth media was supplemented with, as all cells expressing recombinant GS should also express Mab due to the difference in strength of the promoters the transcription of both expression cassettes are transcribed from.

The variance of Mab titre and HC and LC expression levels seen with non-TEE vectors, compared to the similarity of production levels and HC/LC abundance of pools produced with TEE vectors (sections 5.2.4.4 and 5.2.4.5), shows the importance of knowing and controlling the transcription level of a recombinant gene within a cell. The continuously open-confirmation of a stable recombinant gene conferred by a TEE effectively controls across cell lines for transgene transcription levels, whereas in non-TEE plasmids this is not the case, with transcription levels depending entirely upon the transcriptional activity of the random genomic site of integration. Whilst selection and amplification should produce some level of control, it is not at the level provided by a TEE.

## 5.5. Discussion and Conclusions.

Directed evolution was used as a strategy by which to effect large-scale genetic changes upon the CHO cell. It was hypothesised that evolving CHO cells against a compound that inhibits the secretory pathway would result in myriad fine transcriptomic and proteomic changes within the cell - many more than could be instigated by direct ectopic expression/repression of genes – such that the cell population could overcome the inhibitory effect, with the aim of producing a CHO cell with an enhanced secretory pathway and thus an enhanced productivity phenotype.

A parental CHO cell line was separately evolved against two secretory blocking agent that were selected based on previous literature and efficiency testing in CHO cells (section 5.3). The resultant evolved cell lines were transfected with model recombinant proteins to determine whether the directed evolution process had changed cell productivity capacity. Cells were stained to determine any morphological changes that might have resulted in any phenotypic change. RT-PCR and Western blotting was performed to probe any potential changes in the targeted areas of the cell lines' transcriptomic and proteomic profile.

Evolution of CHO cells against BFA resulted in a cell line that, when transiently expressing ETE and DTE model Mabs, saw an overall decrease in productivity compared to MedI-CHO. MedI-BA saw a small but statistically significant increase in growth rate of 18% with both Mabs. However, overall titre and specific productivity levels both decreased, with total MedI-BA titres reaching only 80% (ETE) and 62% (DTE) that of MedI-CHO, with specific productivity of each evolved cell seeing a drop to 73% (ETE) and 50% (DTE) the level of that seen in the parental cell line (figure 5-6).

Stable production of model Mabs in MedI-BA and MedI-CHO was performed so as to better analyse cell line productivity. Use of a normal expression plasmid resulted in a wide range of results across biological replicates and as such could not be used to support a definitive conclusion as to the productivity capacity of an evolved cell line (figure 5-7, 5-8). Production of a stable pool with a plasmid containing a TEE allowed better control of transgene transcription levels, resulting in better data upon which to base conclusions (figure 5-9). Productivity and growth curves of MedI-CHO and MedI-BA stably expressing ETE and DTE Mabs showed that MedI-CHO out-produced MedI-BA by 2.2-fold (ETE) and 11-fold (DTE) in terms of maximum titre at day 10 of growth, despite MedI-BA slightly out-growing MedI-CHO between days 5-7 of culture (figure 5-10, 5-11). The evolutionary process also appeared to have little effect upon CHO cell morphology, as shown by there being little difference in secretory organelle imaging and BiP protein levels between the two cell lines (figures 5-14, 5-15).

A decrease in Mab productivity in BFA-evolved cells compared to parental cells disproves the hypothesis that evolution of CHO cells against BFA would affect myriad small genetic and phenotypic changes within the cell, resulting in an enhanced secretory pathway which in turn would produced an enhanced productivity phenotype. A possible reason for this is that BFA blocking of COPI-mediated ER-Golgi and intra-Golgi transport results in the induction of the Golgi stress response (Yoshida 2009).

It was initially hypothesised that cells overcoming induced stresses would have increased capacity for protein folding, processing and vesicle transport so as to be better able to survive these stresses, resulting in a cell line with an enhanced productivity phenotype. This response to increased UPR activation has been seen in plasma cells (Shaffer et al. 2004; Shapiro-Shelef and Calame 2005). However, in evolved CHO cells there was no significant increase in either the transcript or protein levels of the ER stress marker BiP, suggesting that any major enhancement in stress response brought about by the addition of an evolutionary agent is muted over the course of an evolutionary process (figure 5-16). This further suggests that if exposed to an inhibitory pressure over an extended duration, a cell line will eventually become resilient to the inhibitory compound's presence and thus revert to normal cellular process characteristics.

There is some precedence for this response to chronic ER stress. In rat liver cells, long-term exposure to stress (and thus repeated UPR activation) results in feedback-mediated suppression of mRNAs involved in the UPR (including BiP and some ER chaperones). This leads to a 'new normal' status for the cell in which it becomes more resistant to normal UPR responses as it has experienced them many times before (Gomez and Rutkowski 2016). That this does not happen in plasma cells is likely due to their terminally differentiated nature.

Another possibility is that, as BFA concentrations were gradually increased in a step-wise manner, the cells are only ever put under a mild ER stress, under which the UPR's cell survival responses are favored (Rutkowski et al. 2006; Rutkowski and Kaufman 2007). As such all cells will be able to survive and so there is little evolutionary advantage for cells within the population with enhanced biosynthetic or secretory capacity. Were cells to be treated with higher levels of the inhibitory agent those cells with inherently weaker protein processing and secretory capacities would be quickly directed towards apoptosis, leaving only cells with the enhanced productivity phenotype to survive. Sequencing of the BFA-susceptible Sec7 domain of GBF1 was attempted but was unsuccessful. Were this to be successful it would provide important information as to whether the BFA resistance of evolved cell lines had come about due to a mutation in the Sec7 domain meaning it is unaffected by the presence of BFA.

Evolution and selection of CHO cells against the secretory blocker FLI-06 resulted in a cell line that, when transiently transfected with model Mabs, saw a general increase in productivity compared to MedI-CHO. Despite an 80% (ETE) and 96% (DTE) decrease in growth rate, MedI-FLI saw a 1.7-fold increase in overall ETE titre, representing a 3.7-fold increase in Qp. When expressing a DTE Mab, MedI-FLI's overall titre decreased to 55% that of MedI-CHO, however MedI-FLI's Qp increased to 2.2-fold that of the parental cell line (figure 5-18). Stable pools were produced in an attempt to further corroborate this result, but data produced was not sufficient enough to allow a conclusion to be made (figure 5-20).

That FLI-06 evolved cells saw an increase in specific productivity of both ETE and DTE Mabs suggests that the hypothesis of directed evolution of the secretory pathway has not been completely disproven but that the outcome varies depending upon the inhibitory compound used. This could be due to the mode of action of the specific compound, but also the specific target it interacts with and the role of this within the targeted pathway, as highlighted by FLI-06 more directly targeting ER machinery and function more specifically than BFA.

However, the FLI-06 data does come with some caveats. Due to the inability of stable pools to produce consistent results, conclusions can only be drawn from the transient data and thus a comparison of growth and productivity over an extended period of time can not be made. However, that transient and stable transfection of BFA cell lines produced similar results does somewhat mitigate this. A further issue is that, whilst ETE-expressing MedI-FLI produced a higher titre than MedI-CHO, the overall titre of DTE was lower, despite a higher Qp. This lower overall titre would make the cell line unsuitable from a manufacturing process point of view where the trade-off between an enhanced Qp and reduced cell growth has to result in an increase in overall titre.

It has also been suggested that evolution/selection against a chemical agent can result in surviving cells that are sick, with reduced function after the evolutionary process (Majors et al. 2009; Nicolaides et al. 2005). Furthermore, treatment of CHO cells with chemical compounds can result in chromosomal abnormalities (Radha and Natarajan 1998). Neither of these outcomes would be welcome within an industrial cell line and so do raise some questions as to the practicability of using chemical compounds in cell line development.

To allow a firm conclusion to be made as to the suitability of FLI-06 as a compound with which to mediate directed evolution with the aim of producing an enhanced productivity phenotype in CHO cells, stable pools would have to be created so as to definitively assay MedI-FLI's productivity capacity. Using a TEE or TI system to produce these stable pools would ensure the transcriptional activity of the transgene is comparable across all test subjects and replicates, allowing

much more accurate data upon which to base this conclusion, as was seen in TEE-containing MedI-BA stable pools. Furthermore, RT-PCR of the UPR-related proteins could be carried out on MedI-FLI cells to see whether the evolutionary process has had an impact upon the CHO cell's UPR status. Whilst the exact molecular target of FLI-06 is unknown the relative transcript and protein levels of genes and proteins involved in ERES formation and cargo loading (e.g. PREB, Sar1 and COPII coat subunits) could be determined by RT-PCR/Western blotting to better determine the effect FLI-06 has upon the CHO cell at a transcriptomic and/or proteomic level. Furthermore robust, high sensitivity techniques could be used to sequence specific genes within the secretory and UPR pathways of evolved cell lines to better determine whether any compound resistance could be due to a mutation within the target of a compound providing an evolutionary benefit with no concomitant change in productivity phenotype (Cartwright et al. 2018).

Despite these results, the idea of using directed evolution against a chemical compound to speed up the normal evolution of long-term CHO cultures is not entirely redundant. Directed evolution against temperature reduction has resulted in CHO cells with an increased productivity phenotype, long-term passaging can improve cell growth levels and evolution against certain growth inhibitors produced in bioreactor culture have resulted in cell lines with enhanced growth and productivity characteristics (Syddall et al., manuscript in preparation; Fernandez-Martell et al. 2017; Majors et al. 2009). Over-exposure to ER stress (in this case mediated by an inhibitory compound) can result in many repeated cycles of UPR activation, dulling the cell's response to the UPR, resulting in a 'reset' (Gomez and Rutkowski 2016). Also, gentle increasing of selective pressure may not have allowed cells with a highly-secreting phenotype to out-compete those with a normal secretory capacity (Rutkowski et al. 2006).

This data has lead to the conclusion that treating CHO cells with gradually increasing concentration of an inhibitory agent may bring about resistance by mutation, or cells reverting to their normal phenotype. As such a long evolutionary process may not be suitable for production of a high-producing cell line, especially for an industrial environment. As such it is hypothesised that exposure to an inhibitory agent at a high concentration for a short period of time (e.g. only until cell growth has returned towards normal levels and no longer) may result in poor-secreting cells being filtered out of the overall population, with high-secreting cells surviving the selective pressure. This would allow the high-producers to pervade within the population, resulting in a larger number of high producers from which a high-producing single cell clone could be selected.

# 6) Short-term chemical selection of CHO cell populations to enhance cellular productivity levels.

*It was hypothesised that a short-term treatment of CHO cells with secretory blocking agents would prove a more effective strategy in producing an enhanced productivity phenotype than the long-term evolutionary process used in chapter 5. This strategy would again aim to exploit differences within the cell population but over a shorter time-frame would not allow resistance to the chemicals used to develop, instead resulting in the death of cells unable to survive the blocking of the secretory pathway. After undergoing short-term chemical selection cell productivity was assayed by transient transfection with a model DTE Mab, with Western blotting and RT-PCR carried out to better understand how cells had responded to the chemical filter.*

## 6.1. Introduction

The directed evolution approach to enhancing CHO cell productivity used in chapter 5 did not, in the case of Brefeldin A, produce a cell line with enhanced secretory capacity. Transient and stable production of both ETE- and DTE-Mabs was reduced in the MedI-BA cell line compared to non-evolved control cells. Evolution against FLI-06 resulted in a slight productivity increase in MedI-FLI evolved cells transiently expressing DTE-Mab, though this could not be replicated in stable pool production. This finding, especially that of the BFA-evolved cell line, was contrary to the initial hypothesis that blocking of a certain region of the constitutive secretory pathway would result in the CHO cell population overcoming this blockage, resulting in a cell line with a higher secretory (and therefore productivity) phenotype.

As discussed in both chapters 4 and 5, this lack of enhanced phenotype could be due to the overexposure of the CHO cell to the selective pressure. Long-term exposure to an inhibitory agent may have resulted in the development of mutations within that compound's target pathway. These mutations would convey resistance to the inhibitory compound, giving them a selective advantage over the rest of the cell population and allowing the mutation to take hold and prevail within the cell population. Any mutation(s) would therefore allow cell survival and perseverance within the population without any change in phenotype.

It has also been shown that growing cells under the same conditions for an extended period of time allows genetic drift resulting in changes in cell phenotype

(Davies et al. 2013; Fernandez-Martell et al. 2017). Whilst this effect was controlled against during the evolutionary processes used in chapter 5 it is feasible that this could have an effect upon cell growth and productivity that is unrelated to the addition of the secretory inhibitor.

Finally, long-term exposure to BFA and FLI-06 could instigate some level of chronic ER stress upon the cell population, resulting in feedback suppression of the unfolded protein response (Gomez and Rutkowski 2016). This response could also be replicated within the Golgi stress response. Were this the case the cell would eventually overcome the imposed stress without affecting any phenotypic change. Finally, focusing upon a very specific molecular target may only result in a small change within the cell's proteotype, having little impact upon overall phenotype. Selection of an inhibitory compound with a more global effect may result in a more wide-ranging transcriptomic/proteomic changes within the cell, resulting in a more concerted phenotypic change.

To specifically enhance CHO cell productivity through targeting of the secretory pathway, it was hypothesised that a more condensed evolutionary process might better enhance the overall productivity of a CHO cell population than a process that was drawn-out. Short-term treatment of a cell population with a secretory blocking agent would have an effect of filtering the cell population, removing cells with a lower secretory capacity and thus allowing those with better capacity to prevail, resulting in an overall population with an enhanced secretory capacity.

The advantage this short-term selection would have over a longer-term evolution is that it would allow the process of overcoming a blockage to occur without a long enough exposure to allow mutations to take hold within the population, or the effect of chronic ER stress dampening the ER stress response. Whilst application of a chemical inhibitor as a selective pressure (as in directed evolution) may allow a genotypic change to occur (e.g. point mutation), resulting in resistance to the compound without a change in phenotype occurring. This genotypic change could feasibly result in an increase of the compound's specific target and not pathways around it, resulting only in the enhancement of a specific section of the biosynthetic pathway, merely shifting the bottleneck to elsewhere within the biosynthetic pathway and not enhancing overall secretion.

Furthermore, a shorter time-frame required compared to a full evolutionary strategy would potentially allow high-producing cell lines to be selected earlier in the cell line development process, shortening time-lines in the development of a high-producing CHO cell chassis with a high-secreting phenotype (as discussed in Fan et al. 2013). Other strategies discussed included evolution against an ER-stress inducer to instigate more global changes around the biosynthetic and secretory pathways, and secondary evolution, where cell surface host-cell

proteins essential for cell function and survival were blocked. These (e.g. essential amino acid transporters) are transported to the cell surface via the constitutive secretory pathway and so blocking their presence would result in increased constitutive secretion to ensure their presence at the cell surface. As such this increase in constitutive transport may also increase flux of recombinant product.

## 6.2. Short-term chemical selection process

The same compounds shown to impact upon cell growth in chapter 5 – Brefeldin A and FLI-06 – were used to chemically select parental CHO cell populations with an aim to enhance their overall productivity (Orci et al. 1991; Yonemura et al. 2016). A single compound chemical selection with just BFA was initially performed. A second dual chemical selection was also performed using both BFA and FLI-06 in concert to target two specific regions of the secretory pathway at the same time; Golgi-originating transport vesicle formation and ER cargo loading respectively. Cells were treated with BFA or a cocktail of BFA and FLI-06 (henceforth referred to as BFA/FLI) with each compound at three lethal dose levels: LD25, LD50 and LD75. These concentrations were determined by kill curves in sections 5.3.2.3 and 5.3.2.4, figures 5-4 and 5-5. The calculated lethal dose concentrations of both FLI-06 and BFA are shown in table 5-2.

The short-term chemical selection process was designed to briefly treat cells with the secretory blocking compounds to quickly and efficiently remove cells with a reduced secretory capacity from the overall population. The main difference between this process and that of the full evolution processes used in chapter 5 are:

1. Cells are only treated with one concentration of drug. In the evolutionary process cells were initially treated with a low concentration of the inhibitory compound which was gradually increased as the process proceeded.
2. Once cell growth has recovered cell culture is stopped. During the evolutionary process once cells had adapted to the inhibitory compound they were maintained in its presence for varying lengths of time. This extended period would likely induce chronic effects upon the cell and its organelles.

The short-term chemical selection strategy is designed not to drastically change the overall phenotype of a cell population through wider changes in the transcriptomic and/or proteomic profile of a cell, as with a straight evolution strategy where full inhibition of a target was performed to overexpress the target (Claude et al. 1999; van der Linden et al. 2010). It is designed to allow for the exploitation of the inherent phenotypic heterogeneity within a CHO cell

population to filter out the best-producing cell lines, whilst not exposing the cells to the two conditions described above.

Within a population of cells there is some level of heterogeneity of cell phenotypes (Davies et al. 2013). The spread of these phenotypes depends upon various cell functions and parameters (e.g. those that are tangible such as growth rate, cell volume; those that are less tangible, such as secretory capacity). It is a combination of these parameters that result in the cell's productivity phenotype, and these parameters are expected to be normally distributed. As such, there will be some cells in the heterogeneous population – those towards the right of the normal distribution – that will have a desired phenotype for a combination of the required functions (figure 6-1). Cells with desirable phenotypes (e.g. high Qp) are normally selected and carried forward during the clone screening process. However, it may be possible to use a chemical to filter out cells with non-desirable phenotypes, enriching the population for desirable phenotypes.

As such, it is hypothesised that utilising a short-term chemical selection on the cell population will enrich the population for cells with a favourable phenotype, without resulting in any major genotypic changes. This could possibly shorten the clone screening process used in industrial cell line development pathways.

### 6.2.1. Short-term chemical selection protocol design

Host MedI-CHO cells were split from the same initial population into 10 mL cultiflask cultures. Cultures were treated with the required concentration of BFA or BFA/FLI, at LD25, LD50 or LD75 (table 5-2). Passaging continued in the presence of the required inhibitors and was generally a reactionary process. Passages longer than the standard four-days duration were used if cell recovery to the presence of the inhibitor required an extended period of time, with some no-inhibitor passages used for the same purpose if required. Cells were continuously passaged until the growth rate under selective pressure had recovered, for a two passages duration, to that of the non-treated parental control cells. At this point cells were immediately banked. Untreated cells were maintained as a control. Each drug and dose combination (including untreated controls) was repeated in three separate cultures so as to provide three biological replicates.

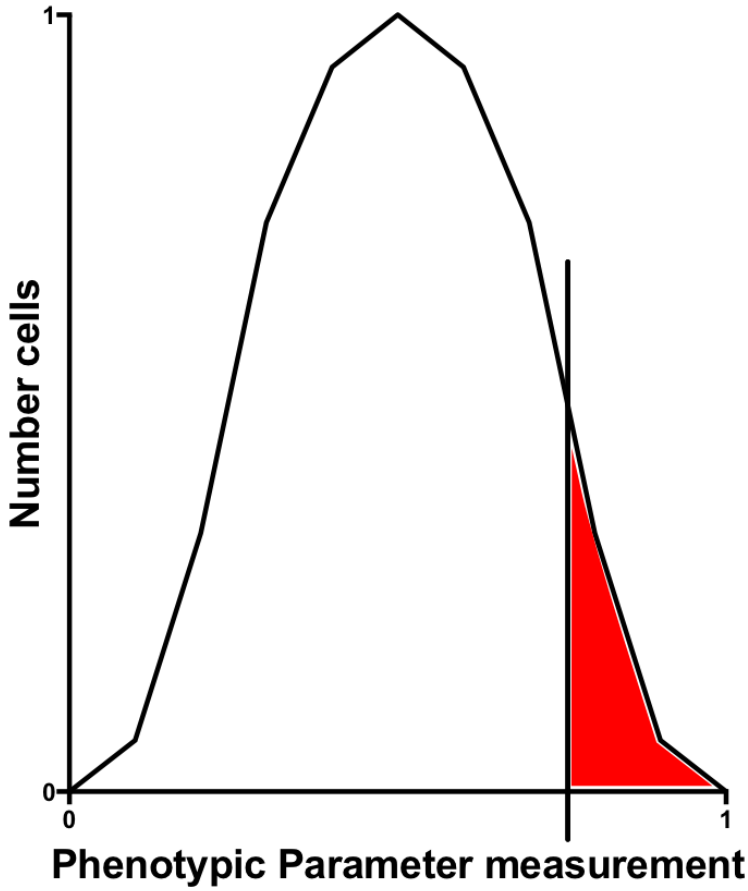


**Figure 6-1. Visualisation of the normal distribution of a CHO cell population.**

*Phenotypic parameter measurement is ranked on the x-axis between a scale of 0 (poor) to 1 (excellent). The section highlighted in red represents the region of the population with the phenotype for which we are trying to select, with the section in white expected to be removed during the process of the chemical filter.*

## 6.3. Single short-term chemical selection of CHO cells with Brefeldin A

The BFA concentrations used for the single BFA chemical filter are as shown in table 5-2. The progress of cell growth was tracked throughout the filtering process to determine cell recovery to the presence of BFA. The growth of the filtered cells is shown in figure 6-2.

### 6.3.1. Cell growth during BFA-filtering process

Due to the presence of BFA in the chemically filtered cells, cell growth was slower than that of the control. As such the generation number of the chemically filtered cells progressed slower than that of the control cells, despite undergoing a similar number of passages (10 passages for LD75 and LD50 cultures, compared to 11 passages for control cells). Treatment with LD25 BFA resulted in no obvious change in culture growth and viability compared to the control and as such cells were banked after three passages of treatment (figure 6-2 a, b).

The cells selected with LD50 and LD75 BFA concentrations reacted very similarly despite the 0.372 μM difference between the two. Cell growth for both was very low for the first 3 - 4 passages, growth rate stabilising at approximately 0.6 $\mu$ $day^{-1}$ after 10 generations of treatment. Recovery of the selected cells to the level of the control cells occurred at approximately generation 25, after passages 9-10. At this point in time growth rate matched that of the control cells, though these were at generation 40 due to the slower growth rates and extended doubling time of the filtered cells.

### 6.3.2. Productivity analysis of BFA-selected CHO cells.

All BFA-filtered cells were revived from cell banks for productivity analysis by transient transfection with plasmid DNA encoding DTE-Mab, using the Lonza nucleofector™ cuvette system according to the protocol described in section 2.5.1. Titre was determined using the Valita™TITER assay. Growth and productivity data for the transfected cells are shown in figure 6-3, with data shown as a fold-change of the average value of the un-filtered (LD0) control cells.

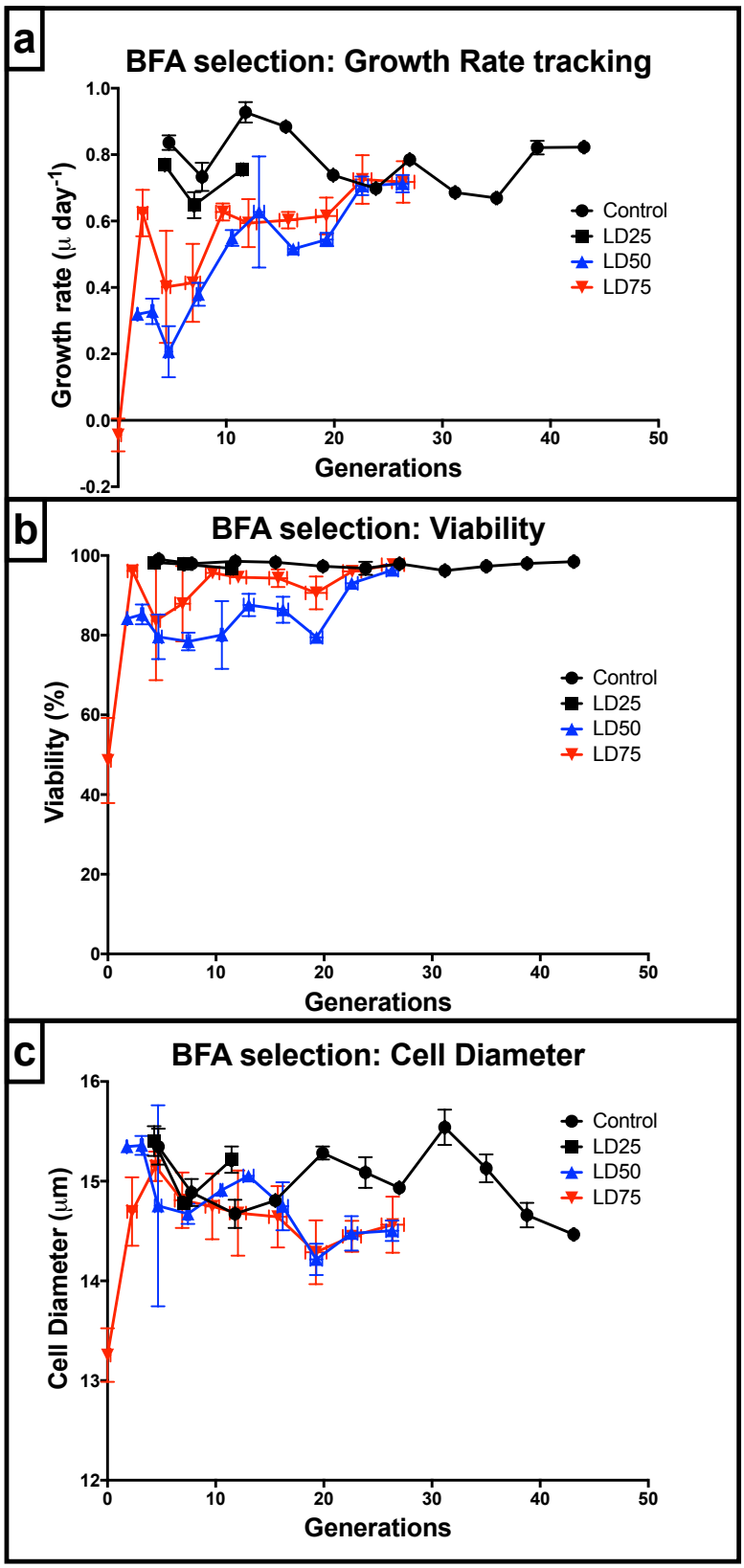


**Figure 6-2. Cell growth metrics of CHO cells during short-term BFA selection process.**

*.CHO cells were treated with BFA at concentrations equivalent to LD25, LD50 and LD75. Cell growth data was collected at the end of every passage to track the [a]cell growth rate, [b] cell viability and [c] cell diameter throughout the chemical selection process. Each point shows data collected at the end of a single passage, showing the mean of three biological replicates. Error bars show the SEM of the growth metric measured (y-axis) and the generation number (x-axis).*

Post-transfection all three selected cell lines saw a reduction in growth-rate compared to that of the un-selected control cells (with an average μ = 0.67). Whilst cells selected at LD25 and LD50 BFA levels see a decrease in growth rate to approximately 75% that of the control, growth rate of cells selected at LD75 BFA level dropped to 61% of that of the host line. All of these decreases in growth are statistically significant (figure 6-3 a). Viability of all transfected cells is quite similar to that of the control cells (average of 95.9%), although a drop (to 98.3% of control) in the LD25 cell lines is statistically significant (figure 6-3 b).

The DTE-Mab productivity levels of the LD25 and LD75 selected cells do not differ significantly to that of the non-selected control, which produced an average titre of 4.6 mg/mL. However, the LD75 selected cell line produced an average titre of 5.8 mg/mL, a 1.26-fold increase compared to the titre level of the control (figure 6-3 c). When taking into account the reduced growth rate of all the selected cell lines, these titre values result in a 1.6-fold increase in Qp for the LD25 and LD50 cell lines from the control (which saw an average Qp value of 0.69 pg $cell^{-1}$ $day^{-1}$). The LD75 selected cells however saw an average Qp increase to 2.05 pg $cell^{-1}$ $day^{-1}$, a close to three-fold increase in Qp when compared to the control.

As with the evolved cell lines produced in chapter 5, the process of nucleofection with a model Mab resulted in a significant decrease in the growth of the chemically filtered cells (sections 5.4.2.1, 5.4.3.3). To ensure that the selection process had not resulted in a reduction in the ability of a cell line to be efficiently transfected, transfection efficiency was compared between the control and LD75 cell populations. Each cell line was co-transfected with a total of 5 μg of DTE-Mab and pMAX-GFP in a 1:1 molar ratio. Transfection efficiency was measured by flow cytometry, with a non-transfected cell line used as a negative control for gating analysis of cell fluorescence (section 2.7.2). One of the LD0 cell lines did not grow as expected due to contamination and as such the data for this cell line was removed, resulting in three biological replicates of the LD75 cell line and two of the control LD0 cell line.

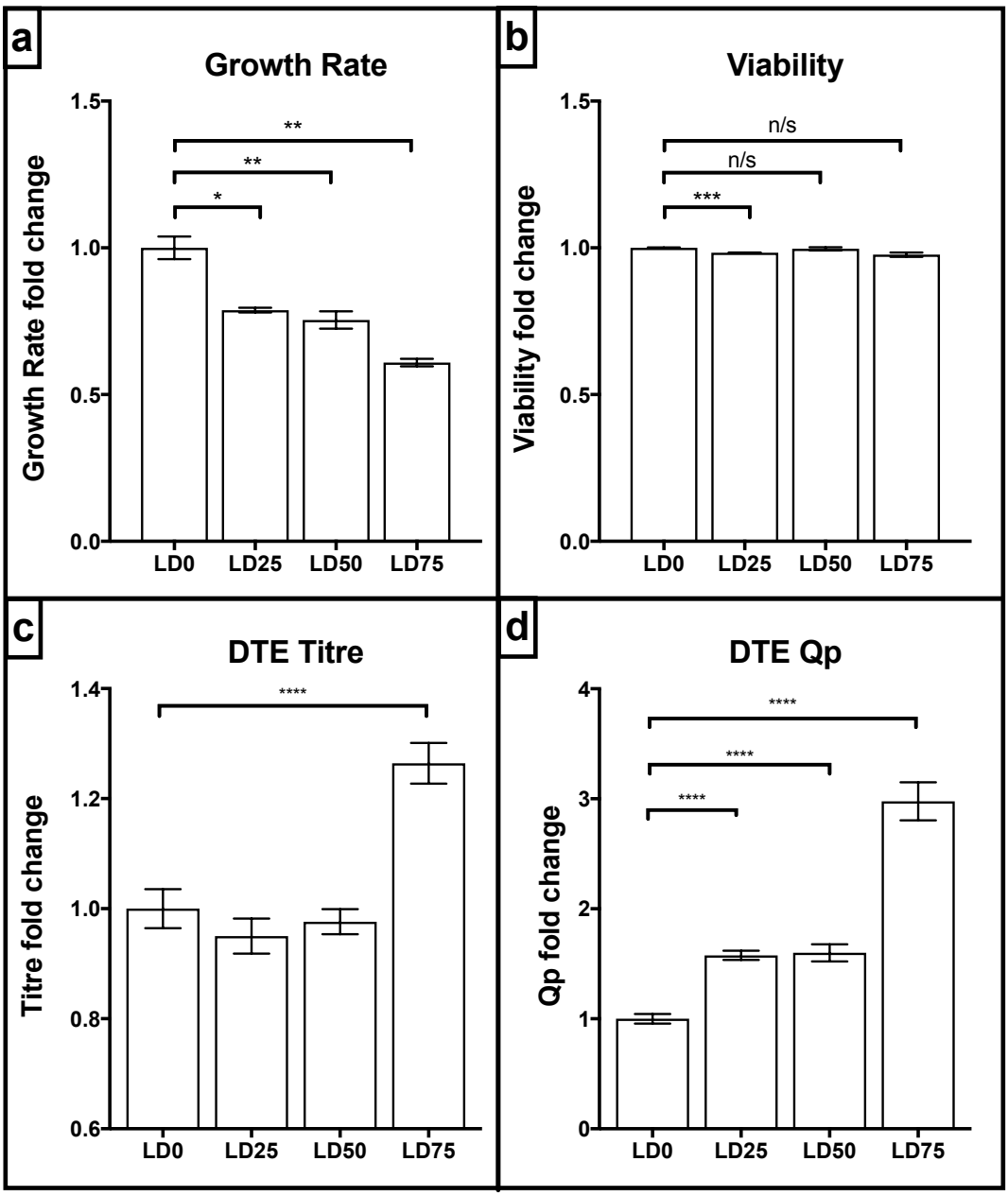


**Figure 6-3. Productivity of BFA-selected CHO cells transiently expressing a DTE-Mab.**

*BFA-selected cells were transiently transfected with plasmid DNA encoding DTE-Mab. After four days' growth cultures were sampled for growth and productivity data. Data is shown as a fold change compared to the control (LD0) cells for [a] cell growth rate, [b] viability, [c] overall titre and [d] specific productivity. For growth/viability, columns show the mean average of three biological replicates. For titre and Qp, columns show mean average of four technical replicates of each of three biological replicates (n = 12). All error bars show SEM.*

As previously discussed (section 5.4.2.1), GFP-transfection showed that there is no significant difference in transfection efficiency between the treated (in this case selected) and parental cell lines (figure 6-4 a). Although transfection efficiency was lower than expected (58% for the control), all cell lines were comparable.

This reduction in efficiency may well be due to pMAX-GFP expression diminishing by day four of culture, though diminishing expression would be replicable across all cultures. As with the Mab-only transfected cells, growth rate in the BFA-selected cells reduced to 68% (61% in Mab-only transfected cells) of that seen in the control cells with only a small change in viability, which remained at 97% of that of the control in both experiments (figures 6-4 b, c; 6-3 a, b). Both experiments saw a significant increase in overall titre, though this was more pronounced (1.9-fold increase) in the co-transfection experiment than the single transfection (1.3-fold; Figures 6-4 d; 6-4 c). This may be due to the use of an ELISA to measure productivity level in the co-transfection experiment (due to stock issues) whilst Valita™TITER HS was used to measure productivity in the single-transfection experiment. However, whilst exact titre measurements may vary across the two assay types, fold-changes and trends between the two separate assays had been seen to be comparable (Thompson et al. 2017).

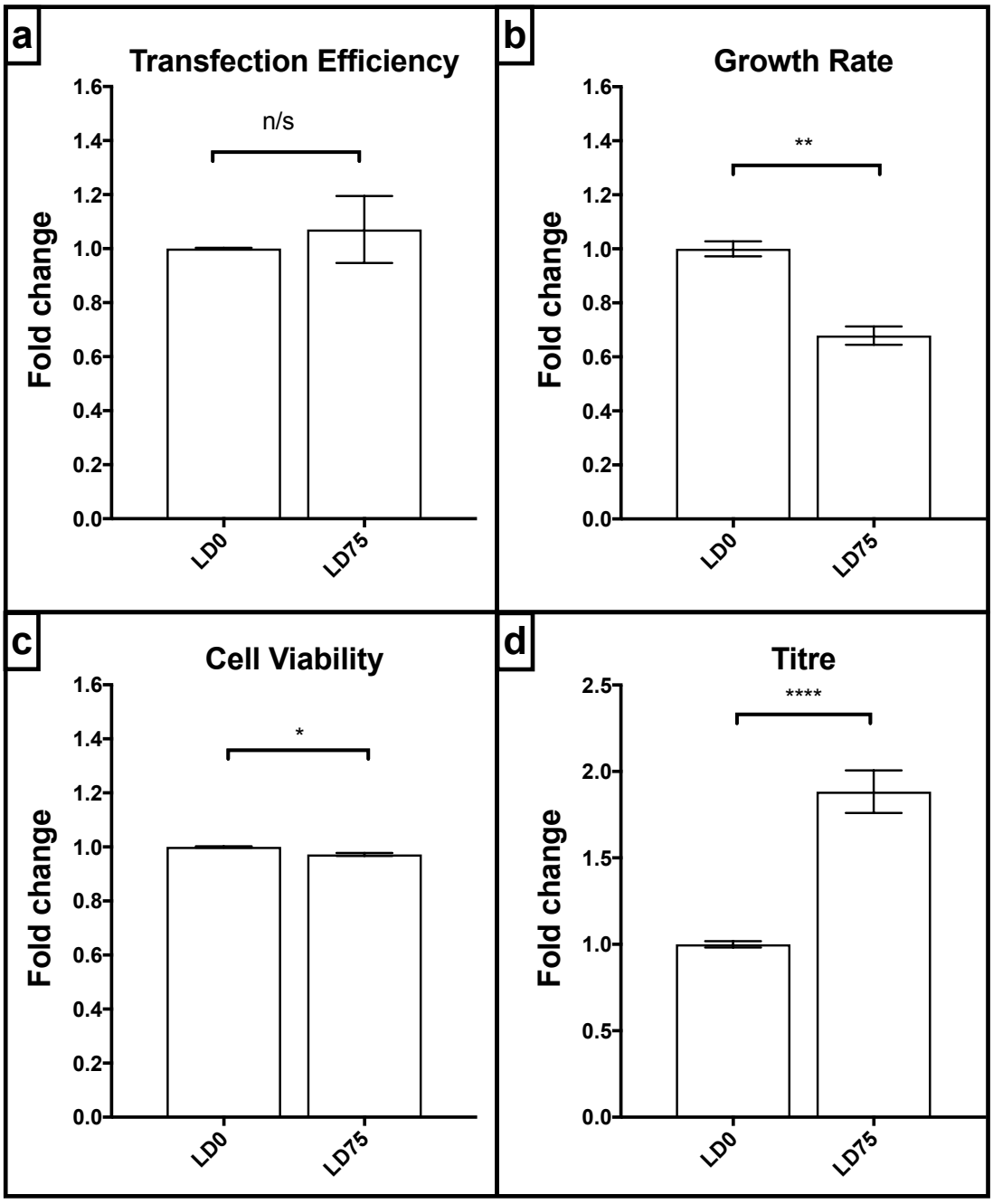


**Figure 6-4. Transient co-transfection of BFA-selected cells with DTE-Mab and GFP.**

*BFA selected cells (LD0 control and LD75) were transiently transfected with DTE-Mab and pMAX-GFP plasmid DNA. After four days' growth cells were sampled for [a] transfection efficiency, [b, c] cell growth and viability and [d] productivity. Bars for growth data and transfection efficiency show the mean average of two (LD0) or three (LD75) biological replicates. Bars for titre show the mean of three technical replicates of two (LD0) or three (LD75) biological replicates (n = 6/9 respectively. Error bars show SEM.*

### 6.3.3. Stability of BFA-selected CHO cells.

The stability of the BFA-selected cells was checked to ensure that the removal of the selection chemical and the cell-banking process had not resulted in the cells losing their BFA resistance. A kill-curve was performed on all cell lines to determine BFA-resistance levels and judge how the different cell lines reacted after ten passages without the presence of BFA. Kill curves were carried out in 96 deep well plate culture at 0, 0.5 and 1.0 μM BFA, with growth analysed by PrestoBlue® (section 2.3.2).

Kill curve growth data is shown in figure 6-5 as a fold-change from the requisite cell lines cultured without the presence of BFA. The kill curve shows that the unselected control cells, alongside the LD25 cells, have no resistance to BFA. Whilst at 1.0 μM BFA the LD75 cell line sees a slight reduction in growth rate compared to that of the same cell line grown without BFA (94% of LD75 cells with no BFA present) this decrease is not at the same level as that seen from the LD50 cell line at the same BFA concentration (which had 72% the growth rate of the same cell line grown without BFA). The kill curve highlights the difference in BFA resistance between the three differently selected cell lines, showing that even after removal of BFA the BFA-resistance phenotype of the LD75 cell line perseveres. However due to determine whether full resistance levels had been maintained the kill curve would have to be extended to the LD75 level (1.22 μM), with a maximum value of 1.0 μM BFA used for ease of preparation.

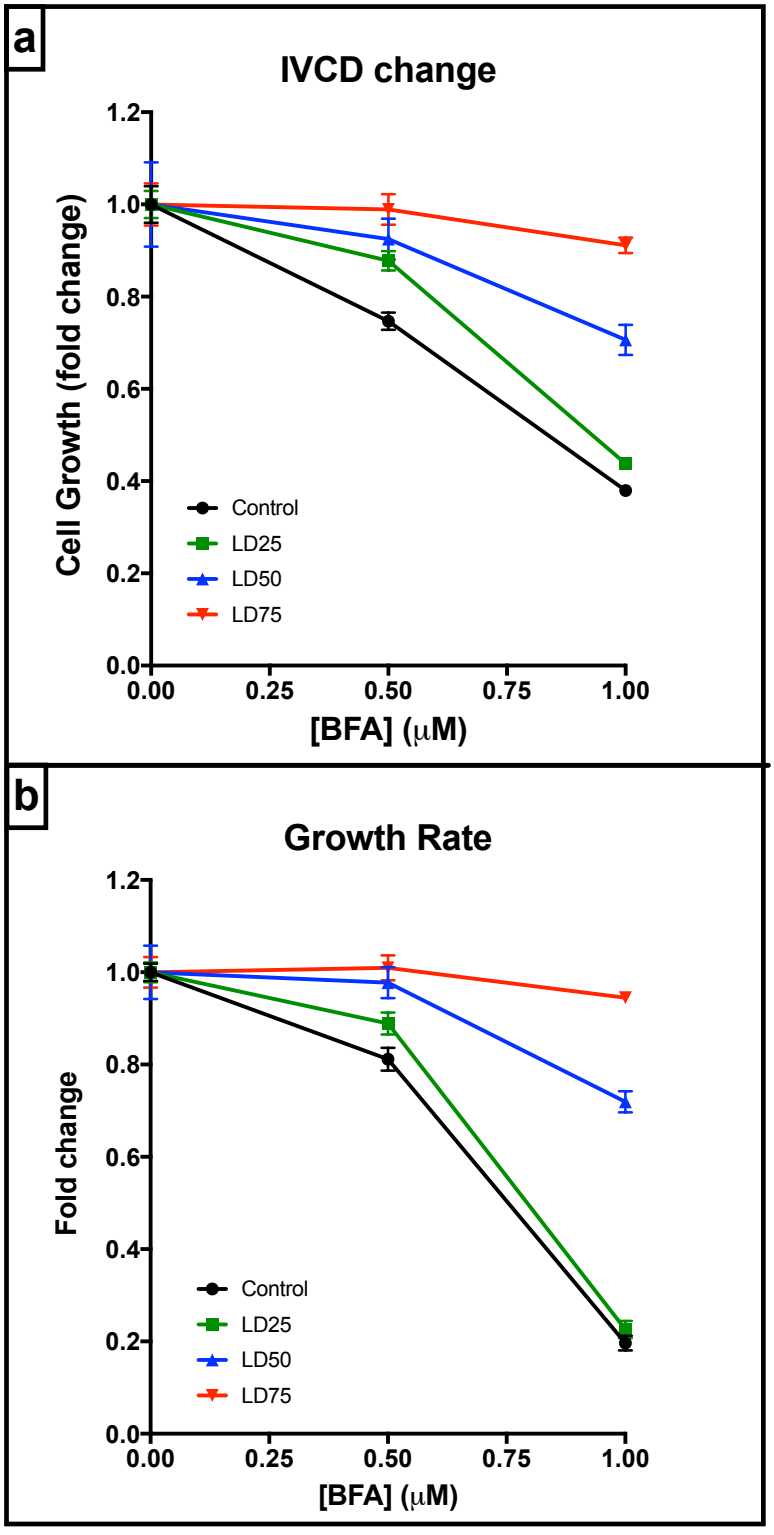


**Figure 6-5. Kill-curve showing BFA-resistance of BFA-selected cells**

*After ten passages of growth without the presence of BFA, the BFA resistance of CHO cells selected with BFA at control, LD25, LD50 and LD75 levels were tested in a kill curve. All the different cell lines were grown with BFA in a shaken 96 deep well plate at 0, 0.5 and 1.0 μM BFA. Cells were grown for three days and cell growth measured by PrestoBlue®, allowing [a] IVCD change over the course of three days' growth and [b] growth rate to be determined across all cell lines and BFA concentrations. Each data point shows the average of two technical replicates of three biological replicates (n = 6). Error bars show SEM.*

## 6.4. Dual short-term chemical selection of CHO cells with Brefeldin A and FLI-06.

The BFA-only (single) short-term selection of CHO cells resulted in a 1.2-fold increase in titre and a 3-fold increase in Qp of DTE-Mab (figures 6-3, 6-4). Directed evolution of CHO against the ERES blocking agent FLI-06 was shown to decrease CHO cell growth upon transient transfection, but also increasing CHO productivity (section 5.4.3). It was hypothesised that a dual short-term chemical selection with both BFA and FLI-06 would select cells based on enhanced secretory capacity at two different points within the secretory pathway at the same time – at ERES formation (FLI-06-mediated) and Golgi-ER and inter-Golgi transport (BFA).

As with the BFA-only single selection, CHO cells were treated with both compounds at LD25, LD50 and LD75 concentrations. Treatment with both compounds was performed concurrently in an attempt to increase the power of the selection process by imposing harsher conditions upon the cell population. The LD values for each of the two compounds used in the chemical selection were the same as those calculated from the initial kill curves (sections 5.3.2.3 and 5.3.2.4; table 5-2). The LD values of each individual compound were used as the effect of BFA and FLI-06 upon the CHO cell is quite different. FLI-06 is more potent than BFA, with FLI-06's LD100 value (0.32 μM) being approximately five-fold less than that of BFA (1.59 μM; table 5-2). Furthermore FLI-06 affects CHO cell growth over a smaller concentration range (0.15 – 0.25 μM) than BFA (0.2 – 1.2 μM; figures 5-4 b and 5-5 b). With these differences in compound efficiency a dual kill curve would prove difficult to perform with each FLI-06 concentration having to be tested at each BFA concentration and *vice versa*. As such elucidation of an exact ratio of the two compounds in combination that was specifically measureable in terms of LD value would prove difficult and variable.

### 6.4.1. Cell growth during BFA and FLI-06 dual short-term chemical selection.

As with the BFA-only chemical selection, the presence of two compounds inhibitory to CHO cell growth slowed cell growth compared to that of the control cells, reducing the number of generations the filtered cells progressed through despite undergoing a similar number of passages. Cell cultures were passaged as required with the BFA/FLI-06 combined levels maintained throughout with the exception of an early passage where recovery of cells to a sufficient level at which they could be passaged. As FLI-06 was dissolved in DMSO, two DMSO-only control cultures were maintained alongside two untreated controls. The growth of these two separate controls varied little throughout the selection process. Growth data from all cultures was taken throughout the selection process to track and compare cell line progression. Selected cells were banked once their growth rates had recovered to within 10% of the parental cell line (figure 6-6).

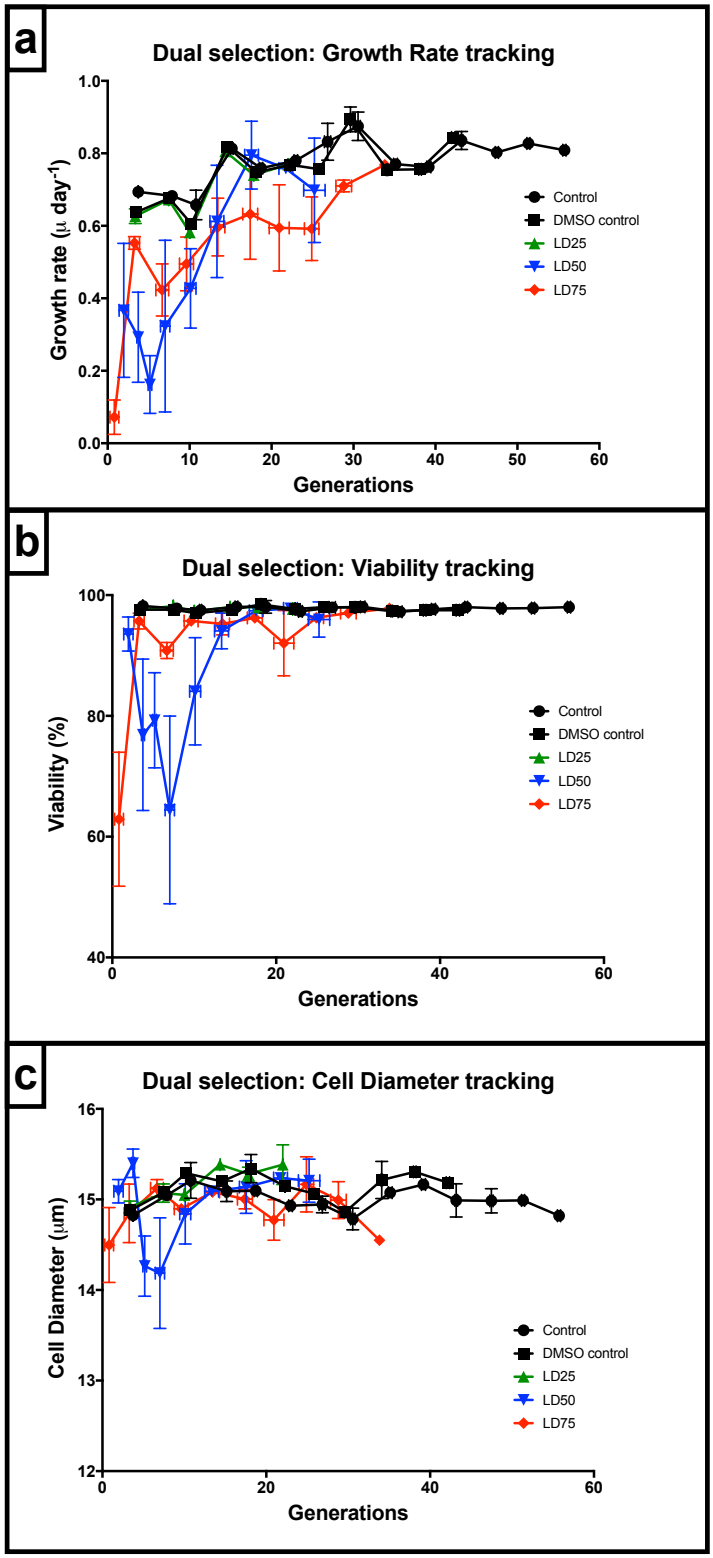


**Figure 6-6. Tracking of BFA and FLI-06 dual-short-term selection process**

*Parental MedI-CHO cells were treated with BFA and FLI-06 in combination at concentrations equivalent to LD25, LD50 and LD75. Cell growth data was collected at the end of each passage to track the [a] growth rate, [b] cell viability and [c] average cell diameter throughout the chemical filtration process. Each data point shows data collected at the end of a single passage, showing the mean of three biological replicates. Error bars show the SEM of the growth metric measured (y-axis) and the generation number (x-axis).*

As with the BFA-only short-term selection process, the growth metrics of the LD25-treated cells were comparable to those of the control cells (figure 6-6 a). The LD50 and LD75 cell lines, as in the BFA single selection process, took between six to ten passages (20-30 generations) for growth to recover to the level of the control cells. In general the LD50 cells recovered quicker than LD75 cells in terms of both growth rate and viability (figure 6-6 a, b). Whilst the cell size of some selected cells varied early in the selection process, this also stabilised within roughly five passages (figure 6-6 c).

### 6.4.2. Productivity analysis of short-term dual-selected CHO cells.

All dual-selected cells (biological duplicates from control and DMSO controls; biological triplicates from LD25, LD50 and LD75 test subjects) were revived from cell banks and grown for at least four passages without the presence of any selective pressure. After three days of post-passage growth, cells from each culture was transiently transfected by nucleofection, using the Lonza Amaxa™ cuvette system (section 2.5.1). Cells from each culture were transfected with plasmid DNA made up of a 1:1 molar ratio of DTE-Mab and pMAXGFP such that both cell productivity and transfection efficiency could be assayed concurrently. Post-transfection cells were grown for four days before sampling for growth, transfection efficiency (by flow cytometry) and titre (by Valita™TITER). Growth and productivity data are shown in figure 6-7, with data shown as a fold change of the average value of the control cells. Control data from both untreated and DMSO-controlled cells has been combined due to the lack of differentiation between these two separate controls in terms of both cell growth during the filtration process and behaviour post-transfection.

Transfection efficiency of the control cells is slightly lower than previously observed (54%), although as in section 6.2.2 this may be due to the four-day expression process. However, as all cell lines were treated the same they are directly comparable. The transfection efficiency of the LD25 cells does not differ to that of the control. However, the transfection efficiency of the LD50 and LD75 dual-selected cell lines see a significant increase, with a 17% and 21% increase in GFP-positive cells respectively (figure 6-7 a). In contradiction to growth rates seen in the BFA selected and FLI-06-evolved cells, in the dual selection the post-transfection growth rate of the LD50 cells sees a significant 15% increase. LD75 cells also see a 12% increase though as this is not significant it could be due to noise (figure 6-7 b). There is no change in viability between the cultures, though cell diameter does significantly increase by 5% in the LD50 and LD75 cultures (data not shown, $p < 0.01$ and $< 0.05$ respectively).

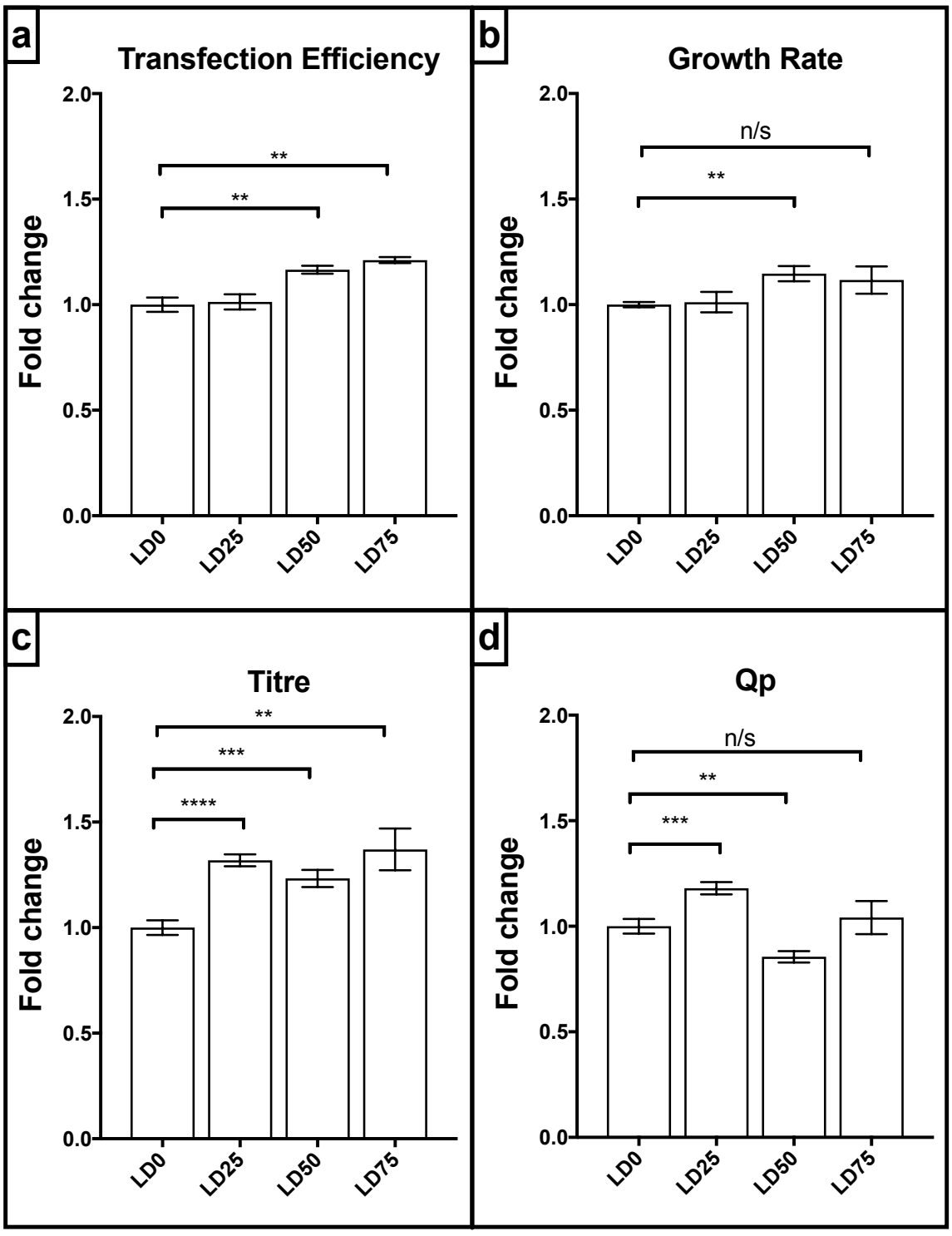


**Figure 6-7. Co-transfection of dual-selected CHO cells with DTE-Mab and GFP.**

*BFA and FLI-06 dual-selected cells were transiently transfected with plasmid DNA encoding DTE-Mab and GFP. After four days' growth cultures were sampled for growth, productivity and transfection efficiency data. Data is shown as a fold change of the control (LD0) data for [a] transfection efficiency, [b] growth rate, [c] overall titre and [d] specific productivity. For growth/viability, columns show the mean average of three biological replicates (four for control). For titre and Qp, columns show mean of three technical replicates of each of the biological replicates (n = 9 for test data; n = 12 for control data). All error bars show SEM.*

Titre results show that all three selection concentration levels see a statistically significant increase in DTE-Mab titre when compared to the control (figure 6-7 c). Whilst LD25 and LD75 cells see a 32% and 37% increase in titre from the control respectively, the LD50 cells only see a 23% increase in overall titre. When taking this data alongside that of cell growth, the LD75 cells see no increase in Qp compared to the control. This is in contradiction to the LD25 cells which see a statistically significant Qp increase of 18%, whilst LD50 cells see a significant decrease in Qp to 86% of the level seen in the control cells (figure 6-7 d).

In this short-term dual selection strategy, altering the concentrations of the two inhibitory agents has no effect upon overall productivity. With LD50 cells slightly out-growing LD75 cells post-transfection it is feasible that there is an ideal level of the inhibitory cocktail between these two LD values that may instigate a maximal growth rate response. That transfection efficiency is higher in the more harshly filtered cells than the control/LD25 cells further highlights that treating cells with secretory inhibitors does not negatively impact upon transfection efficiency. However it is not known why this process may result in increased transfection efficiency after four days' growth, especially with the growth rate of better-transfected cells increases slightly. This may show that increased levels of GFP remaining within the cell is not an artifact of a reduction in cell growth rate and metabolic turnover.

### 6.4.3. Stability analysis of short-term dual-selected CHO cells.

As in section 6.2.3, the stability of the selected cells was tested to determine to what level resistance to the BFA/FLI-06 cocktail persevered after cell banking, revival and six passages without the presence of selective pressure. A kill curve was carried out to test all biological replicates in duplicate at untreated, LD50 and LD75 concentrations of the dual inhibitors. Kill curves were carried out in a 96 deep-well plates with VCD measured by PrestoBlue® at the start and after three days' growth to calculate growth rate (sections 2.3.2; 2.1.2.2; 2.1.5.1). Growth data is shown in Figure 6-8 as a fold-change of the requisite cell line grown without the presence of the inhibitor mixture.

Treatment with a LD50 level of the inhibitor cocktail resulted in a severe reduction of growth rate and IVCD change in the LD25 and LD0 cell lines. Treatment of the LD75 cells results in a reduction in IVCD change over three days to 73% of that of untreated LD75 cells. When treated with an LD75 level of the inhibitor mixture the IVCD of LD75 cells reduces to 64% of the untreated cells. A similar pattern is seen in the LD50 cells, with a larger decrease in IVCD change seen when these cells are treated with the LD75-level of the inhibitor mixture, showing that both cell lines are unable to work at the untreated level when treated with the LD75 level of the inhibitor mixture.

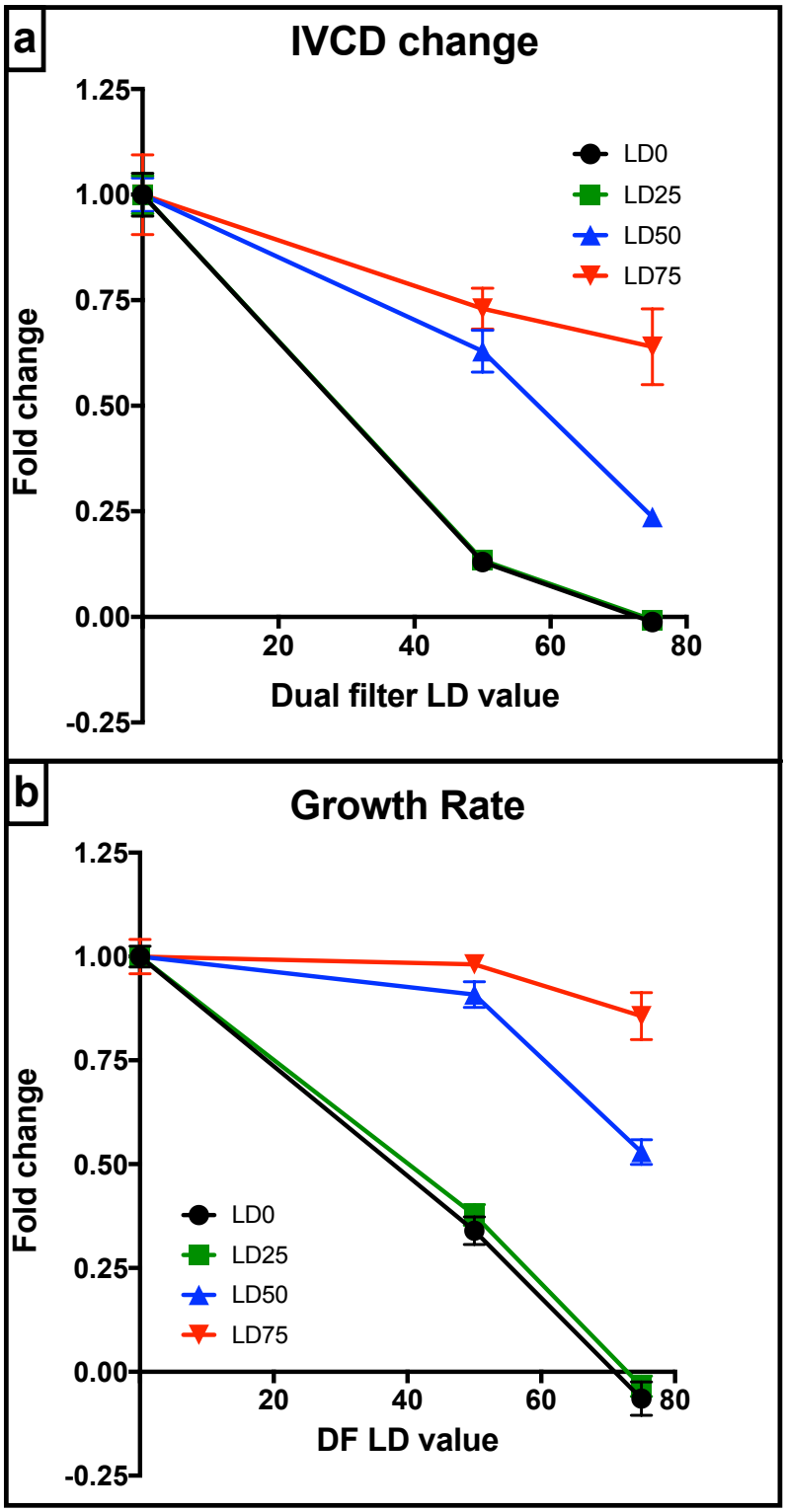


**Figure 6-8. Kill curves of BFA/FLI-06 short-term dual-selected CHO cells.**

*After six passages of growth from vial, without the presence of BFA/FLI-06, the resistance of dual-selected cells to the BFA/FLI-06 cocktail was determined by kill curve at LD0, LD50 and LD75 levels of the inhibitor cocktail. Cells were grown for three days in a 96 deep-well shaken plate with start- and end-point VCD samples taken by PrestoBlue® analysis so [a] IVCD change over three days' growth and [b] growth rate could be determined across all cell lines and BFA concentrations. Each data point shows the average of a maximum of two technical replicates of three biological replicates (n = 6). Error bars show SEM*

This data suggests that the cells have not maintained full resistance to the drug mixture at the levels used (figure 6-8 a). Growth-rate levels of the LD50 and LD75 cell lines also show that resistance to the BFA/FLI combination has decreased somewhat. At the LD50 concentration the growth of the LD50 cell line is 90% of the same cell line with no drug treatment, decreasing to 53% when in the presence of LD75 of the inhibitor cocktail. Growth rate of the LD75 cell line is barely effected by LD50 levels of the two compounds, though growth rate decreases to 85% of the untreated levels at a LD75 dose of the compound cocktail (figure 6-8 b). Whilst these are not large differences they do show that some of the resistance to both compounds in combination has been lost.

## 6.5. Transcriptomic and proteomic analysis of short-term chemical-selected cells.

As with fully evolved cell lines (section 5.2.5.2), short-term selected cells underwent basic proteomic and transcriptomic analysis to better decipher the effect the short-term selection process had upon CHO cells. RNA and whole cell protein samples were taken from LD75 cells from both the BFA- and BFA/FLI-06 dual-selected cell lines for analysis by RT-PCR and Western blotting respectively. The same techniques and protocols were used to quantitate BiP protein levels and the mRNA transcript abundance of BFA targets GBF1, ArfGef1 and ArfGef2, and the UPR markers ATF6, BiP and XBP1. The procedures used are the same as used in section 5.2.5, as described in 2.7.3 and 2.8.

There is no statistical difference in the level of BiP protein seen in the BFA LD75 selected cells compared to the control cells (figure 6-9), suggesting the short-term selection process has had little effect upon the cell's UPR response, as was previously seen with the BFA-evolved cell line (figure 5-15). This suggests that the BFA short-term selection and evolution processes were not too dissimilar in their outcome with regards to the CHO cell response.

However, in the dual-selection cells, whilst the BiP band intensity appears to be greater than that of the other cell lines (figure 6-9 a), once the intensity data is normalised for β-actin housekeeping levels there is a significant decrease in BiP protein level to 64% of that seen in the control (figure 6-9 b). This suggests that cells treated with both secretion-inhibiting compounds the ER stress response and UPR has been knocked down. β-actin was used as a housekeeping gene as it is historically used as such and antibodies against it are readily available. However more recent studies have suggested that it may actually display divergent expression levels that make it unsuitable for this role. As such this data may be somewhat skewed by use of β-actin as a housekeeping gene instead of a gene selected from a wide-ranging CHO proteomic study (Brown et al. 2018).

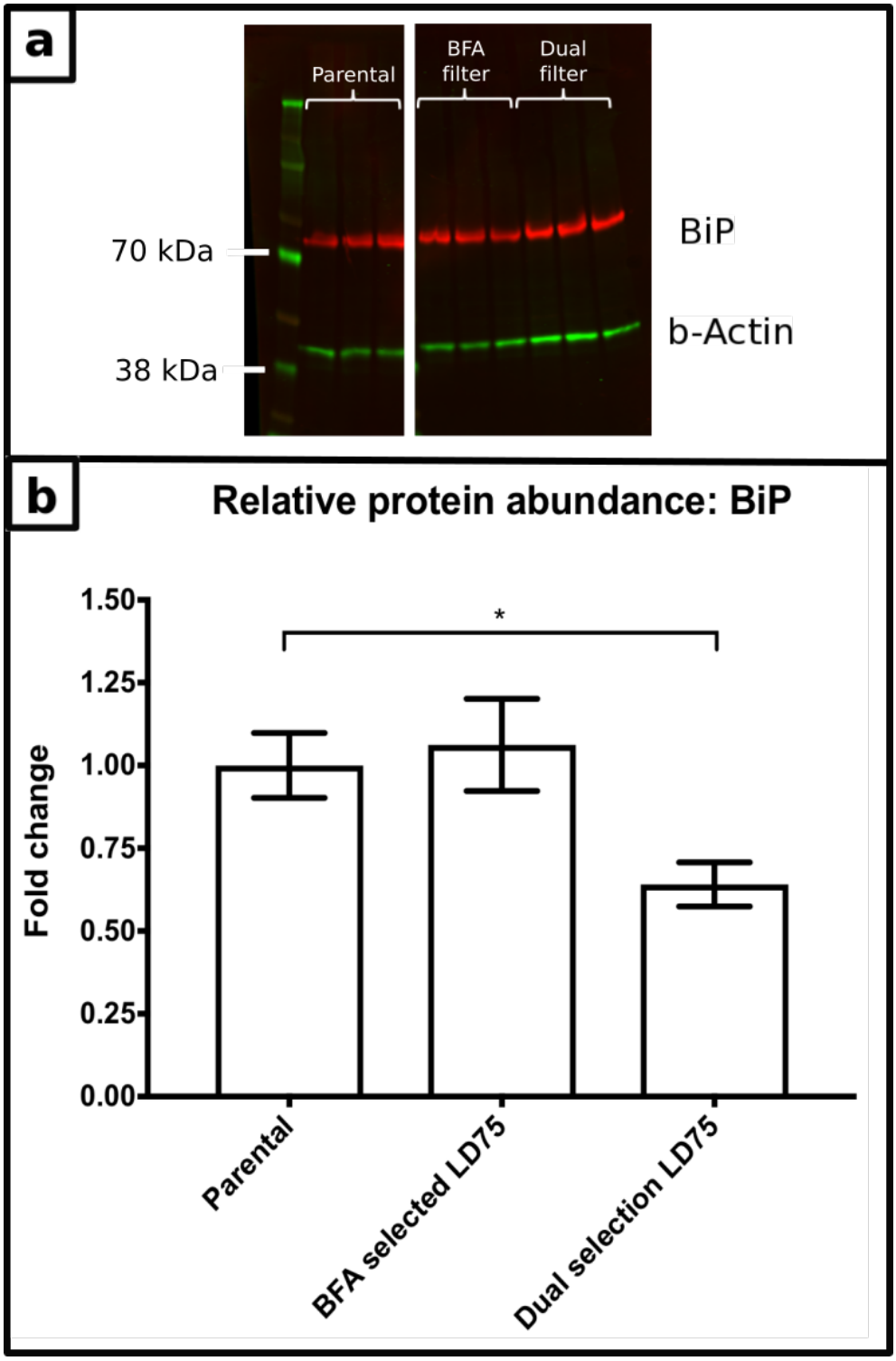


**Figure 6-9. Western blot analysis of BiP levels in chemically selected CHO cells.**

*The UPR status of MedI-CHO and BFA- and FLI-06-filtered CHO cell lines were analysed by quantitative western blotting of the UPR marker BiP, with β-actin was used as a housekeeping control. [a] Protein bands at the requisite sizes were seen for BiP (72.4 kDa) and β-actin (41.7 kDa). Image is indicative of technical replicates [b] Image Studio software was used to quantify band intensity. Protein samples were initially normalised for cell number, with BiP intensity values further normalised to the value of the housekeeping gene before calculation of the protein abundance level as a fold-change compared to MedI-CHO control cells. Columns show the mean average of two technical replicates of three biological replicates (n = 6). Error bars show SEM.*

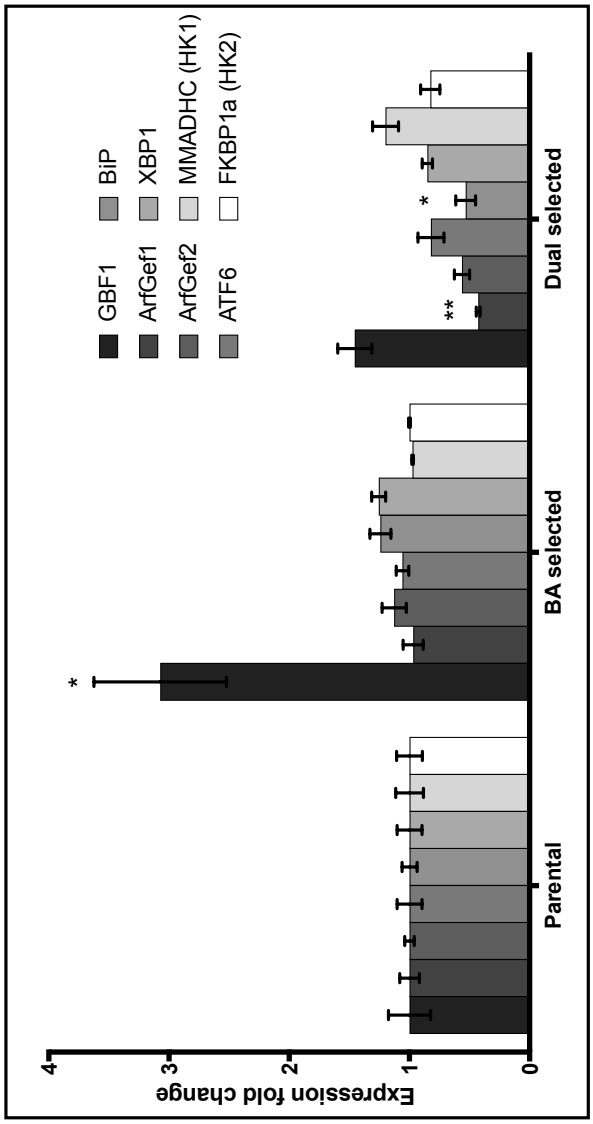


**Figure 6-10. Relative transcript abundance of BFA target and UPR-marker genes in short-term chemically selected CHO cells.**

*The transcript levels of UPR-related genes (ATF6, BiP and XBP1) and BFA targets (GBF1, ArfGef1, ArfGef2) in different CHO cell lines were analysed by RT-PCR. MMADHC and FKBP1a were used as housekeeping genes and positive controls. Protocol and analysis was the same as figure 5-16. Error bars show SEM. Statistical significance was calculated by a Dunnett's multiple comparisons test.*

RT-PCR analysis, taking into account housekeeping gene values for normalisation, shows that the relative transcript abundance of GBF1 in BFA-filtered cells is 3.07-fold greater than in the control. Whilst ArfGef1 and ATF6 levels stay stable compared to the control, ArfGef2 levels increase slightly (1.12-fold), with XBP1 and BiP levels seeing a larger increase (both a 1.25-fold increase). However none of these changes are statistically significant (figure 6-10).

Cells that have undergone the dual BFA/FLI-06 selection see a non-significant 1.46-fold increase in GBF1 levels. However, all other test genes see a decrease in transcript abundance level, with levels of ArfGef1 (43% of control) and BiP (53% of control) both being statistically significant. This transcriptomic data, taken alongside the productivity data, suggests that both a slight upregulation or down-regulation of the CHO UPR stress response can result in a slight increase in CHO productivity (figures 6-4, 6-7 and 6-10).

Although the shortened chemical exposure time of the selection strategy should limit the opportunity for resistance through mutation to occur, this outcome is still a possibility. As in section 5.2.5.3, an attempt was made to sequence the Sec7 domain of the GBF1 gene to determine whether a mutation may have resulted in the BFA resistance of the filtered cell lines. However, the PCR and product purification steps produced too little PCR product for sequencing to be carried out. Repetition of the experiment was planned but could not be preformed due to time constraints.

## 6.6. Discussion and conclusion.

Data from both single (BFA only) and dual (BFA/FLI) short-term selection experiments suggests that secretory blocking agents can be used to filter a population of CHO cells, resulting in a population selected for an increased productivity phenotype, resulting in a small increase in titre and Qp. However, that there is no stepwise increase in overall titre or Qp across both selection experiments (i.e. productivity increases as the LD value increases) suggests that the selection effect is not necessarily as directly linked to the concentration of the secretory inhibitor used in the selection process. This may be due to some flexibility being allowed in the selection process to ensure cells could recover growth after chemical treatment.

Whilst the overall outcome is somewhat satisfactory it does appear to be a somewhat trial-and-error process. Whilst the short-term selection process has a shorter duration than the evolutionary process (approximately four weeks compared to 8-12 weeks) it may be feasible to fit in to an early stage of a cell line chassis development. However it is unlikely to be suitable for use when looking to optimise the production of a specific Mab.

Overall Mab and Mab-GFP transfection data shows that the treatment of CHO cells with BFA does not reduce transfection efficiency (figure 6-4 a). However, these cells do still see a reduction in cell growth to approximately 60-70% the level of untreated control cells after transfection, despite there being no difference during normal cell growth. This would suggest that the selection process has brought about a change within the cell that favours the productivity of a recombinant Mab over cell growth.

That the reduction in cell growth in DTE-Mab and GFP co-transfected cells was less than that seen in DTE-Mab-only transfected cells (figures 6-3 a, 6-4 b) also points towards this as DTE-Mab-only transfected cells were transfected with a higher total level of Mab-encoding plasmid than the co-transfected cells. As such it can be suggested that the Mab-only cells were producing more Mab than the co-transfected cells. As such these cells had more Mab to process and so directed more cellular resources to this (and away from cell growth) than the co-transfected cells. To better compare the productivity of the differently selected cell lines stable pools could be produced from each cell line to better analyse the productivity of each cell line of a variety of model recombinant proteins, and how this productivity impacts upon cell function.

That there was no reduction in growth level in cells that had been dual-selected is somewhat contradictory to what was seen in the single filter experiment. There is also no obvious reason why the selection process has resulted in an increase in the transfection levels of CHO cells (figure 6-7 a, b). It is interesting that the dual selected cells do not retain resistance to the inhibitory cocktail to the same level seen in cells that have only been selected against one secretory-blocking chemical (figure 6-8). Perhaps the increased stress put upon the dual-selected cells by two inhibitory compounds results in a genetically more unstable population that will therefore drift – potentially back towards the 'normal' quicker than an unselected population.

As more of the secretory pathway is being targeted by the dual selection it could also prove more efficient at selecting out the high secreting phenotypes. However there is some opportunity for the two compounds to interact with each other, which could limit their function or impact upon CHO cell function by an unexpected interaction. However, that the dual and single BFA short-term selections result in a similar titre increase (1.4-fold for both), but differing Qp values (slight reduction and three-fold increase respectively; figures 6-3, 6-7), this suggests that the selection process differs between the two compound mixtures.

This again highlights that the short-term selection process is somewhat unpredictable. A scientific decision can be made as to a pathway to target (and the compound used to target it), but with the short-term selection process being

performed on biological entities it is difficult to predict exactly how they will respond. However, all experiments were repeated with three biological replicates across which there was little variance, suggesting that the results are replicable.

Comparing proteomic and transcriptomic data of BFA-selected and BFA-evolved CHO cells suggests that there is little change in the transcription or protein abundance of the genes/proteins analysed and that therefore the two cell lines are transcriptionally similar (figures 5-15, 5-16, 6-9 and 6-10). However, despite this there is still a difference in the productivity between the two different treatment types, with BFA-evolution resulting in reduced transient production whilst BFA short-term selection resulted in an enhancement (figures 5-6, 6-3, 6-7). With little difference in transcriptomic levels this points towards any changes within the cell being at a proteomic level.

However, changes at an RNA transcript or proteomic level are likely to take longer to occur than the duration of a short-term selection process (Majors et al. 2009). The extended exposure of evolved cells to BFA when compared to the short exposure time of the chemical selection suggests that a BFA-resistant mutation may have persevered within the population and, whilst an increase in GBF1 transcription levels was seen, this mutation ensured the same change was not required at a proteomic level. This could be better confirmed by sequencing of the GBF1 Sec7 domain to confirm any mutation, or proteomic analysis of GBF1 levels within the BFA-evolved and -filtered populations.

Short-term selection of CHO cells with both BFA and FLI-06 resulted in a much more marked difference in transcriptomic profile compared to BFA-only treated cells (figure 6-10). This could be explained by the targeting of two sections of the secretory pathway, although this is not entirely clear. A FLI-06 only selected cell line would need to be produced for this, whilst comparison of UPR marker transcript abundance of this, FLI-06-evolved cells and the dual selection would be required to allow a firm conclusion to be made. However it is again worth noting that, despite the differences in (an albeit limited) transcriptomic profile, FLI-06-selected cells saw the same increase in DTE transient production as the BFA-selected cells.

Overall, this data suggests that short-term chemical selection could be used as a method to enhance productivity in transient culture, which can be used in biologic production (Daramola et al. 2014). If the secretory pathway has indeed been enhanced then the improvement in titre level should be seen across the range of model products as the secretory pathway sections targeted are constitutive, not specific. Testing the effect the short-term selection process has upon stable production would provide information as to whether the selection strategy is suitable for production of a chassis CHO cell line for stable production.

Beyond chassis development, performing short-term chemical selection on a cell line stably expressing a model recombinant protein would allow an insight into whether the selection process can have the same effect when a cell is already producing a recombinant molecule. This would also allow better comparison between the single/multi-gene engineering and evolution/selection approaches to cell engineering , as was discussed in chapter 4.

However, that variations in transcriptomic data showed little change in relative productivity levels (see BFA-selected cells compared to dual selected cells, figures 6-3, 6-7, 6-10), whilst elsewhere similarities in transcriptomic data did not result in similar productivity levels (compare transient productivity of BFA evolved cells – figure 5-6 - to BFA filtered cells – figure 6-3) does not allow an easy resolution on the short-term selection and evolution processes to be made. Whilst there are some productivity benefits (as well as those of time) to a short, sharp filtering process over a long-term evolution, I would hypothesise that to a certain extent this depends upon the initial cell population and can not be as well designed and predicted as was initially thought.

# 7) General conclusions and further work

*Major observations from the thesis are briefly listed. In depth conclusions to the work within this thesis are made at the end of each chapter. This section brings together the work across the entire project by briefly summarising it and highlighting overall trends. The work's place within the overall literature is discussed and suggestions for future directions made.*

## 7.1. Major observations

### 7.1.1. Gene target selection (chapter 3)

- There are few, if any, obvious changes in the transcriptomic profile of the biosynthetic and secretory pathway of the CHO cell when comparing producing and non producing cells, or cells with high-or low- production levels.
- This differs to the plasma cell, where there are many large changes in transcriptional levels seen when comparing naïve B cells with the terminally differentiated, high-producing plasma cell.

### 7.1.2. Genetic engineering of CHO secretory and biosynthetic pathways (chapter 4)

- Ectopic expression of single genes targeting specific areas of the CHO secretory pathway had little effect upon CHO cell productivity of a transiently expressed DTE-Mab.
- Expression of transcription factors that cut across several regions of the CHO secretory and biosynthetic pathways can increase the productivity of CHO cells transiently expressing a DTE-Mab.
- Combination of up to three genes targeting a specific region of the CHO secretory pathway is not sufficient to enhance CHO cell productivity.
- Combination of globally-effective transcription factors can have a concomitant effect upon CHO cell transient productivity.
- Ectopic expression of transcription factors in a stably-producing CHO cell pool does not have the same effect as seen in transiently-expressing CHO cells. This may be due to the chronic stress stable expression has 're-setting' the UPR of CHO cells, making them impervious to UPR engineering.

### 7.1.3. Directed evolution and chemical selection of CHO cells with secretory blocking compounds (chapters 5 and 6)

- Hamster-derived cell lines react differently to secretory blocking agents that function in other mammalian cell lineages.
- Long-term treatment/evolution of CHO cells with secretory blocking compounds did not result in an increased secretory phenotype.
- Long-term treatment/evolution of CHO cells against Brefeldin A resulted in a CHO cell line with a reduced productivity phenotype.
- Short-term treatment of CHO cells with chemical secretory pathway blockers resulted in a population of cells with an increased productivity phenotype.

## 7.2. Overall conclusions

The overall aim of this project and thesis was to engineer the secretory pathway of the CHO cell, removing secretory bottlenecks and leading to an enhancement of productivity levels. CHO cells are widely used in the manufacture of biologics due to being well defined but genetically malleable; easy to culture; able to produce human-like protein modifications; and their regulatory acceptance. However they are by no means a Mab producing and secretion specialised cell.

By enhancing the CHO secretory pathway it was hypothesised that the titre and specific productivity of model recombinant biologics would be enhanced. Furthermore, the engineering strategies used might enhance the capacity of CHO cells to produce the increasing array of more difficult to express biologics (e.g. bispecifics, fusion proteins) that are proliferating within the pipelines of biopharmaceutical companies.

This project builds upon genetic engineering strategies that have previously been used to enhance CHO cell productivity. There are many examples within the literature of direct genetic engineering approaches increasing CHO productivity. Whilst there have been examples secretory pathway engineering these have only targeted a single section of the secretory pathway. Here, using high-throughput techniques, a large number of genes targeting many different sections of the CHO biosynthetic and secretory pathways could be screened. Novel targets were discovered though utilisation of plasma cell transcriptomic data, with some increasing CHO productivity levels. Enhancement of the secretory pathway through directed evolutionary and selection processes was performed, a method that has yet to be fully investigated in the literature. This has lead to the novel strategy of a short-term chemical selection, utilising secretory blocking compounds and exploiting the heterogeneous nature of the cell population to bring about phenotypic change rather than trying to instigate ‘omic changes.

### 7.2.1. Single- and multi-gene engineering of CHO cells

In chapters 1 and 3 prior knowledge from the literature was combined with transcriptomic and proteomic data from CHO cells to better understand the CHO biosynthetic and secretory systems. Comparison of 'omic data from high- and low-producing CHO cells (productivity difference being due to either cell line development and selection, different growth phases or different expression processes) showed there was little 'omic difference between high- and low-producing CHO populations when looking at genes/proteins involved in secretion.

However, transcriptomic data comparing naïve B cells to their terminally differentiated state, the Mab producing and high-secreting plasma cell, highlighted myriad changes in gene expression that are required to turn a low producing cell in to one specialising in high production and secretion levels. From 'omic data, general secretory pathway knowledge and previous CHO engineering strategies, a panel of genes were selected with which to engineer the CHO biosynthetic and secretory pathways.

Genetic engineering of the CHO cell with the selected gene targets was performed in chapter 4. The effect of each of the single genes at different dose levels upon CHO productivity of a transiently expressed DTE-Mab was tested. This showed that a few of the genes selected had a positive impact upon CHO productivity of a DTE-Mab. In general more global engineering targets, such as transcription factors, enhanced CHO cell productivity. Whilst ectopic expression of XBP1s has previously been shown to enhance CHO productivity, transcription factors upregulated in plasma cells – CREB3L2 and CRELD2 – were shown to also enhance CHO productivity, albeit not to the same level as XBP1s.

Overexpression of some ER folding machinery components also slightly enhanced CHO productivity levels. Single gene data was used to inform multigene engineering strategies to further boost CHO productivity levels. Alongside a data-driven combination of transcription factors, knowledge-driven two/three gene combinations were also designed to target specific sections of the secretory pathway.

Expression of many genes selected to directly engineer specific sections of the biosynthetic and secretory pathways had little impact upon the level of CHO production of a transiently expressed DTE-Mab. However, effector genes which had a more global effect in enhancing transcription levels of many genes, had a concerted positive impact upon CHO cell productivity. With some effector genes CHO productivity was further enhanced upon expression of two transcription factors when compared to single-gene expression.

The single gene screen was expressed in a stably-expressing cell pool. The positive effects of some targets in a transiently-expressing cell line were not seen in a stably expressing pool. Whilst XBP1, CREB3L2 and CRELD2 all enhanced transient DTE-Mab expression levels, they had little effect upon production levels of a stably-expressed DTE-Mab.

We hypothesise that continuous expression of a recombinant protein, especially one considered difficult to express, generates a chronic ER stress response. This can perpetuate feedback suppression of proteins involved in unfolded protein response activation. This may explain the lack of response of stably expressing pools to the expression of transcription factors that are understood to function within the UPR.

Results from direct genetic engineering approaches show that directed single- and multi-gene engineering of complicated molecular pathways is generally not suitable for enhancing CHO productivity. There are many molecular targets involved in different areas of the biosynthetic secretory pathway, whilst only a small number can feasibly be used to engineer the cell. As such the overexpression of only two or three genes can cause stoichiometric imbalances reducing molecular function, possibly introducing more bottlenecks within the cell.

However, use of more global targets such as transcription factors that cut across the many different sections of the CHO biosynthetic and secretory processes can perpetuate wider transcriptional changes, resulting in the cell-driven widening of the targeted pathway as a whole. This approach is more likely to remove pathway bottlenecks and perpetuate an overall increase in productivity.

Interestingly sole expression of some targets previously shown to enhance CHO productivity (e.g. CERT, YY1) did not have the same effect within this gene screen. This highlights that the differences between different CHO cell lines, model recombinant proteins and expression systems can have all impact upon a gene's effect upon cell productivity. This again highlights that a directed genetic engineering solution for one problem may not necessarily solve another similar problem, as well as highlighting the importance of reproducibility between research and manufacturing environments to ensure processes are reliable and reproducible.

### 7.2.2. Directed evolution and short-term chemical selection of CHO cells

In chapters 5 and 6 two different directed evolution strategies were used to try and enhance CHO productivity levels. It was hypothesised that blocking of a section of the secretory pathway would perpetuate myriad transcriptomic and proteomic changes within the cell so as to overcome the blockage, with these

changes resulting in an enhanced secretory phenotype. A literature review was undertaken to select potential secretory blocking compounds that could be used.

The selected compounds were tested in CHO cells to confirm they impacted upon cell growth and secretory pathway morphology. Two compounds were selected and shown to have an effect upon CHO cells: Brefeldin A inhibits vesicle formation at the Golgi for both anterograde and recycling transport; FLI-06 inhibits cargo loading at ER exit sites. Host CHO cells underwent separate evolution against both compounds, with the concentration of both being increased step-wise throughout the evolutionary process. The effectiveness of the directed evolution strategy was determined by assaying the productivity of resultant cell lines through transient and stable transfection with a difficult- and easy-to-express model Mabs.

In chapter 5 it was hypothesised that blocking of the secretory pathway at a specific point for an extended period of time would result in either the upregulation of molecular machinery within the region targeted, and/or enhancement of other pathways to overcome the blockage. Both outcomes would result in an increase in cellular secretory capacity upon removal of the selection agent, increasing the flux of protein transport and producing an enhanced productivity phenotype. However, evolution of CHO against Brefeldin A resulted in a decrease in both transient and stable productivity levels of two model Mabs.

Whilst evolution against FLI-06 resulted in an increase in transient Mab titres, it was not possible to collect representative data regarding stable Mab expression. From the Brefeldin A evolution results it was hypothesised that a long-term exposure to increasing levels of this selective agent brought about mutations within the compound's target pathway. These produced resistance to Brefeldin A, perpetuating a Brefeldin A resistance phenotype without the hypothesised concomitant secretory enhancements. As such it was suggested that a long-term exposure to an inhibitory compound is more likely to bring about resistance within a cell line through mutations of a small number of targets rather than the more global upregulation of pathway components hypothesised.

The inability of the BFA evolutionary process to instigate an increased secretory and productivity phenotype led to the hypothesis that a short-term treatment of CHO cells with a high concentration of a secretory inhibitor could act as a "chemical selection". This would have the effect of attenuating the growth of cells with a low secretory phenotype, allowing cells with a high secretory phenotype to pervade within the population. Rather than trying to change the 'omic profile of CHO cells as attempted by directed evolution, the short-term selection process would take advantage of the natural heterogeneity within the cell population in selecting for higher producers.

This hypothesis was tested in chapter 6 where host CHO cells were treated with a single concentration of inhibitor until cell growth recovered. Two short-term chemical selections were tested – a BFA-only selection and a dual selection where BFA and FLI-06 were used in tandem. The effectiveness of the two short-term selection processes was tested by transient expression of a DTE Mab. This showed that an overall increase in Qp was seen in CHO cells treated with a high dose of inhibitory compound/cocktail for a short period of time. However the post-transfection growth rate was reduced in the selected cells, resulting in a smaller overall titre increase being seen.

## 7.3. Future directions

Within the scope of this project there are still some lose ends that could be tied up. Firstly the stable productivity levels of the FLI-06-evolved cell line, as well as the short-term chemically selected cell lines, would need to be assayed. This was attempted in the thesis but cell productivity levels varied too much for a conclusion to be made. Use of TEE technology would normalise Mab transcription levels across test and control cell lines, providing clearer data on the productivity of the FLI-06 evolved and short-term selected cell lines. Whilst transient expression studies have suggested that chemically selected cell lines have an increased Qp, production of stable pools would provide more conclusive information as to the effectiveness of this process in increasing productivity, both from a research and industry viewpoint.

Stable expression of multiple genes – encoding both effector and product - from a single vector could simplify gene engineering process. Controlling component levels by placing them under the transcriptional control of synthetic promoters would also allow tunable expression of genes. This would allow tighter control of product and effector levels, providing a process by which processing load and stoichiometry within the cell could be closely controlled. This could limit the amount of overloading and stoichiometric side-effects upon the CHO biosynthetic and secretory pathways which may well come about from gene overexpression.

Further work is required to better understand the transcriptomic and proteomic changes that have produced the phenotypes of the evolved and short-term selection produced cell lines. Sequencing of the Sec7 region of the BFA target GBF1 was attempted but unsuccessful. This information would prove whether the BFA resistance of BFA-evolved cells is due to a point mutation and not a general overcoming of the secretory blockage. Further proteomic work – such as Western blotting – would provide quantifiable information on the protein levels of BFA targets, as well as UPR effectors, allowing a deeper insight into the effects evolutionary and filtering processes have had upon the cell. This was performed

but analysis of GBF1 levels was unsuccessful. The application and effect of evolutionary and filtering processes still need to be defined to ensure they are not a trial-and-error approach to cell line development. Evolution of a stably-expressing cell line could be tested to better determine the step of cell line development at which these processes could be best deployed.

Further to the scope of this project there are other methods by which the CHO secretory pathway could be engineered. Better CHO-relevant secretory pathway maps could be produced by utilising gene knock-down and screening technologies such as siRNA. This could better inform the selection of gene and pathway engineering targets. Further to this, the knock-down of host cell proteins that undergo constitutive secretion, but are not required for the specific function of the CHO cell, could free up space within the CHO secretory pathway for recombinant proteins. The reduction in secretory cargo could reduce the pressure upon secretory bottlenecks, as well as easing the process of product purification.

Furthermore, the use of genome editing tools such as CRISPR could be used to engineer endogenous genes. For example the XBP1 (ensuring it is always within a spliced conformation), CREB3L2 (ensuring only its transcriptionally active cytosolic domain is present) or CERT (engineering the phosphorylation site that determines its functionality such that it can always transfer ceramide) genes used in the project were designed to be transcriptionally active. Modification of the endogenous gene by CRISPR would allow genes to be in a more active confirmation without altering stoichiometry levels or increasing the transcriptomic burden upon the cell through expression of recombinant genes.

# 8) Bibliography


Abbott WM, Middleton B, Kartberg F, Claesson J, Roth R, Fisher D. 2015. Optimisation of a simple method to transiently transfect a CHO cell line in high-throughput and at large scale. Protein Expression and Purification 116:113-119.

Aggarwal SR. 2014. What's fueling the biotech engine-2012 to 2013. Nature Biotechnology 32(1):32-39.

Ahmad M, Hirz M, Pichler H, Schwab H. 2014. Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production. Applied Microbiology and Biotechnology 98(12):5301-5317.

Al-Fageeh MB, Marchant RJ, Carden MJ, Smales CM. 2006. The cold-shock response in cultured mammalian cells: Harnessing the response for the improvement of recombinant protein production. Biotechnology and Bioengineering 93(5):829-835.

Allan BB, Moyer BD, Balch WE. 2000a. Rab1 recruitment of p115 into a cis-SNARE complex: Programming budding COPII vesicles for fusion. Science 289(5478):444-448.

Allan BB, Weissman J, Aridor M, Moyer B, Chen CD, Yoo JS, Balch WE. 2000b. State-specific assays to study biosynthetic cargo selection and role of SNAREs in export from the endoplasmic reticulum and delivery to the Golgi. Methods-a Companion to Methods in Enzymology 20(4):411-416.

Allan VJ, Thompson HM, McNiven MA. 2002. Motoring around the Golgi. Nature Cell Biology 4(10):E236-E242.

Almo SC, Love JD. 2014. Better and faster: improvements and optimization for mammalian recombinant protein production. Current Opinion in Structural Biology 26:39-43.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. BASIC LOCAL ALIGNMENT SEARCH TOOL. Journal of Molecular Biology 215(3):403-410.

Alvarez C, Garcia-Mata R, Hauri HP, Sztul E. 2001. The p115-interactive proteins GM130 and giantin participate in endoplasmic reticulum-Golgi traffic. Journal of Biological Chemistry 276(4):2693-2700.

Anders N, Juergens G. 2008. Large ARF guanine nucleotide exchange factors in membrane trafficking. Cellular and Molecular Life Sciences 65(21):3433-3445.

Anderson MJ, Whitcomb PJ. 2007. DOE Simplified second edition: Practical Tools for Effective Experimentation.

Aoki R, Kitaguchi T, Oya M, Yanagihara Y, Sato M, Miyawaki A, Tsuboi T. 2010. Duration of fusion pore opening and the amount of hormone released are regulated by myosin II during kiss-and-run exocytosis. Biochemical Journal 429:497-504.

Appenzeller C, Andersson H, Kappeler F, Hauri HP. 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. Nature Cell Biology 1(6):330-334.

Appenzeller-Herzog C, Hauri HP. 2006. The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. Journal of Cell Science 119(11):2173-2183.

Arakel EC, Schwappach B. 2018. Formation of COPI-coated vesicles at a glance. Journal of Cell Science 131(5).

Aridor M. 2018. COPII gets in shape: Lessons derived from morphological aspects of early secretion. Traffic (Copenhagen, Denmark).

Baik JY, Dahodwala H, Oduah E, Talman L, Gemmill TR, Gasimli L, Datta P, Yang B, Li G, Zhang F and others. 2015. Optimization of bioprocess conditions improves production of a CHO cell-derived, bioengineered heparin. Biotechnology Journal 10(7):1067-1081.

Baines AC, Zhang B. 2007. Receptor-mediated protein transport in the early secretory pathway. Trends in Biochemical Sciences 32(8):381-388.

Baker RW, Hughson FM. 2016. Chaperoning SNARE assembly and disassembly. Nature Reviews Molecular Cell Biology 17(8):465-479.

Baldi L, Hacker DL, Adam M, Wurm FM. 2007. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. Biotechnology Letters 29(5):677-684.

Bao H, Das D, Courtney NA, Jiang Y, Briguglio JS, Lou X, Roston D, Cui Q, Chanda B, Chapman ER. 2018. Dynamics and number of trans-SNARE complexes determine nascent fusion pore properties. Nature 554(7691):260-+.

Barbosa S, Fasanella G, Carreira S, Llarena M, Fox R, Barreca C, Andrew D, O'Hare P. 2013. An Orchestrated Program Regulating Secretory Pathway Genes and Cargos by the Transmembrane Transcription Factor CREB-H. Traffic 14(4):382-398.
Bard F, Malhotra V. 2006. The formation of TGN-to-plasma-membrane transport carriers. Annual Review of Cell and Developmental Biology 22:439-455.
Barlowe C, Helenius A. 2016. Cargo Capture and Bulk Flow in the Early Secretory Pathway. Annual Review of Cell and Developmental Biology, Vol 32 32:197-222.
Barlowe C, Orci L, Yeung T, Hosobuchi M, Hamamoto S, Salama N, Rexach MF, Ravazzola M, Amherdt M, Schekman R. 1994. COPII - A MEMBRANE COAT FORMED BY SEC PROTEINS THAT DRIVE VESICLE BUDDING FROM THE ENDOPLASMIC-RETICULUM. Cell 77(6):895-907.
Barlowe C, Schekman R. 1993. SEC12 ENCODES A GUANINE-NUCLEOTIDE-EXCHANGE FACTOR ESSENTIAL FOR TRANSPORT VESICLE BUDDING FROM THE ER. Nature 365(6444):347-349.
Baron CL, Malhotra V. 2002. Role of diacylglycerol in PKD recruitment to the TGN and protein transport to the plasma membrane. Science 295(5553):325-328.
Barr FA, Nakamura N, Warren G. 1998. Mapping the interaction between GRASP65 and GM130, components of a protein complex involved in the stacking of Golgi cisternae. Embo Journal 17(12):3258-3268.
Barron N, Kumar N, Sanchez N, Doolan P, Clarke C, Meleady P, O'Sullivan F, Clynes M. 2011a. Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. Journal of Biotechnology 151(2):204-211.
Barron N, Sanchez N, Kelly P, Clynes M. 2011b. MicroRNAs: tiny targets for engineering CHO cell phenotypes? Biotechnology Letters 33(1):11-21.
Barrowman J, Bhandari D, Reinisch K, Ferro-Novick S. 2010. TRAPP complexes in membrane traffic: convergence through a common Rab. Nature Reviews Molecular Cell Biology 11(11):759-763.
Baumann M, Gludovacz E, Sealover N, Bahr S, George H, Lin N, Kayser K, Borth N. 2017. Preselection of recombinant gene integration sites enabling high transcription rates in CHO cells using alternate start codons and recombinase mediated cassette exchange. Biotechnology and Bioengineering 114(11):2616-2627.
Bebbington CR, Renner G, Thomson S, King D, Abrams D, Yarranton GT. 1992. HIGH-LEVEL EXPRESSION OF A RECOMBINANT ANTIBODY FROM MYELOMA CELLS USING A GLUTAMINE-SYNTHETASE GENE AS AN AMPLIFIABLE SELECTABLE MARKER. Bio-Technology 10(2):169-175.
Beck R, Ravet M, Wieland FT, Cassel D. 2009. The COPI system: Molecular mechanisms and function. Febs Letters 583(17):2701-2709.
Becker E, Florin L, Pfizenmaier K, Kaufmann H. 2008. An XBP-1 dependent bottle-neck in production of IgG subtype antibodies in chemically defined serum-free Chinese hamster ovary (CHO) fed-batch processes. Journal of Biotechnology 135(2):217-223.
Beckmann H, Su LK, Kadesch T. 1990. TFE3 - A HELIX LOOP HELIX PROTEIN THAT ACTIVATES TRANSCRIPTION THROUGH THE IMMUNOGLOBULIN ENHANCER MU-E3 MOTIF. Genes & Development 4(2):167-179.
Ben-Tekaya H, Miura K, Pepperkok R, Hauri HP. 2005. Live imaging of bidirectional traffic from the ERGIC. Journal of Cell Science 118(2):357-367.
Benton T, Chen T, McEntee M, Fox B, King D, Crombie R, Thomas TC, Bebbington C. 2002. The use of UCOE vectors in combination with a preadapted serum free, suspension cell line allows for rapid production of large quantities of protein. Cytotechnology 38(1-2):43-46.
Bertolotti A, Zhang YH, Hendershot LM, Harding HP, Ron D. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nature Cell Biology 2(6):326-332.
Bethune J, Wieland FT. 2018. Assembly of COPI and COPII Vesicular Coat Proteins on Membranes. Annual review of biophysics.
Betts Z, Dickson AJ. 2015. Assessment of UCOE on Recombinant EPO Production and Expression Stability in Amplified Chinese Hamster Ovary Cells. Molecular Biotechnology 57(9):846-858.
Betts Z, Dickson AJ. 2016. Ubiquitous Chromatin Opening Elements (UCOEs) effect on transgene position and expression stability in CHO cells following methotrexate (MTX) amplification. Biotechnology Journal 11(4):554-564.

Bi X, Mancias JD, Goldberg J. 2007. Insights into COPII coat nucleation from the structure of Sec23 center dot Sar1 complexed with the active fragment of sec31. Developmental Cell 13(5):635-645.
Binder JX, Pletscher-Frankild S, Tsafou K, Stolte C, O'Donoghue SI, Schneider R, Jensen LJ. 2014. COMPARTMENTS: unification and visualization of protein subcellular localization evidence. Database-the Journal of Biological Databases and Curation.
Birch JR, Racher AJ. 2006. Antibody production. Advanced Drug Delivery Reviews 58(5-6):671-685.
Bittner MA, Aikman RL, Holz RW. 2013. A Nibbling Mechanism for Clathrin-mediated Retrieval of Secretory Granule Membrane after Exocytosis. Journal of Biological Chemistry 288(13):9177-9188.
Boal F, Guetzoyan L, Sessions RB, Zeghouf M, Spooner RA, Lord JM, Cherfils J, Clarkson GJ, Roberts LM, Stephens DJ. 2010. LG186: An Inhibitor of GBF1 Function that Causes Golgi Disassembly in Human and Canine Cells. Traffic 11(12):1537-1551.
Bock JB, Matern HT, Peden AA, Scheller RH. 2001. A genomic perspective on membrane compartment organization. Nature 409(6822):839-841.
Boehm M, Bonifacino JS. 2002. Genetic analyses of adaptin function from yeast to mammals. Gene 286(2):175-186.
Boncompain G, Perez F. 2013. The many routes of Golgi-dependent trafficking. Histochemistry and Cell Biology 140(3):251-260.
Bonifacino JS, Glick BS. 2004. The mechanisms of vesicle budding and fusion. Cell 116(2):153-166.
Borth N, Mattanovich D, Kunert R, Katinger H. 2005. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. Biotechnology Progress 21(1):106-111.
Boyce M, Bryant KF, Jousse C, Long K, Harding HP, Scheuner D, Kaufman RJ, Ma DW, Coen DM, Ron D and others. 2005. A selective inhibitor-of eIF2 alpha dephosphorylation protects cells from ER stress. Science 307(5711):935-939.
Brewer JW, Diehl JA. 2000. PERK mediates cell-cycle exit during the mammalian unfolded protein response. Proceedings of the National Academy of Sciences of the United States of America 97(23):12625-12630.
Brown AJ, Gibson S, Hatton D, James DC. 2018. Transcriptome-Based Identification of the Optimal Reference CHO Genes for Normalisation of qPCR Data. Biotechnology Journal 13(1).
Brown AJ, James DC. 2016. Precision control of recombinant gene transcription for CHO cell synthetic biology. Biotechnology Advances 34(5):492-503.
Brown AJ, Sweeney B, Mainwaring DO, James DC. 2014. Synthetic Promoters for CHO Cell Engineering. Biotechnology and Bioengineering 111(8):1638-1647.
Brown FC, Schindelhaim CH, Pfeffer SR. 2011. GCC185 plays independent roles in Golgi structure maintenance and AP-1-mediated vesicle tethering. Journal of Cell Biology 194(5):779-787.
Browne SM, Al-Rubeai M. 2007. Selection methods for high-producing mammalian cell lines. Trends in Biotechnology 25(9):425-432.
Brownhill K, Wood L, Allan V. 2009. Molecular motors and the Golgi complex: Staying put and moving through. Seminars in Cell & Developmental Biology 20(7):784-792.
Burgess TL, Kelly RB. 1987. CONSTITUTIVE AND REGULATED SECRETION OF PROTEINS. Annual Review of Cell Biology 3:243-293.
Cacciatore JJ, Chasin LA, Leonard EF. 2010. Gene amplification and vector engineering to achieve rapid and high-level therapeutic protein production using the Dhfr-based CHO cell selection system. Biotechnology Advances 28(6):673-681.
Cai C, Rajaram M, Zhou X, Liu Q, Marchica J, Li J, Powers RS. 2012. Activation of multiple cancer pathways and tumor maintenance function of the 3q amplified oncogene FNDC3B. Cell Cycle 11(9):1773-1781.
Cai H, Reinisch K, Ferro-Novick S. 2007a. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. Developmental Cell 12(5):671-682.
Cai H, Yu S, Menon S, Cai Y, Lazarova D, Fu C, Reinisch K, Hay JC, Ferro-Novick S. 2007b. TRAPPI tethers COPII vesicles by binding the coat subunit Sec23. Nature 445(7130):941-944.
Cai Y, Chin HF, Lazarova D, Menon S, Fu C, Cai H, Sclafani A, Rodgers DW, De La Cruz EM, Ferro-Novick S and others. 2008. The structural basis for activation of the Rab Ypt1p by the TRAPP membrane-tethering complexes. Cell 133(7):1202-1213.

Cain K, Peters S, Hailu H, Sweeney B, Stephens P, Heads J, Sarkar K, Ventom A, Page C, Dickson A. 2013. A CHO cell line engineered to express XBP1 and ERO1-L has increased levels of transient protein expression. Biotechnology Progress 29(3):697-706.
Carlage T, Hincapie M, Zang L, Lyubarskaya Y, Madden H, Mhatre R, Hancock WS. 2009. Proteomic Profiling of a High-Producing Chinese Hamster Ovary Cell Culture. Analytical Chemistry 81(17):7357-7362.
Carlage T, Kshirsagar R, Zang L, Janakiraman V, Hincapie M, Lyubarskaya Y, Weiskopf A, Hancock WS. 2012. Analysis of dynamic changes in the proteome of a Bcl-XL overexpressing Chinese hamster ovary cell culture during exponential and stationary phases. Biotechnology Progress 28(3):814-823.
Carrouel F, Couble ML, Vanbelle C, Staquet MJ, Magloire H, Bleicher F. 2008. HUGO(FNDC3A): a new gene overexpressed in human odontoblasts. Journal of Dental Research 87(2):131-136.
Cartwright JF, Anderson K, Longworth J, Lobb P, James DC. 2018. Highly sensitive detection of mutations in CHO cell recombinant DNA using multi-parallel single molecule real-time DNA sequencing. Biotechnology and Bioengineering 115(6):1485-1498.
Celik E, Calik P. 2012. Production of recombinant proteins by yeast cells. Biotechnology Advances 30(5):1108-1118.
Chalmers F, van Lith M, Sweeney B, Cain K, Bulleid NJ. 2017. Inhibition of IRE1alpha-mediated XBP1 mRNA cleavage by XBP1 reveals a novel regulatory process during the unfolded protein response. Wellcome open research 2:36-36.
Charaniya S, Karypis G, Hu W-S. 2009. Mining Transcriptome Data for Function-Trait Relationship of Hyper Productivity of Recombinant Antibody. Biotechnology and Bioengineering 102(6):1654-1669.
Chardin P, Paris S, Antonny B, Robineau S, BeraudDufour S, Jackson CL, Chabre M. 1996. A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains. Nature 384(6608):481-484.
Chatellard P, Pankiewicz R, Meier E, Durrer L, Sauvage C, Imhof MO. 2007. The IE2 promoter/enhancer region from mouse CMV provides high levels of therapeutic protein expression in mammalian cells. Biotechnology and Bioengineering 96(1):106-117.
Cheung P-YP, Pfeffer SR. 2016. Transport Vesicle Tethering at the Trans Golgi Network: Coiled Coil Proteins in Action. Frontiers in cell and developmental biology 4:18-18.
Chi X, Wang S, Huang Y, Stamnes M, Chen J-L. 2013. Roles of Rho GTPases in Intracellular Transport and Cellular Transformation. International Journal of Molecular Sciences 14(4):7089-7108.
Chiba Y, Jigami Y, Kuroda K, Kobayashi K, Ichikawa K, Nonaka K, Suzuki T. 2012. Method for high-level secretory production of protein. Official Gazette of the United States Patent and Trademark Office Patents.
Chin CL, Chin HK, Chin CSH, Lai ET, Ng SK. 2015. Engineering selection stringency on expression vector for the production of recombinant human alpha1-antitrypsin using Chinese Hamster ovary cells. Bmc Biotechnology 15.
Chung BK-S, Yusufi FNK, Mariati, Yang Y, Lee D-Y. 2013. Enhanced expression of codon optimized interferon gamma in CHO cells. Journal of Biotechnology 167(3):326-333.
Chung C-Y, Wang Q, Yang S, Ponce SA, Kirsch BJ, Zhang H, Betenbaugh MJ. 2017. Combinatorial genome and protein engineering yields monoclonal antibodies with hypergalactosylation from CHO cells. Biotechnology and Bioengineering 114(12):2848-2856.
Chung JY, Lim SW, Hong YJ, Hwang SO, Lee GM. 2004. Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant Chinese hamster ovary cells. Biotechnology and Bioengineering 85(5):539-546.
Clarke C, Doolan P, Barron N, Meleady P, O'Sullivan F, Gammell P, Melville M, Leonard M, Clynes M. 2011. Large scale microarray profiling and coexpression network analysis of CHO cells identifies transcriptional modules associated with growth and productivity. Journal of Biotechnology 155(3):350-359.
Claude A, Zhao BP, Kuziemsky CE, Dahan S, Berger SJ, Yan JP, Armold AD, Sullivan EM, Melancon P. 1999. GBF1: A novel Golgi-associated BFA-resistant guanine nucleotide exchange factor that displays specificity for ADP-ribosylation factor 5. Journal of Cell Biology 146(1):71-84.
Cleves A, McGee T, Bankaitis V. 1991. Phospholipid transfer proteins: a biological debut. Trends in cell biology 1(1):30-4.

Cockett MI, Bebbington CR, Yarranton GT. 1990. HIGH-LEVEL EXPRESSION OF TISSUE INHIBITOR OF METALLOPROTEINASES IN CHINESE HAMSTER OVARY CELLS USING GLUTAMINE-SYNTHETASE GENE AMPLIFICATION. Bio-Technology 8(7):662-667.
Collen D, Lijnen HR. 2004. Tissue-type plasminogen activator: a historical perspective and personal account. Journal of Thrombosis and Haemostasis 2(4):541-546.
Collen D, Stassen JM, Marafino BJ, Builder S, Decock F, Ogez J, Tajiri D, Pennica D, Bennett WF, Salwa J and others. 1984. BIOLOGICAL PROPERTIES OF HUMAN TISSUE-TYPE PLASMINOGEN-ACTIVATOR OBTAINED BY EXPRESSION OF RECOMBINANT DNA IN MAMMALIAN-CELLS. Journal of Pharmacology and Experimental Therapeutics 231(1):146-152.
Cong L, Ran FA, Cox D, Lin SL, Barretto R, Habib N, Hsu PD, Wu XB, Jiang WY, Marraffini LA and others. 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 339(6121):819-823.
D'Agostino M, Risselada HJ, Lurick A, Ungermann C, Mayer A. 2017. A tethering complex drives the terminal stage of SNARE-dependent membrane fusion. Nature.
Dalton AC, Barton WA. 2014. Over-expression of secreted proteins from mammalian cell lines. Protein Science 23(5):517-525.
Daramola O, Stevenson J, Dean G, Hatton D, Pettman G, Holmes W, Field R. 2014. A High-Yielding CHO Transient System: Coexpression of Genes Encoding EBNA-1 and GS Enhances Transient Protein Expression. Biotechnology Progress 30(1):132-141.
Davies SL, Lovelady CS, Grainger RK, Racher AJ, Young RJ, James DC. 2013. Functional heterogeneity and heritability in CHO cell populations. Biotechnology and Bioengineering 110(1):260-274.
Davis R, Schooley K, Rasmussen B, Thomas J, Reddy P. 2000. Effect of PDI overexpression on recombinant protein secretion in CHO cells. Biotechnology Progress 16(5):736-743.
De Jesus M, Wurm FM. 2011. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. European Journal of Pharmaceutics and Biopharmaceutics 78(2):184-188.
de la Torre BG, Albericio F. 2018. The Pharmaceutical Industry in 2017. An Analysis of FDA Drug Approvals from the Perspective of Molecules. Molecules 23(3).
De Matteis MA, Luini A. 2008. Exiting the Golgi complex. Nature Reviews Molecular Cell Biology 9(4):273-284.
Deer JR, Allison DS. 2004. High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1 alpha gene. Biotechnology Progress 20(3):880-889.
Derouazi M, Girard P, Van Tilborgh F, Iglesias K, Muller N, Bertschinger M, Wurm FM. 2004. Serum-free large-scale transient transfection of CHO cells. Biotechnology and Bioengineering 87(4):537-545.
Dinnis DM, James DC. 2005. Engineering mammalian cell factories for improved recombinant monoclonal antibody production: Lessons from nature? Biotechnology and Bioengineering 91(2):180-189.
Dinter A, Berger EG. 1998. Golgi-disturbing agents. Histochemistry and Cell Biology 109(5-6):571-590.
Doolan P, Melville M, Gammell P, Sinacore M, Meleady P, McCarthy K, Francullo L, Leonard M, Charlebois T, Clynes M. 2008. Transcriptional profiling of gene expression changes in a PACE-transfected CHO DUKX cell line secreting high levels of rhBMP-2. Molecular Biotechnology 39(3):187-199.
Dorner AJ, Kaufman RJ. 1994. THE LEVELS OF ENDOPLASMIC-RETICULUM PROTEINS AND ATP AFFECT FOLDING AND SECRETION OF SELECTIVE PROTEINS. Biologicals 22(2):103-112.
Dorner AJ, Krane MG, Kaufman RJ. 1988. REDUCTION OF ENDOGENOUS GRP78 LEVELS IMPROVES SECRETION OF A HETEROLOGOUS PROTEIN IN CHO CELLS. Molecular and Cellular Biology 8(10):4063-4070.
Du Z, Treiber D, McCarter JD, Fomina-Yadlin D, Saleem RA, McCoy RE, Zhang Y, Tharmalingam T, Leith M, Follstad BD and others. 2015. Use of a Small Molecule Cell Cycle Inhibitor to Control Cell Growth and Improve Specific Productivity and Product Quality of Recombinant Proteins in CHO Cell Cultures. Biotechnology and Bioengineering 112(1):141-155.

Du Z, Treiber D, McCoy RE, Miller AK, Han M, He F, Domnitz S, Heath C, Reddy P. 2013. Non-invasive UPR monitoring system and its applications in CHO production cultures. Biotechnology and Bioengineering 110(8):2184-2194.
Dubuke ML, Munson M. 2016. The Secret Life of Tethers: The Role of Tethering Factors in SNARE Complex Regulation. Frontiers in cell and developmental biology 4:42-42.
Durocher Y, Butler M. 2009. Expression systems for therapeutic glycoprotein production. Current Opinion in Biotechnology 20(6):700-707.
Ecker DM, Jones SD, Levine HL. 2015. The therapeutic monoclonal antibody market. Mabs 7(1):9-14.
Faini M, Beck R, Wieland FT, Briggs JAG. 2013. Vesicle coats: structure, function, and general principles of assembly. Trends in Cell Biology 23(6):279-288.
Fan L, Kadura I, Krebs LE, Hatfield CC, Shaw MM, Frye CC. 2012. Improving the efficiency of CHO cell line generation using glutamine synthetase gene knockout cells. Biotechnology and Bioengineering 109(4):1007-1015.
Fan L, Kadura I, Krebs LE, Larson JL, Bowden DM, Frye CC. 2013. Development of a highly-efficient CHO cell line generation system with engineered SV40E promoter. Journal of Biotechnology 168(4):652-658.
Farhan H, Weiss M, Tani K, Kaufman RJ, Hauri H-P. 2008. Adaptation of endoplasmic reticulum exit sites to acute and chronic increases in cargo load. Embo Journal 27(15):2043-2054.
Fasshauer D. 2003. Structural insights into the SNARE mechanism. Biochimica Et Biophysica Acta-Molecular Cell Research 1641(2-3):87-97.
Feng Y, Jadhav AP, Rodighiero C, Fujinaga Y, Kirchhausen T, Lencer WI. 2004. Retrograde transport of cholera toxin from the plasma membrane to the endoplasmic reticulum requires the trans-Golgi network but not the Golgi apparatus in Exo2-treated cells. Embo Reports 5(6):596-601.
Feng Y, Yu S, Lasell TKR, Jadhav AP, Macia E, Chardin P, Melancon P, Roth M, Mitchison T, Kirchhausen T. 2003. Exo1: A new chemical inhibitor of the exocytic pathway. Proceedings of the National Academy of Sciences of the United States of America 100(11):6469-6474.
Fernandez-Martell A, Johari YB, James DC. 2017. Metabolic phenotyping of CHO cells varying in cellular biomass accumulation and maintenance during fed-batch culture. Biotechnology and bioengineering.
Finger FP, Hughes TE, Novick P. 1998. Sec3p is a spatial landmark for polarized secretion in budding yeast. Cell 92(4):559-571.
Fischer S, Handrick R, Otte K. 2015. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. Biotechnology Advances 33(8):1878-1896.
Fischer S, Marquart KF, Pieper LA, Fieder J, Gamer M, Gorr I, Schulz P, Bradl H. 2017. miRNA engineering of CHO cells facilitates production of difficult-to-express proteins and increases success in cell line development. Biotechnology and Bioengineering 114(7):1495-1510.
Flach H, Rosenbaum M, Duchniewicz M, Kim S, Zhang SL, Cahalan MD, Mittler G, Grosschel R. 2010. Mzb1 Protein Regulates Calcium Homeostasis, Antibody Secretion, and Integrin Activation in Innate-like B Cells. Immunity 33(5):723-735.
Florin L, Pegel A, Becker E, Hausser A, Olayioye MA, Kaufmann H. 2009. Heterologous expression of the lipid transfer protein CERT increases therapeutic protein productivity of mammalian cells. Journal of Biotechnology 141(1-2):84-90.
Fokin AI, Brodsky IB, Burakov AV, Nadezhdina ES. 2014. Interaction of early secretory pathway and Golgi membranes with microtubules and microtubule motors. Biochemistry-Moscow 79(9):879-893.
Fox RM, Hanlon CD, Andrew DJ. 2010. The CrebA/Creb3-like transcription factors are major and direct regulators of secretory capacity. Journal of Cell Biology 191(3):479-492.
Fucini RV, Chen JL, Sharma C, Kessels MM, Stamnes M. 2002. Golgi vesicle proteins are linked to the assembly of an actin complex defined by mAbp1. Molecular Biology of the Cell 13(2):621-631.
Fucini RV, Navarrete A, Vadakkan C, Lacomis L, Erdjument-Bromage H, Tempst P, Stamnes M. 2000. Activated ADP-ribosylation factor assembles distinct pools of actin on Gel membranes. Journal of Biological Chemistry 275(25):18824-18829.
Fugmann T, Hausser A, Schoeffler P, Schmid S, Pfizenmaier K, Olayioye MA. 2007. Regulation of secretory transport by protein kinase D-mediated phosphorylation of the ceramide transfer protein. Journal of Cell Biology 178(1):15-22.

Fukuda M. 2008. Regulation of secretory vesicle traffic by Rab small GTPases. Cellular and Molecular Life Sciences 65(18):2801-2813.
Fussenegger M, Schlatter S, Datwyler D, Mazur X, Bailey JE. 1998. Controlled proliferation by multigene metabolic engineering enhances the productivity of Chinese hamster ovary cells. Nature Biotechnology 16(5):468-472.
Gahn TA, Schildkraut CL. 1989. THE EPSTEIN-BARR VIRUS ORIGIN OF PLASMID REPLICATION, ORIP, CONTAINS BOTH THE INITIATION AND TERMINATION SITES OF DNA-REPLICATION. Cell 58(3):527-535.
Garcia-Mata R, Szul T, Alvarez C, Elizabeth S. 2003. ADP-ribosylation factor/COPI-dependent events at the endoplasmic reticulum-Golgi interface are regulated by the guanine nucleotide exchange factor GBF1. Molecular Biology of the Cell 14(6):2250-2261.
Gardner BM, Pincus D, Gotthardt K, Gallagher CM, Walter P. 2013. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. Cold Spring Harbor Perspectives in Biology 5(3).
Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. Walker J.M. (eds) Humana Press.
Gatti MDL, Wlaschin KF, Nissom PM, Yap M, Hu W-S. 2007. Comparative transcriptional analysis of mouse hybridoma and recombinant Chinese hamster ovary cells undergoing butyrate treatment. Journal of Bioscience and Bioengineering 103(1):82-91.
Geoghegan D, Arnall C, Hatton D, Noble-Longster J, Sellick C, Senussi T, James DC. 2018. Control of Amino Acid Transport into CHO Cells. Biotechnology and bioengineering.
Gerst JE. 2003. SNARE regulators: matchmakers and matchbreakers. Biochimica Et Biophysica Acta-Molecular Cell Research 1641(2-3):99-110.
Gilchrist A, Au CE, Hiding J, Bell AW, Fernandez-Rodriguez J, Lesimple S, Nagaya H, Roy L, Gosline SJC, Hallett M and others. 2006. Quantitative proteomics analysis of the secretory pathway. Cell 127(6):1265-1281.
Gillingham AK, Munro S. 2016. Finding the Golgi: Golgin Coiled-Coil Proteins Show the Way. Trends in Cell Biology 26(6):399-408.
Gillingham AK, Pfeifer AC, Munro S. 2002. CASP, the alternatively spliced product of the gene encoding the CCAAT-displacement protein transcription factor, is a Golgi membrane protein related to giantin. Molecular Biology of the Cell 13(11):3761-3774.
Ginn SL, Amaya AK, Alexander IE, Edelstein M, Abedi MR. 2018. Gene therapy clinical trials worldwide to 2017: An update. Journal of Gene Medicine 20(5).
Glick BS, Luini A. 2011. Models for Golgi Traffic: A Critical Assessment. Cold Spring Harbor Perspectives in Biology 3(11).
Glick BS, Nakano A. 2009. Membrane Traffic Within the Golgi Apparatus. Annual Review of Cell and Developmental Biology. p 113-132.
Gomez JA, Rutkowski DT. 2016. Experimental reconstitution of chronic ER stress in the liver reveals feedback suppression of BiP mRNA expression. Elife 5.
Gordon DE, Bond LM, Sahlender DA, Peden AA. 2010. A Targeted siRNA Screen to Identify SNAREs Required for Constitutive Secretion in Mammalian Cells. Traffic 11(9):1191-1204.
Gorur A, Yuan L, Kenny SJ, Baba S, Xu K, Schekman R. 2017. COPII-coated membranes function as transport carriers of intracellular procollagen I. Journal of Cell Biology 216(6):1745-1759.
Grabski R, Balklava Z, Wyrozumska P, Szul T, Brandon E, Alvarez C, Holloway ZG, Sztul E. 2012. Identification of a functional domain within the p115 tethering factor that is required for Golgi ribbon assembly and membrane trafficking. Journal of Cell Science 125(8):1896-1909.
Graham TR, Burd CG. 2011. Coordination of Golgi functions by phosphatidylinositol 4-kinases. Trends in Cell Biology 21(2):113-121.
Guetzoyan LJ, Spooner RA, Boal F, Stephens DJ, Lord JM, Roberts LM, Clarkson GJ. 2010. Fine tuning Exo2, a small molecule inhibitor of secretion and retrograde trafficking pathways in mammalian cells. Molecular Biosystems 6(10):2030-2038.
Guillemain I, Exton JH. 1997. Effects of brefeldin A on phosphatidylcholine phospholipase D and inositolphospholipid metabolism in HL-60 cells. European Journal of Biochemistry 249(3):812-819.
Guo Y, Punj V, Sengupta D, Linstedt AD. 2008. Coat-tether interaction in Golgi organization. Molecular Biology of the Cell 19(7):2830-2843.

Guo ZH, Liu LX, Cafiso D, Castle D. 2002. Perturbation of a very late step of regulated exocytosis by a secretory carrier membrane protein (SCAMP2)-derived peptide. Journal of Biological Chemistry 277(38):35357-35363.
Ha TK, Hansen AH, Kol S, Kildegaard HF, Lee GM. 2018. Baicalein Reduces Oxidative Stress in CHO Cell Cultures and Improves Recombinant Antibody Productivity. Biotechnology Journal 13(3).
Hacker DL, De Jesus M, Wurm FM. 2009. 25 years of recombinant proteins from reactor-grown cells - Where do we go from here? Biotechnology Advances 27(6):1023-1027.
Halic M, Gartmann M, Schlenker O, Mielke T, Pool MR, Sinning I, Beckmann R. 2006. Signal recognition particle receptor exposes the ribosomal translocon binding site. Science 312(5774):745-747.
Hammond S, Kaplarevic M, Borth N, Betenbaugh MJ, Lee KH. 2012. Chinese hamster genome database: An online resource for the CHO community at http://www.CHOgenome.org. Biotechnology and Bioengineering 109(6):1353-1356.
Hanna MG, Mela I, Wang L, Henderson RM, Chapman ER, Edwardson JM, Audhya A. 2016. Sar1 GTPase Activity Is Regulated by Membrane Curvature. Journal of Biological Chemistry 291(3):1014-1027.
Hansen HG, Pristovsek N, Kildegaard HF, Lee GM. 2017. Improving the secretory capacity of Chinese hamster ovary cells by ectopic expression of effector genes: Lessons learned and future directions. Biotechnology Advances 35(1):64-76.
Harakuge S, Kuge O, Orci L, Amherdt M, Ravazzola M, Wieland FT, Rothman JE. 1994. EN-BLOC INCORPORATION OF COATOMER SUBUNITS DURING THE ASSEMBLY OF COP-COATED VESICLES. Journal of Cell Biology 124(6):883-892.
Harding HP, Zhang YH, Bertolotti A, Zeng HQ, Ron D. 2000. Perk is essential for translational regulation and cell survival during the unfolded protein response. Molecular Cell 5(5):897-904.
Haredy AM, Nishizawa A, Honda K, Ohya T, Ohtake H, Omasa T. 2011. ATF4 over-expression increased IgG1 productivity in Chinese hamster ovary cells. BMC proceedings 5 Suppl 8:O11-O11.
Haredy AM, Nishizawa A, Honda K, Ohya T, Ohtake H, Omasa T. 2013. Improved antibody production in Chinese hamster ovary cells by ATF4 overexpression. Cytotechnology 65(6):993-1002.
Harreither E, Hackl M, Pichler J, Shridhar S, Auer N, Labaj PP, Scheideler M, Karbiener M, Grillari J, Kreil DP and others. 2015. Microarray profiling of preselected CHO host cell subclones identifies gene expression patterns associated with in-creased production capacity. Biotechnology Journal 10(10):1625-1638.
Hartley CL, Edwards S, Mullan L, Bell PA, Fresquet M, Boot-Handford RP, Briggs MD. 2013. Armet/Manf and Creld2 are components of a specialized ER stress response provoked by inappropriate formation of disulphide bonds: implications for genetic skeletal diseases. Human molecular genetics 22(25):5262-5275.
Hatton D, Forrest-Owen W, Dean G, Gibson S, Crook T, Lunney A, Ruddock S, Davis A, Daramola L, Field R. 2010. High-Yielding CHO Cell Pools for Rapid Production of Recombinant Antibodies. Noll T, editor. 239-244 p.
Hausser A, Storz P, Martens S, Link G, Toker A, Pfizenmaier K. 2005. Protein kinase D regulates vesicular transport by phosphorylating and activating phosphatidylinositol-4 kinase III beta at the Golgi complex. Nature Cell Biology 7(9):880-U24.
Hay JC, Klumperman J, Oorschot V, Steegmaier M, Kuo CS, Scheller RH. 1998. Localization, dynamics, and protein interactions reveal distinct roles for ER and Golgi SNAREs. Journal of Cell Biology 141(7):1489-1502.
He Y, Sun S, Sha H, Lu Z, Yang L, Xue Z, Chen H, Qi L. 2010. Emerging Roles for XBP1, a sUPeR Transcription Factor. Gene Expression 15(1):13-25.
He Z, Proudfoot C, Mileham AJ, McLaren DG, Whitelaw CBA, Lillico SG. 2015. Highly Efficient Targeted Chromosome Deletions Using CRISPR/Cas9. Biotechnology and Bioengineering 112(5):1060-1064.
Henne WM, Liou J, Emr SD. 2015. Molecular mechanisms of inter-organelle ER-PM contact sites. Current Opinion in Cell Biology 35:123-130.
Hirschberg K, Miller CM, Ellenberg J, Presley JF, Siggia ED, Phair RD, Lippincott-Schwartz J. 1998. Kinetic analysis of secretory protein traffic and characterization of Golgi to plasma

membrane transport intermediates in living cells. Journal of Cell Biology 143(6):1485-1503.
Hoang H-D, Maruyama J-i, Kitamoto K. 2015. Modulating Endoplasmic Reticulum-Golgi Cargo Receptors for Improving Secretion of Carrier-Fused Heterologous Proteins in the Filamentous Fungus Aspergillus oryzae. Applied and Environmental Microbiology 81(2):533-543.
Hogwood CEM, Bracewell DG, Smales CM. 2014. Measurement and control of host cell proteins (HCPs) in CHO cell bioprocesses. Current Opinion in Biotechnology 30:153-160.
Hong W, Lev S. 2014. Tethering the assembly of SNARE complexes. Trends in Cell Biology 24(1):35-43.
Horgan CP, McCaffrey MW. 2011. Rab GTPases and microtubule motors. Biochemical Society Transactions 39:1202-1206.
Hu M-C, Gong H-Y, Lin G-H, Hu S-Y, Chen MH-C, Huang S-J, Liao C-F, Wu J-L. 2007. XBP-1, a key regulator of unfolded protein response, activates transcription of IGF1 and Akt phosphorylation in zebrafish embryonic cell line. Biochemical and Biophysical Research Communications 359(3):778-783.
Huang Y-M, Hu W, Rustandi E, Chang K, Yusuf-Makagiansar H, Ryll T. 2010. Maximizing Productivity of CHO Cell-Based Fed-Batch Culture Using Chemically Defined Media Conditions and Typical Manufacturing Equipment. Biotechnology Progress 26(5):1400-1410.
Hui N, Nakamura N, Sonnichsen B, Shima DT, Nilsson T, Warren G. 1997. An isoform of the Golgi t-SNARE, syntaxin 5, with an endoplasmic reticulum retrieval signal. Molecular Biology of the Cell 8(9):1777-1787.
Huitema K, van den Dikkenberg J, Brouwers J, Holthuis JCM. 2004. Identification of a family of animal sphingomyelin synthases. Embo Journal 23(1):33-44.
Hunziker W, Whitney JA, Mellman I. 1991. SELECTIVE-INHIBITION OF TRANSCYTOSIS BY BREFELDIN-A IN MDCK CELLS. Cell 67(3):617-627.
Hussain H, Fisher DI, Abbott WM, Roth RG, Dickson AJ. 2017. Use of a Protein Engineering Strategy to Overcome Limitations in the Production of 'Difficult to Express' Recombinant Proteins. Biotechnology and bioengineering.
Hussain H, Maldonado-Agurto R, Dickson AJ. 2014. The endoplasmic reticulum and unfolded protein response in the control of mammalian recombinant protein production. Biotechnology Letters 36(8):1581-1593.
ICH-Q5D. 1998. International Conference on Harmonisation; guidance on quality of biotechnological/biological products: derivation and characterization of cell substrates used for production of biotechnological/biological products; availability. Notice. Food and Drug Administration, HHS. Federal register 63(182):50244-9.
Iordanova JV, Fasshauer D. 2016. Analysis of Scfd2-A New Member of the SM Protein Family. Biophysical Journal 110(3):597A-597A.
Ishaque A, Thrift J, Murphy JE, Konstantinov K. 2007. Over-expression of Hsp70 in BHK-21 cells engineered to produce recombinant factor VIII promotes resistance to apoptosis and enhances secretion. Biotechnology and Bioengineering 97(1):144-155.
Ivan V, de Voer G, Xanthakis D, Spoorendonk KM, Kondylis V, Rabouille C. 2008. Drosophila Sec16 Mediates the Biogenesis of tER Sites Upstream of Sar1 through an Arginine-Rich Motif. Molecular Biology of the Cell 19(10):4352-4365.
Ivanov AI. 2014. Pharmacological Inhibitors of Exocytosis and Endocytosis: Novel Bullets for Old Targets. Exocytosis and Endocytosis, 2nd Edition 1174:3-18.
Jackson CL, Casanova JE. 2000. Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. Trends in Cell Biology 10(2):60-67.
Jackson LP, Kelly BT, McCoy AJ, Gaffry T, James LC, Collins BM, Hoening S, Evans PR, Owen DJ. 2010. A Large-Scale Conformational Change Couples Membrane Recruitment to Cargo Binding in the AP2 Clathrin Adaptor Complex. Cell 141(7):1220-U213.
Jacob A, Jing J, Lee J, Schedin P, Gilbert SM, Peden AA, Junutula JR, Prekeris R. 2013. Rab40b regulates trafficking of MMP2 and MMP9 during invadopodia formation and invasion of breast cancer cells. Journal of Cell Science 126(20):4647-4658.
Jadhav V, Hackl M, Druz A, Shridhar S, Chung C-Y, Heffner KM, Kreil DP, Betenbaugh M, Shiloach J, Barron N and others. 2013. CHO microRNA engineering is growing up: Recent successes and future challenges. Biotechnology Advances 31(8):1501-1513.

Jaeger R, Bertrand MJM, Gorman AM, Vandenabeele P, Samali A. 2012. The unfolded protein response at the crossroads of cellular life and death during endoplasmic reticulum stress. Biology of the Cell 104(5):259-270.
Jensen D, Schekman R. 2011. COPII-mediated vesicle formation at a glance. Journal of Cell Science 124(1):1-4.
Jerome-Majewska LA, Achkar T, Luo L, Lupu F, Lacy E. 2010. The trafficking protein Tmed2/p24 beta(1) is required for morphogenesis of the mouse embryo and placenta. Developmental Biology 341(1):154-166.
Jin Y, Sultana A, Gandhi P, Franklin E, Hamamoto S, Khan AR, Munson M, Schekman R, Weisman LS. 2011. Myosin V Transports Secretory Vesicles via a Rab GTPase Cascade and Interaction with the Exocyst Complex. Developmental Cell 21(6):1156-1170.
Johari YB, Estes SD, Alves CS, Sinacore MS, James DC. 2015. Integrated cell and process engineering for improved transient production of a "difficult-to-express" fusion protein by CHO cells. Biotechnology and Bioengineering 112(12):2527-2542.
Josse L, Smales CM, Tuite MF. 2010. Transient Expression of Human TorsinA Enhances Secretion of Two Functionally Distinct Proteins in Cultured Chinese Hamster Ovary (CHO) Cells. Biotechnology and Bioengineering 105(3):556-566.
Jump DB. 2011. Fatty acid regulation of hepatic lipid metabolism. Current Opinion in Clinical Nutrition and Metabolic Care 14(2):115-120.
Junutula JR, Schonteich E, Wilson GM, Peden AA, Scheller RH, Prekeris R. 2004. Molecular characterization of Rab11 interactions with members of the family of Rab11-interacting proteins. Journal of Biological Chemistry 279(32):33430-33437.
Kamal A, Goldstein LSB. 2002. Principles of cargo attachment to cytoplasmic motor proteins. Current Opinion in Cell Biology 14(1):63-68.
Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research 28(1):27-30.
Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Research 42(D1):D199-D205.
Kang S, Ren D, Xiao G, Daris K, Buck L, Enyenihi AA, Zubarev R, Bondarenko PV, Deshpande R. 2014. Cell Line Profiling to Improve Monoclonal Antibody Production. Biotechnology and Bioengineering 111(4):748-760.
Kannan S. 2018. 4 Easy Steps to Analyze Your qPCR Data Using Double Delta Ct Analysis. https://bitesizebio.com/24894/4-easy-steps-to-analyze-your-qpcr-data-using-double-delta-ct-analysis/.
Kee Y, Yoo JS, Hazuka CD, Peterson KE, Hsu SC, Scheller RH. 1997. Subunit structure of the mammalian exocyst complex. Proceedings of the National Academy of Sciences of the United States of America 94(26):14438-14443.
Kelley B. 2009. Industrialization of mAb production technology The bioprocessing industry at a crossroads. Mabs 1(5):443-452.
Kesik-Brodacka M. 2018. Progress in biopharmaceutical development. Biotechnology and Applied Biochemistry 65(3):306-322.
Khaitlina SY. 2014. Intracellular transport based on actin polymerization. Biochemistry-Moscow 79(9):917-927.
Kim JJ, Lipatova Z, Segev N. 2016. TRAPP Complexes in Secretion and Autophagy. Frontiers in cell and developmental biology 4:20-20.
Kim SH, Lee GM. 2009. Development of serum-free medium supplemented with hydrolysates for the production of therapeutic antibodies in CHO cell cultures using design of experiments. Applied Microbiology and Biotechnology 83(4):639-648.
Kinseth MA, Anjard C, Fuller D, Guizzunti G, Loomis WF, Malhotra V. 2007. The golgi-associated protein GRASP is required for unconventional protein secretion during development. Cell 130(3):524-534.
Kirchhoff CF, Wang X-ZM, Conlon HD, Anderson S, Ryan AM, Bose A. 2017. Biosimilars: Key regulatory considerations and similarity assessment tools. Biotechnology and Bioengineering 114(12):2696-2705.
Klein U, Casola S, Cattoretti G, Shen Q, Lia M, Mo TW, Ludwig T, Rajewsky K, Dalla-Favera R. 2006. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. Nature Immunology 7(7):773-782.

Knorr RL, Mizushima N, Dimova R. 2017. Fusion and scission of membranes: Ubiquitous topological transformations in cells. Traffic 18(11):758-761.
Koduri RK, Miller JT, Thammana P. 2001. An efficient homologous recombination vector pTV(I) contains a hot spot for increased recombinant protein expression in Chinese hamster ovary cells. Gene 280(1-2):87-95.
Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23(10):1289-1291.
Kozak M. 1987. AT LEAST 6 NUCLEOTIDES PRECEDING THE AUG INITIATOR CODON ENHANCE TRANSLATION IN MAMMALIAN-CELLS. Journal of Molecular Biology 196(4):947-950.
Kraemer A, Mentrup T, Kleizen B, Rivera-Milla E, Reichenbach D, Enzensperger C, Nohl R, Taeuscher E, Goerls H, Ploubidou A and others. 2013. Small molecules intercept Notch signaling and the early secretory pathway. Nature Chemical Biology 9(11):731-+.
Kraemer O, Klausing S, Noll T. 2010. Methods in mammalian cell line engineering: from random mutagenesis to sequence-specific approaches. Applied Microbiology and Biotechnology 88(2):425-436.
Kreykenbohm V, Wenzel D, Antonin W, Atlachkine V, von Mollard GF. 2002. The SNAREs vti1a and vti1b have distinct localization and SNARE complex partners. European Journal of Cell Biology 81(5):273-280.
Ku SCY, Ng DTW, Yap MGS, Chao S-H. 2008. Effects of overexpression of X-box binding protein 1 on recombinant protein production in Chinese hamster ovary and NS0 myeloma cells. Biotechnology and Bioengineering 99(1):155-164.
Kumar A, Baycin-Hizal D, Wolozny D, Pedersen LE, Lewis NE, Heffner K, Chaerkady R, Cole RN, Shiloach J, Zhang H and others. 2015. Elucidation of the CHO Super-Ome (CHO-SO) by Proteoinformatics. Journal of Proteome Research 14(11):4687-4703.
Kyte J, Doolittle RF. 1982. A SIMPLE METHOD FOR DISPLAYING THE HYDROPATHIC CHARACTER OF A PROTEIN. Journal of Molecular Biology 157(1):105-132.
Larance M, Lomond AI. 2015. Multidimensional proteomics for cell biology. Nature Reviews Molecular Cell Biology 16(5):269-280.
Latham CF, Meunier FA. 2007. Munc 18a: Munc-y business in mediating exocytosis. International Journal of Biochemistry & Cell Biology 39(9):1576-1581.
Lavoie C, Paiement J, Dominguez M, Roy L, Dahan S, Gushue JN, Bergeron JJM. 1999. Roles for alpha(2)p24 and COPI in endoplasmic reticulum cargo exit site formation. Journal of Cell Biology 146(2):285-299.
Lazaro-Dieguez F, Jimenez N, Barth H, Koster AJ, Renau-Piqueras J, Llopis JL, Burger KNJ, Egea G. 2006. Actin filaments are involved in the maintenance of Golgi cisternae morphology and intra-Golgi pH. Cell Motility and the Cytoskeleton 63(12):778-791.
Le Fourn V, Girod P-A, Buceta M, Regamey A, Mermod N. 2014. CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. Metabolic Engineering 21:91-102.
Leamy AK, Egnatchik RA, Shiota M, Ivanova PT, Myers DS, Brown HA, Young JD. 2014. Enhanced synthesis of saturated phospholipids is associated with ER stress and lipotoxicity in palmitate-treated hepatic cells. Journal of Lipid Research 55(7):1478-1488.
Ledford H. 2018. Rush to protect lucrative antibody patents kicks into gear. Nature 557(7707):623-624.
Lee JS, Kallehauge TB, Pedersen LE, Kildegaard HF. 2015. Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. Scientific Reports 5.
Lee MCS, Miller EA, Goldberg J, Orci L, Schekman R. 2004. Bi-directional protein transport between the ER and Golgi. Annual Review of Cell and Developmental Biology 20:87-123.
Leung L, Kwong M, Hou S, Lee C, Chan JY. 2003. Deficiency of the Nrf1 and Nrf2 transcription factors results in early embryonic lethality and severe oxidative stress. Journal of Biological Chemistry 278(48):48021-48029.
Lewis AM, Abu-Absi NR, Borys MC, Li ZJ. 2016. The use of "Omics technology to rationally improve industrial mammalian cell line performance. Biotechnology and Bioengineering 113(1):26-38.
Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C and others. 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. Nature Biotechnology 31(8):759-+.

Li C, Fan Y, Lan T-H, Lambert NA, Wu G. 2012. Rab26 Modulates the Cell Surface Transport of alpha(2)-Adrenergic Receptors from the Golgi. Journal of Biological Chemistry 287(51):42784-42794.
Li C, Wei J, Li Y, He X, Zhou Q, Yan J, Zhang J, Liu Y, Liu Y, Shu H-B. 2013. Transmembrane Protein 214 (TMEM214) Mediates Endoplasmic Reticulum Stress-induced Caspase 4 Enzyme Activation and Apoptosis. Journal of Biological Chemistry 288(24):17908-17917.
Li MQ, Baumeister P, Roy B, Phan T, Foti D, Luo SZ, Lee AS. 2000. ATF6 as a transcription activator of the endoplasmic reticulum stress element: Thapsigargin stress-induced changes and synergistic interactions with NF-Y and YY1. Molecular and Cellular Biology 20(14):5096-5106.
Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R. 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Research 43(W1):W580-W584.
Lieu ZZ, Lock JG, Hammond LA, La Gruta NL, Stow JL, Gleeson PA. 2008. A trans-Golgi network golgin is required for the 14 regulated secretion of TNF in activated macrophages in vivo. Proceedings of the National Academy of Sciences of the United States of America 105(9):3351-3356.
Liljedahl M, Maeda Y, Colanzi A, Ayala I, Van Lint J, Malhotra V. 2001. Protein kinase D regulates the fission of cell surface destined transport carriers from the trans-Golgi network. Cell 104(3):409-420.
Lin C-H, Lin Y-W, Chen Y-C, Liao C-C, Jou Y-S, Hsu M-T, Chen C-F. 2016. FNDC3B promotes cell migration and tumor metastasis in hepatocellular carcinoma. Oncotarget 7(31):49498-49508.
Lindgren K, Salmen A, Lundgren M, Bylund L, Ebler A, Faldt E, Sorvik L, Fenge C, Skoging-Nyberg U. 2009. Automation of cell line development. Cytotechnology 59(1):1-10.
Liu J, Guo W. 2012. The exocyst complex in exocytosis and cell migration. Protoplasma 249(3):587-597.
Liu LX, Guo ZH, Tieu QY, Castle A, Castle D. 2002. Role of secretory carrier membrane protein SCAMP2 in granule exocytosis. Molecular Biology of the Cell 13(12):4266-4278.
Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods 25(4):402-408.
Lomonossoff GP, D'Aoust M-A. 2016. Plant-produced biopharmaceuticals: A case of technical developments driving clinical deployment. Science 353(6305):1237-1240.
Long KR, Yamamoto Y, Baker AL, Watkins SC, Coyne CB, Conway JF, Aridor M. 2010. Sar1 assembly regulates membrane constriction and ER export. Journal of Cell Biology 190(1):115-128.
Lonza. 2008. General Protocol for nucleofectionR of suspension cell l[i]nes. bio.lonza.com/go/op/242.
Lonza. 2009. Amaxa™ 96-well Shuttle™ Optimisation Protocol for Cell Lines. bio.lonza.com/go/op/237.
Love KR, Dalvie NC, Love JC. 2017. The yeast stands alone: the future of protein biologic production. Current opinion in biotechnology 53:50-58.
Lu L, Hannoush RN, Goess BC, Varadarajan S, Shair MD, Kirchhausen T. 2013. The small molecule dispergo tubulates the endoplasmic reticulum and inhibits export. Molecular Biology of the Cell 24(7):1020-1029.
Mahameed M, Tirosh B. 2017. Engineering CHO cells with an oncogenic KIT improves cells growth, resilience to stress and productivity. Biotechnology and bioengineering.
Majors BS, Chiang GG, Betenbaugh MJ. 2009. Protein and Genome Evolution in Mammalian Cells for Biotechnology Applications. Molecular Biotechnology 42(2):216-223.
Majors BS, Chiang GG, Pederson NE, Betenbaugh MJ. 2012. Directed evolution of mammalian anti-apoptosis proteins by somatic hypermutation. Protein Engineering Design & Selection 25(1):27-38.
Malsam J, Soellner TH. 2011. Organization of SNAREs within the Golgi Stack. Cold Spring Harbor Perspectives in Biology 3(10).
Mamat U, Wilke K, Bramhill D, Schromm AB, Lindner B, Kohl TA, Luis Corchero J, Villaverde A, Schaffer L, Head SR and others. 2015. Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins. Microbial Cell Factories 14.
Manneville JB, Etienne-Manneville S, Skehel P, Carter T, Ogden D, Ferenczi M. 2003. Interaction of the actin cytoskeleton with microtubules regulates secretory organelle movement near

the plasma membrane in human endothelial cells. Journal of Cell Science 116(19):3927-3938.
Mansour SJ, Skaug J, Zhao XH, Giordano J, Scherer SW, Melancon P. 1999. p200 ARF-GEP1: A Golgi-localized guanine nucleotide exchange protein whose Sec7 domain is targeted by the drug brefeldin A. Proceedings of the National Academy of Sciences of the United States of America 96(14):7968-7973.
Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM, Stefano GB. 2014. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. Medical science monitor basic research 20:138-42.
Martin RW, Majewska NI, Chen CX, Albanetti TE, Jimenez RBC, Schmelzer AE, Jewett MC, Roy V. 2017. Development of a CHO-Based Cell-Free Platform for Synthesis of Active Monoclonal Antibodies. Acs Synthetic Biology 6(7):1370-1379.
Martincic K, Alkan SA, Cheatle A, Borghesi L, Milcarek C. 2009. Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. Nature Immunology 10(10):1102-U82.
Maslen CL, Babcock D, Redig JK, Kapeli K, Akkari YM, Olson SB. 2006. CRELD2: Gene mapping, alternate splicing, and comparative genomic identification of the promoter region. Gene 382:111-120.
Matanis T, Akhmanova A, Wulf P, Del Nery E, Weide T, Stepanova T, Galjart N, Grosveld F, Goud B, De Zeeuw CI and others. 2002. Bicaudal-D regulates COPI-independent Golgi-ER transport by recruiting the dynein-dynactin motor complex. Nature Cell Biology 4(12):986-992.
Matsumura M, Shimoda M, Arii T, Kataoka H. 1991. ADAPTATION OF HYBRIDOMA CELLS TO HIGHER AMMONIA CONCENTRATION. Cytotechnology 7(2):103-112.
Matthews CB, Wright C, Kuo A, Colant N, Westoby M, Love JC. 2017. Reexamining opportunities for therapeutic protein production in eukaryotic microorganisms. Biotechnology and Bioengineering 114(11):2432-2444.
Maurisse R, De Semir D, Emamekhoo H, Bedayat B, Abdolmohammadi A, Parsi H, Gruenert DC. 2010. Comparative transfection of DNA into primary and transformed mammalian cells from different lineages. Bmc Biotechnology 10.
McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. 2013. Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Research 41(W1):W597-W600.
Meleady P, Doolan P, Henry M, Barron N, Keenan J, O'Sullivan F, Clarke C, Gammell P, Melville MW, Leonard M and others. 2011. Sustained productivity in recombinant Chinese Hamster Ovary (CHO) cell lines: proteome analysis of the molecular basis for a process-related phenotype. Bmc Biotechnology 11.
Micaroni M, Stanley AC, Khromykh T, Venturato J, Wong CXF, Lim JP, Marsh BJ, Storrie B, Gleeson PA, Stow JL. 2013. Rab6a/a ' Are Important Golgi Regulators of Pro-Inflammatory TNF Secretion in Macrophages. Plos One 8(2).
Miller E, Antonny B, Hamamoto S, Schekman R. 2002. Cargo selection into COPII vesicles is driven by the Sec24p subunit. Embo Journal 21(22):6105-6113.
Mironov AA, Beznoussenko GV. 2012. The Kiss-and-Run Model of Intra-Golgi Transport. International Journal of Molecular Sciences 13(6):6800-6819.
Mishev K, Dejonghe W, Russinova E. 2013. Small Molecules for Dissecting Endomembrane Trafficking: A Cross-Systems View. Chemistry & Biology 20(4):475-486.
Misumi Y, Miki K, Takatsuki A, Tamura G, Ikehara Y. 1986. NOVEL BLOCKADE BY BREFELDIN-A OF INTRACELLULAR-TRANSPORT OF SECRETORY PROTEINS IN CULTURED RAT HEPATOCYTES. Journal of Biological Chemistry 261(24):1398-1403.
Miyata S, Mizuno T, Koyama Y, Katayama T, Tohyama M. 2013. The Endoplasmic Reticulum-Resident Chaperone Heat Shock Protein 47 Protects the Golgi Apparatus from the Effects of O-Glycosylation Inhibition. Plos One 8(7).
Mizuno-Yamasaki E, Rivera-Molina F, Novick P. 2012. GTPase Networks in Membrane Traffic. Annual Review of Biochemistry, Vol 81 81:637-659.
Mohrmann R, Sorensen JB. 2012. SNARE Requirements En Route to Exocytosis: from Many to Few. Journal of Molecular Neuroscience 48(2):387-394.
MolecularProbes™. 2003. NBD- and BODIPY Dye–Labeled Sphingolipids. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/mp01154.pdf.
MolecularProbes™. 2005. ER Tracker™ Dyes for Live-Cell Endoplasmic Reticulum Labeling. https://http://www.thermofisher.com/order/catalog/product/E34250.

Moore BA, Robinson HH, Xu Z. 2007. The crystal structure of mouse Exo70 reveals unique features of the mammalian exocyst. Journal of Molecular Biology 371(2):410-421.
Moore KA, Hollien J. 2012. The Unfolded Protein Response in Secretory Cell Function. Annual Review of Genetics, Vol 46 46:165-183.
Mora A, Zhang S, Carson G, Nabiswa B, Hossler P, Yoon S. 2018. Sustaining an efficient and effective CHO cell line development platform by incorporation of 24-deep well plate screening and multivariate analysis. Biotechnology Progress 34(1):175-186.
Mori K, Kuni-Karnochi R, Yamane-Ohnuki N, Wakitani M, Yamano K, Imai H, Kanda Y, Niwa R, Iida S, Uchida K and others. 2004. Engineering Chinese hamster ovary cells to maximize effector function of produced antibodies using FUT8 siRNA. Biotechnology and Bioengineering 88(7):901-908.
Morris JA, Dorner AJ, Edwards CA, Hendershot LM, Kaufman RJ. 1997. Immunoglobulin binding protein (BiP) function is required to protect cells from endoplasmic reticulum stress but is not required for the secretion of selective proteins. Journal of Biological Chemistry 272(7):4327-4334.
Morsomme P, Prescianotto-Baschong C, Riezman H. 2003. The ER v-SNAREs are required for GPI-anchored protein sorting from other secretory proteins upon exit from the ER. The Journal of cell biology 162(3):403-412.
Mougiakos I, Bosma EF, de Vos WM, van Kranenburg R, van der Oost J. 2016. Next Generation Prokaryotic Engineering: The CRISPR-Cas ToolKit. Trends in Biotechnology 34(7):575-587.
Mozley OL, Thompson BC, Fernandez-Martell A, James DC. 2014. A Mechanistic Dissection of Polyethylenimine Mediated Transfection of CHO Cells: to Enhance the Efficiency of Recombinant DNA Utilization. Biotechnology Progress 30(5):1161-1170.
Munro S. 2011. The Golgin Coiled-Coil Proteins of the Golgi Apparatus. Cold Spring Harbor Perspectives in Biology 3(6).
Muzyczka N. 1992. USE OF ADENOASSOCIATED VIRUS AS A GENERAL TRANSDUCTION VECTOR FOR MAMMALIAN-CELLS. Current Topics in Microbiology and Immunology 158:97-129.
Nakamura T, Omasa T. 2015. Optimization of cell line development in the GS-CHO expression system using a high-throughput, single cell-based clone selection system. Journal of Bioscience and Bioengineering 120(3):323-329.
Nakano A, Brada D, Schekman R. 1988. A MEMBRANE GLYCOPROTEIN, SEC12P, REQUIRED FOR PROTEIN-TRANSPORT FROM THE ENDOPLASMIC-RETICULUM TO THE GOLGI-APPARATUS IN YEAST. Journal of Cell Biology 107(3):851-863.
Nelson DS, Alvarez C, Gao YS, Garcia-Mata R, Fialkowski E, Sztul E. 1998. The membrane transport factor TAP/p115 cycles between the Golgi and earlier secretory compartments and contains distinct domains required for its localization and function. Journal of Cell Biology 143(2):319-331.
Neville JJ, Orlando J, Mann K, McCloskey B, Antoniou MN. 2017. Ubiquitous Chromatin-opening Elements (UCOEs): Applications in biomanufacturing and gene therapy. Biotechnology Advances 35(5):557-564.
Ng SK, Wang DIC, Yap MGS. 2007. Application of destabilizing sequences on selection marker for improved recombinant protein productivity in CHO-DG44. Metabolic Engineering 9(3):304-316.
Nicolaides NC, Ebel W, Kline B, Chao QM, Routhier E, Sass PM, Grasso L. 2005. Morphogenics as a tool for target discovery and drug development. Tumor Progression and Therapeutic Resistance 1059:86-96.
Nishimiya D, Mano T, Miyadai K, Yoshida H, Takahashi T. 2013. Overexpression of CHOP alone and in combination with chaperones is effective in improving antibody production in mammalian cells. Applied Microbiology and Biotechnology 97(6):2531-2539.
Nissom PM, Sanny A, Kok YJ, Hiang YT, Chuah SH, Shing TK, Lee YY, Wong KTK, Hu W-s, Sim MYG and others. 2006. Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. Molecular Biotechnology 34(2):125-140.
Niu TK, Pfeifer AC, Lippincott-Schwartz J, Jackson CL. 2005. Dynamics of GBF1, a brefeldin A-sensitive Arf1 exchange factor at the Golgi. Molecular Biology of the Cell 16(3):1213-1222.
Noh SM, Sathyamurthy M, Lee GM. 2013. Development of recombinant Chinese hamster ovary cell lines for therapeutic protein production. Current Opinion in Chemical Engineering 2(4):391-397.

Nyfeler B, Reiterer V, Wendeler MW, Stefan E, Zhang B, Michnick SW, Hauri H-P. 2008. Identification of ERGIC-53 as an intracellular transport receptor of alpha(1)-antitrypsin. Journal of Cell Biology 180(4):705-712.
Nyfeler B, Zhang B, Ginsburg D, Kaufman RJ, Hauri HP. 2006. Cargo selectivity of the ERGIC‐53/MCFD2 transport receptor complex. Traffic 7(11):1473-1481.
O'Callaghan PM, James DC. 2008. Systems biotechnology of mammalian cell factories. Briefings in Functional Genomics & Proteomics 7(2):95-110.
Oh-hashi K, Koga H, Ikeda S, Shimada K, Hirata Y, Kiuchi K. 2009. CRELD2 is a novel endoplasmic reticulum stress-inducible gene. Biochemical and Biophysical Research Communications 387(3):504-510.
Ohashi Y, Iijima H, Yamaotsu N, Yamazaki K, Sato S, Okamura M, Sugimoto K, Dan S, Hirono S, Yamori T. 2012. AMF-26, a Novel Inhibitor of the Golgi System, Targeting ADP-ribosylation Factor 1 (Arf1) with Potential for Cancer Therapy. Journal of Biological Chemistry 287(6):3885-3897.
Ohoka N, Yoshii S, Hattori T, Onozaki K, Hayashi H. 2005. TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. Embo Journal 24(6):1243-1255.
Ohya T, Hayashi T, Kiyama E, Nishii H, Miki H, Kobayashi K, Honda K, Omasa T, Ohtake H. 2008. Improved production of recombinant human antithrombin III in Chinese hamster ovary cells by ATF4 overexpression. Biotechnology and Bioengineering 100(2):317-324.
Oku M, Tanakura S, Uemura A, Sohda M, Misumi Y, Taniguchi M, Wakabayashi S, Yoshida H. 2011. Novel Cis-acting Element GASE Regulates Transcriptional Induction by the Golgi Stress Response. Cell Structure and Function 36(1):1-12.
Ono M, Mannen K, Shimada T, Kuwano M, Mifune K. 1985. EFFECT OF MONENSIN ON THE SYNTHESIS, MATURATION AND SECRETION OF VESICULAR STOMATITIS-VIRUS PROTEINS IN A MONENSIN-RESISTANT CHINESE-HAMSTER OVARY CELL-LINE. Cell Structure and Function 10(3):279-294.
Oprins A, Duden R, Kreis TE, Geuze HJ, Slot JW. 1993. BETA-COP LOCALIZES MAINLY TO THE CIS-GOLGI SIDE IN EXOCRINE PANCREAS. Journal of Cell Biology 121(1):49-59.
Orci L, Tagaya M, Amherdt M, Perrelet A, Donaldson JG, Lippincottschwartz J, Klausner RD, Rothman JE. 1991. BREFELDIN-A, A DRUG THAT BLOCKS SECRETION, PREVENTS THE ASSEMBLY OF NON-CLATHRIN-COATED BUDS ON GOLGI CISTERNAE. Cell 64(6):1183-1195.
Orellana CA, Marcellin E, Gray PP, Nielsen LK. 2017. Overexpression of the regulatory subunit of glutamate-cysteine ligase enhances monoclonal antibody production in CHO cells. Biotechnology and Bioengineering 114(8):1825-1836.
Orellana CA, Marcellin E, Schulz BL, Nouwens AS, Gray PP, Nielsen LK. 2015. High-Antibody-Producing Chinese Hamster Ovary Cells Up-Regulate Intracellular Protein Transport and Glutathione Synthesis. Journal of Proteome Research 14(2):609-618.
Ortiz JA, Castillo M, del Toro ED, Mulet J, Gerber S, Valor LM, Sala S, Sala F, Gutierrez LM, Criado M. 2005. The cysteine-rich with EGF-Like domains 2 (CRELD2) protein interacts with the large cytoplasmic domain of human neuronal nicotinic acetylcholine receptor alpha 4 and beta 2 subunits. Journal of Neurochemistry 95(6):1585-1596.
Owen DJ, Collins BM, Evans PR. 2004. Adaptors for clathrin coats: Structure and function. Annual Review of Cell and Developmental Biology 20:153-191.
Pan H, Yu J, Zhang L, Carpenter A, Zhu H, Li L, Ma D, Yuan J. 2008. A Novel Small Molecule Regulator of Guanine Nucleotide Exchange Activity of the ADP-ribosylation Factor and Golgi Membrane Trafficking. Journal of Biological Chemistry 283(45):31087-31096.
Panagopoulos I, Moller E, Dahlen A, Isaksson M, Mandahl N, Vlamis-Gardikas A, Mertens F. 2007. Characterization of the native CREB3L2 transcription factor and the FUS/CREB3L2 chimera. Genes Chromosomes & Cancer 46(2):181-191.
Pang ZP, Sun J, Rizo J, Maximov A, Sudhof TC. 2006. Genetic analysis of synaptotagmin 2 in spontaneous and Ca2+-triggered neurotransmitter release. Embo Journal 25(10):2039-2050.
Papanikou E, Glick BS. 2014. Golgi compartmentation and identity. Current Opinion in Cell Biology 29:74-81.
Patrucco L, Chiesa A, Soluri MF, Fasolo F, Takahashi H, Carninci P, Zucchelli S, Santoro C, Gustincich S, Sblattero D and others. 2015. Engineering mammalian cell factories with SINEUP noncoding RNAs to improve translation of secreted proteins. Gene 569(2):287-293.

Paul AJ, Handrick R, Ebert S, Hesse F. 2018. Identification of process conditions influencing protein aggregation in Chinese hamster ovary cell culture. Biotechnology and Bioengineering 115(5):1173-1185.
Pauloin A, Adenot P, Hue-Beauvais C, Chanat E. 2016. The perilipin-2 (adipophilin) coat of cytosolic lipid droplets is regulated by an Arf1-dependent mechanism in HC11 mouse mammary epithelial cells. Cell Biology International 40(2):143-155.
Peden AA, Park GY, Scheller RH. 2001. The di-leucine motif of vesicle-associated membrane protein 4 is required for its localization and AP-1 binding. Journal of Biological Chemistry 276(52):49183-49187.
Pelish HE, Peterson JR, Salvarezza SB, Rodriguez-Boulan E, Chen J-L, Stamnes M, Macia E, Feng Y, Shair MD, Kirchhausen T. 2006. Secramine inhibits Cdc42-dependent functions in cells and Cdc42 activation in vitro. Nature chemical biology 2(1):39-46.
Peng R-W, Abellan E, Fussenegger M. 2011. Differential Effect of Exocytic SNAREs on the Production of Recombinant Proteins in Mammalian Cells. Biotechnology and Bioengineering 108(3):611-620.
Peng R-W, Fussenegger M. 2009a. Engineering the Secretory Pathway in Mammalian Cells. AlRubeai M, editor. 233-248 p.
Peng R-W, Fussenegger M. 2009b. Molecular Engineering of Exocytic Vesicle Traffic Enhances the Productivity of Chinese Hamster Ovary Cells. Biotechnology and Bioengineering 102(4):1170-1181.
Peyroche A, Antonny B, Robineau S, Acker J, Cherfils J, Jackson CL. 1999. Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: Involvement of specific residues of the Sec7 domain. Molecular Cell 3(3):275-285.
Phillips MJ, Voeltz GK. 2016. Structure and function of ER membrane contact sites with other organelles. Nature Reviews Molecular Cell Biology 17(2):69-82.
Pieper LA, Strotbek M, Wenger T, Gamer M, Olayioye MA, Hausser A. 2017a. Secretory pathway optimization of CHO producer cells by co-engineering of the mitosRNA-1978 target genes CerS2 and Tbc1D20. Metabolic Engineering 40:69-79.
Pieper LA, Strotbek M, Wenger T, Olayioye MA, Hausser A. 2017b. ATF6-based fine-tuning of the unfolded protein response enhances therapeutic antibody productivity of Chinese hamster ovary cells. Biotechnology and Bioengineering 114(6):1310-1318.
Plate L, Cooley CB, Chen JJ, Paxman RJ, Gallagher CM, Madoux F, Genereux JC, Dobbs W, Garza D, Spicer TP and others. 2016. Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation. Elife 5.
Pobbati AV, Stein A, Fasshauer D. 2006. N- to C-terminal SNARE complex assembly promotes rapid membrane fusion. Science 313(5787):673-676.
Polishchuk EV, Di Pentima A, Luini A, Polishchuk RS. 2003. Mechanism of constitutive export from the Golgi: Bulk flow via the formation, protrusion, and en bloc cleavage of large trans-golgi network tubular domains. Molecular Biology of the Cell 14(11):4470-4485.
Porter AJ, Racher AJ, Preziosi R, Dickson AJ. 2010. Strategies for Selecting Recombinant CHO Cell Lines for cGMP Manufacturing: Improving the Efficiency of Cell Line Generation. Biotechnology Progress 26(5):1455-1464.
Prentice HL, Ehrenfels BN, Sisk WP. 2007. Improving performance of mammalian cells in fed-batch processes through "bioreactor evolution". Biotechnology Progress 23(2):458-464.
Presley JF, Cole NB, Schroer TA, Hirschberg K, Zaal KJM, LippincottSchwartz J. 1997. ER-to-Golgi transport visualized in living cells. Nature 389(6646):81-85.
Price ER, Jin MJ, Lim D, Pati S, Walsh CT, McKeon FD. 1994. CYCLOPHILIN-B TRAFFICKING THROUGH THE SECRETORY PATHWAY IS ALTERED BY BINDING OF CYCLOSPORINE-A. Proceedings of the National Academy of Sciences of the United States of America 91(9):3931-3935.
Prydz K, Hansen SH, Sandvig K, Vandeurs B. 1992. EFFECTS OF BREFELDIN-A ON ENDOCYTOSIS, TRANSCYTOSIS AND TRANSPORT TO THE GOLGI-COMPLEX IN POLARIZED MDCK CELLS. Journal of Cell Biology 119(2):259-272.
Puck TT, Cieciura SJ, Robinson A. 1958. GENETICS OF SOMATIC MAMMALIAN CELLS .3. LONG-TERM CULTIVATION OF EUPLOID CELLS FROM HUMAN AND ANIMAL SUBJECTS. Journal of Experimental Medicine 108(6):945-&.
Puissegur MP, Eichner R, Quelen C, Coyaud E, Mari B, Lebrigand K, Broccardo C, Nguyen-Khac F, Bousquet M, Brousset P. 2012. B-cell regulator of immunoglobulin heavy-chain

transcription (Bright)/ARID3a is a direct target of the oncomir microRNA-125b in progenitor B-cells. Leukemia 26(10):2224-2232.
Puthenveedu MA, Linstedt AD. 2001. Evidence that Golgi structure depends on a p115 activity that is independent of the vesicle tether components giantin and GM130. Journal of Cell Biology 155(2):227-237.
Pybus LP, Dean G, West NR, Smith A, Daramola O, Field R, Wilkinson SJ, James DC. 2014. Model-Directed Engineering of "Difficult-to-Express" Monoclonal Antibody Production by Chinese Hamster Ovary Cells. Biotechnology and Bioengineering 111(2):372-385.
Quax TEF, Claassens NJ, Soell D, van der Oost J. 2015. Codon Bias as a Means to Fine-Tune Gene Expression. Molecular Cell 59(2):149-161.
Rader RA. 2008. (Re)defining biopharmaceutical. Nature Biotechnology 26(7):743-751.
Radha S, Natarajan AT. 1998. Sodium arsenite-induced chromosomal aberrations in the Xq arm of Chinese hamster cell lines. Mutagenesis 13(3):229-234.
Rahimpour A, Vaziri B, Moazzami R, Nematollahi L, Barkhordari F, Kokabee L, Adeli A, Mahboudi F. 2013. Engineering the Cellular Protein Secretory Pathway for Enhancement of Recombinant Tissue Plasminogen Activator Expression in Chinese Hamster Ovary Cells: Effects of CERT and XBP1s Genes. Journal of Microbiology and Biotechnology 23(8):1116-1122.
Rajendra Y, Hougland MD, Schmitt MG, Barnard GC. 2015. Transcriptional and post-transcriptional targeting for enhanced transient gene expression in CHO cells. Biotechnology Letters 37(12):2379-2386.
Raote I, Ortega Bellido M, Pirozzi M, Zhang C, Melville D, Parashuraman S, Zimmermann T, Malhotra V. 2017. TANGO1 assembles into rings around COP II coats at ER exit sites. Journal of Cell Biology 216(4):901-909.
Rasala BA, Mayfield SP. 2015. Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. Photosynthesis Research 123(3):227-239.
Rawson RB. 2013. The site-2 protease. Biochimica Et Biophysica Acta-Biomembranes 1828(12):2801-2807.
Reinhart D, Sommeregger W, Debreczeny M, Gludovacz E, Kunert R. 2014. In search of expression bottlenecks in recombinant CHO cell lines-a case study. Applied Microbiology and Biotechnology 98(13):5959-5965.
Renna M, Faraonio R, Bonatti S, De Stefano D, Carnuccio R, Tajana G, Remondelli P. 2006. Nitric oxide-induced endoplasmic reticulum stress activates the expression of cargo receptor proteins and alters the glycoprotein transport to the Golgi complex. International Journal of Biochemistry & Cell Biology 38(12):2040-2048.
Ribourtout B, Zandecki M. 2015. Plasma cell morphology in multiple myeloma and related disorders. Morphologie 99(325):38-62.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: The European molecular biology open software suite. Trends in Genetics 16(6):276-277.
Rivera-Molina F, Toomre D. 2013. Live-cell imaging of exocyst links its spatiotemporal dynamics to various stages of vesicle fusion. Journal of Cell Biology 201(5):673-680.
Robinson PJ, Pringle MA, Woolhead CA, Bulleid NJ. 2017. Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation. Journal of Biological Chemistry 292(17):6978-6986.
Rosano GL, Ceccarelli EA. 2014. Recombinant protein expression in Escherichia coli: advances and challenges. Frontiers in microbiology 5:172-172.
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. Genes & Development 28(11):1165-1178.
Rouiller Y, Bielser JM, Bruhlmann D, Jordan M, Broly H, Stettler M. 2016. Screening and assessment of performance and molecule quality attributes of industrial cell lines across different fed-batch systems. Biotechnology Progress 32(1):160-170.
Rutkowski DT, Arnold SM, Miller CN, Wu J, Li J, Gunnison KM, Mori K, Akha AAS, Raden D, Kaufman RJ. 2006. Adaptation to ER stress is mediated by differential stabilities of pro-survival and pro-apoptotic mRNAs and proteins. Plos Biology 4(11):2024-2041.
Rutkowski DT, Kaufman RJ. 2007. That which does not kill me makes me stronger: adapting to chronic ER stress. Trends in Biochemical Sciences 32(10):469-476.

Ryan OW, Cate JHD. 2014. Multiplex Engineering of Industrial Yeast Genomes Using CRISPRm. Use of Crispr/Cas9, Zfns, and Talens in Generating Site-Specific Genome Alterations 546:473-489.
Saenz JB, Sun WJ, Chang JW, Li J, Bursulaya B, Gray NS, Haslam DB. 2009. Golgicide A reveals essential roles for GBF1 in Golgi assembly and function. Nature Chemical Biology 5(3):157-165.
Saito K, Maeda M, Katada T. 2017. Regulation of the Sar1 GTPase Cycle Is Necessary for Large Cargo Secretion from the Endoplasmic Reticulum. Frontiers in cell and developmental biology 5:75-75.
Saito K, Yamashiro K, Shimazu N, Tanabe T, Kontani K, Katada T. 2014. Concentration of Sec12 at ER exit sites via interaction with cTAGE5 is required for collagen export. Journal of Cell Biology 206(6):751-762.
Salaun C, James DJ, Greaves J, Chamberlain LH. 2004. Plasma membrane targeting of exocytic SNAPd, proteins. Biochimica Et Biophysica Acta-Molecular Cell Research 1693(2):81-89.
Samali A, Fitzgerald U, Deegan S, Gupta S. 2010. Methods for monitoring endoplasmic reticulum stress and the unfolded protein response. International journal of cell biology 2010:830307-830307.
Sapperstein SK, Walter DM, Grosvenor AR, Heuser JE, Waters MG. 1995. P115 IS A GENERAL VESICULAR TRANSPORT FACTOR-RELATED TO THE YEAST ENDOPLASMIC-RETICULUM TO GOLGI TRANSPORT FACTOR USO1P. Proceedings of the National Academy of Sciences of the United States of America 92(2):522-526.
Sasaki K, Yoshida H. 2015. Organelle autoregulation-stress responses in the ER, Golgi, mitochondria and lysosome. Journal of Biochemistry 157(4):185-195.
Sato K, Nakano A. 2005. Dissection of COPII subunit-cargo assembly and disassembly kinetics during Sar1p-GTP hydrolysis. Nature Structural & Molecular Biology 12(2):167-174.
Saunders F, Sweeney B, Antoniou MN, Stephens P, Cain K. 2015. Chromatin Function Modifying Elements in an Industrial Antibody Production Platform - Comparison of UCOE, MAR, STAR and cHS4 Elements. Plos One 10(4).
Schalen M, Anyaogu DC, Hoof JB, Workman M. 2016. Effect of secretory pathway gene overexpression on secretion of a fluorescent reporter protein in Aspergillus nidulans. Fungal biology and biotechnology 3:3-3.
Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C-T method. Nature Protocols 3(6):1101-1108.
Schuiki I, Volchuk A. 2012. Diverse roles for the p24 family of proteins in eukaryotic cells. Biomolecular concepts 3(6):561-70.
Schwanhaeusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. Nature 473(7347):337-342.
Sehgal PB, Lee JE. 2011. Protein trafficking dysfunctions: Role in the pathogenesis of pulmonary arterial hypertension. Pulmonary circulation 1(1):17-32.
Seth G, Charaniya S, Wiaschin KF, Hu W-S. 2007. In pursuit of a super producer - alternative paths to high producing recombinant mammalian cells. Current Opinion in Biotechnology 18(6):557-564.
Seuter S, Ryynanen J, Carlberg C. 2014. The ASAP2 gene is a primary target of 1,25-dihydroxyvitamin D-3 in human monocytes and macrophages. Journal of Steroid Biochemistry and Molecular Biology 144:12-18.
Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, Yu X, Yang LM, Tan BK, Rosenwald A and others. 2004. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. Immunity 21(1):81-93.
Shapiro-Shelef M, Calame K. 2005. Regulation of plasma-cell development. Nature Reviews Immunology 5(3):230-242.
Shemesh T, Luini A, Malhotra V, Burger KNJ, Kozlov MM. 2003. Prefission constriction of golgi tubular carriers driven by local lipid metabolism: A theoretical model. Biophysical Journal 85(6):3813-3827.
Shen JS, Chen X, Hendershot L, Prywes R. 2002. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of golgi localization signals. Developmental Cell 3(1):99-111.

Shi W, Liao Y, Willis SN, Taubenheim N, Inouye M, Tarlinton DM, Smyth GK, Hodgkin PD, Nutt SL, Corcoran LM. 2015. Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. Nature Immunology 16(6):663-+.
Short B, Preisinger C, Schaletzky J, Kopajtich R, Barr FA. 2002. The Rab6 GTPase regulates recruitment of the dynactin complex to Golgi membranes. Current Biology 12(20):1792-U5.
Sinha R, Ahmed S, Jahn R, Klingauf J. 2011. Two synaptobrevin molecules are sufficient for vesicle fusion in central nervous system synapses. Proceedings of the National Academy of Sciences of the United States of America 108(34):14318-14323.
Smales CM, Dinnis DM, Stansfield SH, Alete D, Sage EA, Birch JR, Racher AJ, Marshall CT, James DC. 2004. Comparative proteomic analysis of GS-NSO murine myeloma cell lines with varying recombinant monoclonal antibody production rate. Biotechnology and Bioengineering 88(4):474-488.
Smith EA, Weisshaar JC. 2011. Docking, Not Fusion, as the Rate-Limiting Step in a SNARE-Driven Vesicle Fusion Assay. Biophysical Journal 100(9):2141-2150.
Sollner T, Bennett MK, Whiteheart SW, Scheller RH, Rothman JE. 1993. A PROTEIN ASSEMBLY-DISASSEMBLY PATHWAY IN-VITRO THAT MAY CORRESPOND TO SEQUENTIAL STEPS OF SYNAPTIC VESICLE DOCKING, ACTIVATION, AND FUSION. Cell 75(3):409-418.
Solscheid B, Tropschug M. 2000. A novel type of FKBP in the secretory pathway of Neurospora crassa. Febs Letters 480(2-3):118-122.
Sorieul M, Langhans M, Guetzoyan L, Hillmer S, Clarkson G, Lord JM, Roberts LM, Robinson DG, Spooner RA, Frigerio L. 2011. An Exo2 Derivative Affects ER and Golgi Morphology and Vacuolar Sorting in a Tissue-Specific Manner in Arabidopsis. Traffic 12(11):1552-1562.
Sou SN, Lee K, Nayyar K, Polizzi KM, Sellick C, Kontoravdi C. 2018. Exploring cellular behavior under transient gene expression and its impact on mAb productivity and Fc-glycosylation. Biotechnology and Bioengineering 115(2):512-518.
Spang A. 2015. The Road not Taken: Less Traveled Roads from the TGN to the Plasma Membrane. Membranes 5(1):84-98.
Spang A. 2017. Cell biology: Bulky tether proteins aid membrane fusion. Nature.
Spiess C, Zhai Q, Carter PJ. 2015. Alternative molecular formats and therapeutic applications for bispecific antibodies. Molecular Immunology 67(2):95-106.
Spooner RA, Watson P, Smith DC, Boal F, Amessou M, Johannes L, Clarkson GJ, Lord JM, Stephens DJ, Roberts LM. 2008. The secretion inhibitor Exo2 perturbs trafficking of Shiga toxin between endosomes and the trans-Golgi network. Biochemical Journal 414:471-484.
Sprangers J, Rabouille C. 2015. SEC16 in COPII coat dynamics at ER exit sites. Biochemical Society Transactions 43:97-103.
Sriburi R, Bommiasamy H, Buldak GL, Robbins GR, Frank M, Jackowski S, Brewer JW. 2007. Coordinate regulation of phospholipid biosynthesis and secretory pathway gene expression in XBP-1 (S)-induced endoplasmic reticulum biogenesis. Journal of Biological Chemistry 282(10):7024-7034.
Stoyle CL, Stephens PE, Humphreys DP, Heywood S, Cain K, Bulleid NJ. 2017. IgG light chain-independent secretion of heavy chain dimers: consequence for therapeutic antibody production and design. Biochemical Journal 474:3179-3188.
Strating JRPM, Martens GJM. 2009. The p24 family and selective transport processes at the ER-Golgi interface. Biology of the Cell 101(9):495-509.
Strohl WR. 2018. Current progress in innovative engineered antibodies. Protein & Cell 9(1):86-120.
Strutzenberger K, Borth N, Kunert R, Steinfellner W, Katinger H. 1999. Changes during subclone development and ageing of human antibody-producing recombinant CHO cells. Journal of Biotechnology 69(2-3):215-226.
Suedhof TC. 2013. A molecular machine for neurotransmitter release: synaptotagmin and beyond. Nature Medicine 19(10):1227-1231.
Sztul E, Lupashin V. 2006. Role of tethering factors in secretory membrane traffic. American Journal of Physiology-Cell Physiology 290(1):C11-C26.
Sztul E, Lupashin V. 2009. Role of vesicle tethering factors in the ER-Golgi membrane traffic. Febs Letters 583(23):3770-3783.
Szul T, Sztul E. 2011. COPII and COPI Traffic at the ER-Golgi Interface. Physiology 26(5):348-364.

Tamosaitis L, Smales CM. 2018. Meta-Analysis of Publicly Available Chinese Hamster Ovary (CHO) Cell Transcriptomic Datasets for Identifying Engineering Targets to Enhance Recombinant Protein Yields. Biotechnology journal:e1800066-e1800066.
Tamura G, Ando K, Suzuki S, Takatsuki A, Arima K. 1968. ANTIVIRAL ACTIVITY OF BREFELDIN A AND VERRUCARIN A. Journal of Antibiotics 21(2):160-+.
Taniguchi M, Nadanaka S, Tanakura S, Sawaguchi S, Midori S, Kawai Y, Yamaguchi S, Shimada Y, Nakamura Y, Matsumura Y and others. 2015. TFE3 Is a bHLH-ZIP-type Transcription Factor that Regulates the Mammalian Golgi Stress Response. Cell Structure and Function 40(1):13-30.
Taniguchi M, Sasaki-Osugi K, Oku M, Sawaguchi S, Tanakura S, Kawai Y, Wakabayashi S, Yoshida H. 2016. MLX Is a Transcriptional Repressor of the Mammalian Golgi Stress Response. Cell structure and function 41(2):93-104.
Tansey EM, Catterall PP. 1994. MONOCLONAL-ANTIBODIES - A WITNESS SEMINAR IN CONTEMPORARY MEDICAL HISTORY. Medical History 38(3):322-327.
Tartakoff AM. 1983. PERTURBATION OF THE STRUCTURE AND FUNCTION OF THE GOLGI-COMPLEX BY MONO-VALENT CARBOXYLIC IONOPHORES. Methods in Enzymology 98:47-59.
Tastanova A, Schulz A, Folcher M, Tolstrup A, Puklowski A, Kaufmann H, Fussenegger M. 2016. Overexpression of YY1 increases the protein production in mammalian cells. Journal of Biotechnology 219:72-85.
Teissie J, Golzio M, Rols MP. 2005. Mechanisms of cell membrane electropermeabilization: A minireview of our present (lack of ?) knowledge. Biochimica Et Biophysica Acta-General Subjects 1724(3):270-280.
Tellam JT, McIntosh S, James DE. 1995. MOLECULAR-IDENTIFICATION OF 2 NOVEL MUNC-18 ISOFORMS EXPRESSED IN NONNEURONAL TISSUES. Journal of Biological Chemistry 270(11):5857-5863.
ThermoFisher. 2018. qPCR efficiency calculator. https://http://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/qpcr-efficiency-calculator.html.
Thompson B, Clifford J, Jenns M, Smith A, Field R, Nayyar K, James DC. 2017. High-throughput quantitation of Fc-containing recombinant proteins in cell culture supernatant by fluorescence polarization spectroscopy. Analytical Biochemistry 534:49-55.
Thompson BC, Segarra CRJ, Mozley OL, Daramola O, Field R, Levison PR, James DC. 2012. Cell line specific control of polyethylenimine-mediated transient transfection optimized with "Design of experiments" methodology. Biotechnology Progress 28(1):179-187.
Tian X, Jin RU, Bredemeyer AJ, Oates EJ, Blazewska KM, McKenna CE, Mills JC. 2010. RAB26 and RAB3D Are Direct Transcriptional Targets of MIST1 That Regulate Exocrine Granule Maturation. Molecular and Cellular Biology 30(5):1269-1284.
Tigges M, Fussenegger M. 2006. Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. Metabolic Engineering 8(3):264-272.
Torii S, Banno T, Watanabe T, Ikehara Y, Murakami K, Nakayama K. 1995. CYTOTOXICITY OF BREFELDIN-A CORRELATES WITH ITS INHIBITORY EFFECT ON MEMBRANE-BINDING OF COP COAT PROTEINS. Journal of Biological Chemistry 270(19):11574-11580.
Torres M, Zuniga R, Gutierrez M, Vergara M, Collazo N, Reyes J, Berrios J, Carlos Aguillon J, Carmen Molina M, Altamirano C. 2018. Mild hypothermia upregulates myc and xbp1s expression and improves anti-TNF alpha production in CHO cells. Plos One 13(3).
Tran K, Gurramkonda C, Cooper MA, Pilli M, Taris JE, Selock N, Han T-C, Tolosa M, Zuber A, Penalber-Johnstone C and others. 2017. Cell-free production of a therapeutic protein: Expression, purification, and characterization of recombinant streptokinase using a CHO lysate. Biotechnology and bioengineering.
Traub LM. 2005. Common principles in clathrin-mediated sorting at the Golgi and the plasma membrane. Biochimica Et Biophysica Acta-Molecular Cell Research 1744(3):415-437.
Trouillon R, Ewing AG. 2014. Actin Controls the Vesicular Fraction of Dopamine Released During Extended Kiss and Run Exocytosis. Acs Chemical Biology 9(3):812-820.

Tsukumo Y, Tsukahara S, Furuno A, Iemura S-i, Natsume T, Tomida A. 2014. TBL2 Is a Novel PERK-Binding Protein that Modulates Stress-Signaling and Cell Survival during Endoplasmic Reticulum Stress. Plos One 9(11).
Ullrich S, Muench A, Neumann S, Kremmer E, Tatzelt J, Lichtenthaler SF. 2010. The Novel Membrane Protein TMEM59 Modulates Complex Glycosylation, Cell Surface Expression, and Secretion of the Amyloid Precursor Protein. Journal of Biological Chemistry 285(27):20664-20674.
Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Research 40(15).
Urbe S, Huber LA, Zerial M, Tooze SA, Parton RG. 1993. RAB11, A SMALL GTPASE ASSOCIATED WITH BOTH CONSTITUTIVE AND REGULATED SECRETORY PATHWAYS IN PC12-CELLS. Febs Letters 334(2):175-182.
Urlaub G, Chasin LA. 1980. ISOLATION OF CHINESE-HAMSTER CELL MUTANTS DEFICIENT IN DIHYDROFOLATE-REDUCTASE ACTIVITY. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 77(7):4216-4220.
Urlaub G, Kas E, Carothers AM, Chasin LA. 1983. DELETION OF THE DIPLOID DIHYDROFOLATE-REDUCTASE LOCUS FROM CULTURED MAMMALIAN-CELLS. Cell 33(2):405-412.
Valente KN, Lenhoff AM, Lee KH. 2015. Expression of Difficult-to-Remove Host Cell Protein Impurities During Extended Chinese Hamster Ovary Cell Culture and Their Impact on Continuous Bioprocessing. Biotechnology and Bioengineering 112(6):1232-1242.
ValitaTITER. 2018a. ValitaTITER Assay Instructions for Use. https://static1.squarespace.com/static/57daa99bbe65946d7045e871/t/5a82e33ac830259134deae30/1518527312168/ValitaTiter+IFU+BMG.pdf.
ValitaTITER. 2018b. ValitaTITER HS Assay Instructions for Use. https://static1.squarespace.com/static/57daa99bbe65946d7045e871/t/5873bfc59de4bb04613a7407/1483980758829/ValitaTitre+HS+IFU+V01.pdf.
van der Linden L, van der Schaar HM, Lanke KHW, Neyts J, van Kuppeveld FJM. 2010. Differential Effects of the Putative GBF1 Inhibitors Golgicide A and AG1478 on Enterovirus Replication. Journal of Virology 84(15):7535-7542.
Vardjan N, Stenovec M, Jorgacevski J, Kreft M, Grilc S, Zorec R. 2009. The Fusion Pore and Vesicle Cargo Discharge Modulation. In: Zorec R, Vardjan N, Chowdhury HH, Kreft M, Rupnik M, editors. Mechanisms of Exocytosis. p 135-144.
Viaud J, Zeghouf M, Barelli H, Zeeh J-C, Padilla A, Guibert B, Chardin P, Royer CA, Cherfils J, Chavanieu A. 2007. Structure-based discovery of an inhibitor of Arf activation by Sec7 domains through targeting of protein-protein complexes. Proceedings of the National Academy of Sciences of the United States of America 104(25):10370-10375.
Vishwanathan N, Yongky A, Johnson KC, Fu H-Y, Jacob NM, Le H, Yusufi FNK, Lee DY, Hu W-S. 2015. Global Insights Into the Chinese Hamster and CHO Cell Transcriptomes. Biotechnology and Bioengineering 112(5):965-976.
Vollenweider F, Kappeler F, Itin C, Hauri HP. 1998. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. Journal of Cell Biology 142(2):377-389.
von Blume J, Alleaume A-M, Kienzle C, Carreras-Sureda A, Valverde M, Malhotra V. 2012. Cab45 is required for Ca2+-dependent secretory cargo sorting at the trans-Golgi network. Journal of Cell Biology 199(7):1057-1066.
Walsh G. 2002. Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. European Journal of Pharmaceutical Sciences 15(2):135-138.
Walsh G. 2014. Biopharmaceutical benchmarks 2014. Nature Biotechnology 32(10):992-1000.
Walter P, Ron D. 2011. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334(6059):1081-1086.
Wang G, Tang W, Xia J, Chu J, Noorman H, van Gulik WM. 2015a. Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. Engineering in Life Sciences 15(1):20-29.
Wang QMJ. 2006. PKD at the crossroads of DAG and PKC signaling. Trends in Pharmacological Sciences 27(6):317-323.
Wang S, Liu Y, Adamson CL, Valdez G, Guo W, Hsu SC. 2004. The mammalian exocyst, a complex required for exocytosis, inhibits tubulin polymerization. Journal of Biological Chemistry 279(34):35958-35966.

Wang T, Grabski R, Sztul E, Hay JC. 2015b. p115-SNARE Interactions: A Dynamic Cycle of p115 Binding Monomeric SNARE Motifs and Releasing Assembled Bundles. Traffic 16(2):148-171.
Wang T, Li L, Hong W. 2017. SNARE proteins in membrane trafficking. Traffic 18(12):767-775.
Wang W, Singh S, Zeng DL, King K, Nema S. 2007. Antibody structure, instability, and formulation. Journal of Pharmaceutical Sciences 96(1):1-26.
Wang X, Hunter AK, Mozier NM. 2009a. Host Cell Proteins in Biologics Development: Identification, Quantitation and Risk Assessment. Biotechnology and Bioengineering 103(3):446-458.
Wang XZ, Harding HP, Zhang YH, Jolicoeur EM, Kuroda M, Ron D. 1998. Cloning of mammalian Ire1 reveals diversity in the ER stress responses. Embo Journal 17(19):5708-5717.
Wang Z, Gerstein M, Snyder M. 2009b. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10(1):57-63.
Watson P, Forster R, Palmer KJ, Pepperkok R, Stephens DJ. 2005. Coupling of ER exit to microtubules through direct interaction of COPII with dynactin. Nature Cell Biology 7(1):48-+.
Watson P, Townley AK, Koka P, Palmer KJ, Stephens DJ. 2006. Sec16 defines endoplasmic reticulum exit sites and is required for secretory cargo export in mammalian cells. Traffic 7(12):1678-1687.
Welz T, Wellbourne-Wood J, Kerkhoff E. 2014. Orchestration of cell surface proteins by Rab11. Trends in Cell Biology 24(7):407-415.
Whaley KJ, Hiatt A, Zeitlin L. 2011. Emerging antibody products and Nicotiana manufacturing. Human Vaccines 7(3):349-356.
Witkos TM, Lowe M. 2015. The Golgin Family of Coiled-Coil Tethering Proteins. Frontiers in cell and developmental biology 3:86-86.
Witkos TM, Lowe M. 2017. Recognition and tethering of transport vesicles at the Golgi apparatus. Current Opinion in Cell Biology 47:16-23.
Wong M, Munro S. 2014. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. Science 346(6209):601-+.
Wu B, Guo W. 2015. The Exocyst at a Glance. Journal of Cell Science 128(16):2957-2964.
Wu WJ, Erickson JW, Lin R, Cerione RA. 2000. The gamma-subunit of the coatomer complex binds Cdc42 to mediate transformation. Nature 405(6788):800-804.
Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. Nature Biotechnology 22(11):1393-1398.
Wurm FM. 2013. CHO Quasispecies—Implications for Manufacturing Processes. *Processes*. p 296-311.
Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S and others. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nature Biotechnology 29(8):735-U131.
Yamamoto K, Gandin V, Sasaki M, McCracken S, Li W, Silvester JL, Elia AJ, Wang F, Wakutani Y, Alexandrova R and others. 2014. Largen: A Molecular Regulator of Mammalian Cell Size Control. Molecular Cell 53(6):904-915.
Yamasaki A, Menon S, Yu S, Barrowman J, Meerloo T, Oorschot V, Klumperman J, Satoh A, Ferro-Novick S. 2009. mTrs130 Is a Component of a Mammalian TRAPPII Complex, a Rab1 GEF That Binds to COPI-coated Vesicles. Molecular Biology of the Cell 20(19):4205-4215.
Yan JP, Colon ME, Beebe LA, Melancon P. 1994. ISOLATION AND CHARACTERIZATION OF MUTANT CHO CELL-LINES WITH COMPARTMENT-SPECIFIC RESISTANCE TO BREFELDIN-A. Journal of Cell Biology 126(1):65-75.
Yang J, Ruff AJ, Arlt M, Schwaneberg U. 2017. Casting epPCR (cepPCR): A Simple Random Mutagenesis Method to Generate High Quality Mutant Libraries. Biotechnology and Bioengineering 114(9):1921-1927.
Yang Y, Mariati, Chusainow J, Yap MGS. 2010. DNA methylation contributes to loss in productivity of monoclonal antibody-producing CHO cell lines. Journal of Biotechnology 147(3-4):180-185.
Yang Z, Wang S, Halim A, Schulz MA, Frodin M, Rahman SH, Vester-Christensen MB, Behrens C, Kristensen C, Vakhrushev SY and others. 2015. Engineered CHO cells for production of diverse, homogeneous glycoproteins. Nature Biotechnology 33(8):842-+.
Yao J, Weng Y, Dickey A, Wang KY. 2015. Plants as Factories for Human Pharmaceuticals: Applications and Challenges. International Journal of Molecular Sciences 16(12):28549-28565.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. Bmc Bioinformatics 13.
Yeaman C, Ayala MI, Wright JR, Bard F, Bossard C, Ang A, Maeda Y, Seufferlein T, Mellman I, Nelson WJ and others. 2004. Protein kinase D regulates basolateral membrane protein exit from trans-Golgi network. Nature Cell Biology 6(2):106-112.
Yee JC, Gerdtzen ZP, Hu W-S. 2009. Comparative Transcriptome Analysis to Unveil Genes Affecting Recombinant Protein Productivity in Mammalian Cells. Biotechnology and Bioengineering 102(1):246-263.
Yonemura Y, Li X, Mueller K, Kraemer A, Atigbire P, Mentrup T, Feuerhake T, Kroll T, Shomron O, Nohl R and others. 2016. Inhibition of cargo export at ER exit sites and the trans-Golgi network by the secretion inhibitor FLI-06. Journal of Cell Science 129(20):3868-3877.
Yoshida H. 2009. ER Stress Response, Peroxisome Proliferation, Mitochondrial Unfolded Protein Response and Golgi Stress Response. Iubmb Life 61(9):871-879.
Yu S, Satoh A, Pypaert M, Mullen K, Hay JC, Ferro-Novick S. 2006. MBet3p is required for homotypic COPII vesicle tethering in mammalian cells. Journal of Cell Biology 174(3):359-368.
Yu X, Breitman M, Goldberg J. 2012. A Structure-Based Mechanism for Arf1-Dependent Recruitment of Coatomer to Membranes. Cell 148(3):530-542.
Yu Y, Wang L, Jiu Y, Zhan Y, Liu L, Xia Z, Song E, Xu P, Xu T. 2011. HID-1 is a novel player in the regulation of neuropeptide sorting. Biochemical Journal 434:383-390.
Yuk IH, Nishihara J, Walker D, Jr., Huang E, Gunawan F, Subramanian J, Pynn AFJ, Yu XC, Zhu-Shimoni J, Vanderlaan M and others. 2015. More similar than different: Host cell protein production using three null CHO cell lines. Biotechnology and Bioengineering 112(10):2068-2083.
Zhang B, Kaufman RJ, Ginsburg D. 2005. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. Journal of Biological Chemistry 280(27):25881-25886.
Zhang GF, Driouich A, Staehelin LA. 1993. EFFECT OF MONENSIN ON PLANT GOLGI - REEXAMINATION OF THE MONENSIN-INDUCED CHANGES IN CISTERNAL ARCHITECTURE AND FUNCTIONAL ACTIVITIES OF THE GOLGI-APPARATUS OF SYCAMORE SUSPENSION-CULTURED CELLS. Journal of Cell Science 104:819-831.
Zhang H, Wang H, Liu M, Zhang T, Zhang J, Wang X, Xiang W. 2013. Rational development of a serum-free medium and fed-batch process for a GS-CHO cell line expressing recombinant antibody. Cytotechnology 65(3):363-378.
Zhang XM, Ellis S, Sriratana A, Mitchell CA, Rowe T. 2004. Sec15 is an effector for the Rab11 GTPase in mammalian cells. Journal of Biological Chemistry 279(41):43027-43034.
Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X. 2014. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. Plos One 9(1).
Zhou MX, Crawford Y, Ng D, Tung J, Pynn AFJ, Meier A, Yuk IH, Vijayasankaran N, Leach K, Joly J and others. 2011. Decreasing lactate level and increasing antibody production in Chinese Hamster Ovary cells (CHO) by reducing the expression of lactate dehydrogenase and pyruvate dehydrogenase kinases. Journal of Biotechnology 153(1-2):27-34.
Zhou Y, Raju R, Alves C, Gilbert A. 2018. Debottlenecking protein secretion and reducing protein aggregation in the cellular host. Current opinion in biotechnology 53:151-157.

"A bottle of wine contains more philosophy that all the books in the world"

Louis Pasteur

"Because it's there"

George Mallory

# 9) Appendices

## Appendix 1: Gene Target Database.

**Key:**

A = Strong, definite evidence of gene having positive impact upon productivity in CHO cell line/related cell line (e.g. other rodent cell line) expressing recombinant product.

B = No definitive expression evidence available, but evidence from transcriptomic/proteomic data sets and biosynthetic/secretory pathway mapping suggests an increase in gene's expression/protein's abundance is linked to increased productivity.

C = Knockdown may result in positive impact upon productivity.

Lab plasma-cell transcript data is from RNAseq data sets. Lab CHO cell transcriptomic data is from RNAseq data.

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **Chaperones; Protein processing** | | | | | | | |
| [A]**HSP70/ HSPA1a** | P0DMV8 (human) | BHK-21; insect | Factor VIII; stable | Stable over-expression | - | 50% increase in specific productivity in single clone. Due to apoptosis resistance and secretion enhancement. | (Ishaque et al. 2007) |
| [A]**Tor1a** | Q9ER39 (mouse) | CHOK1 | Secreted luciferase | Transient over-expression | - | 2.5-fold increase in luciferase secretion.<br>1.3-fold increase in IgG4 secretion. | (Josse et al. 2010) |
| | | CHOK1SV LB01 | IgG4 | | - | | |
| **FKBP11** | P45878 (mouse) | Plasma cell | - | - | Plasma cell transcript-omic | 4-fold increase in RNA transcript abundance in plasma cells. Decrease in abundance in CHO producing lines. Plays role in ER protein folding in yeast. | (Shaffer et al. 2004; Solscheid and Tropschug 2000) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| [B]**HSP90B1 (Endoplasmin)** | P08113 (mouse) | CHO | IgG | - | CHO transcrip-tomic and Proteomic; Plasma cell transcript-omic | Gene sees a 1.3-fold increase in higher-producing CHO cells grown at 33 °C when compared to lower-producers at 37 °C. Increased proteomic abundance in CHO cells with sustained productivity compared to those with non-sustained productivity. Sees a 4-fold increase in plasma cell transcriptomic abundance. | (Carlage et al. 2012; Harreither et al. 2015; Meleady et al. 2011; Yee et al. 2009) |
| [A]**PDI (A3, A4, A5, A6)** | P07237 (human) | CHO 2F5 | Mab | Stable over-expression | CHO and Plasma transcript-omic | 37% increase in Mab production rate due to 58% increase in PDI levels. All PDI transcripts upregulated in Plasma cells. | (Borth et al. 2005; Harreither et al. 2015) |
| | | CHO-DUKX; bacteria, insect, yeast | IL-15 + TNFR:Fc | Transient over-expression | | Decrease in TNFR:Fc secretion. No effect on IL-15 secretion levels. Overexpression in bacteria, yeast and insect cells resulted in increase in recombinant protein secretion. | (Davis et al. 2000) |
| [A]**CANX (Calnexin)** | P35564 (mouse) | rCHO | Thrombo poietin | Stable tet-induced over-expression (co-expressed) | Plasma and CHO transcript-omic. | 2.8-fold increase in CANX and Calr resulted in 1.9-fold increase in thrombopoietin productivity. Both are associated with increased production capacity in CHO cells. CANX: 1.3-fold increase in plasma cell RNA levels; decrease (0.6-fold) in producing CHO cells Calr: 1.75-fold increase in plasma cell RNA levels; decrease (0.5-fold) in producing CHO cells. | (Chung et al. 2004; Harreither et al. 2015) |
| [A]**Calr (Calreticulin)** | P14211 (mouse) | | | | | | |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **TXNDC5** | Q91W90 (mouse) | CHO (+Bcl-X1) | Mab | - | Plasma and CHO transcript-omic. CHO proteomic | 4-fold increase in plasma cell RNA abundance. TXNDC5 protein levels increased 1.97-fold in a Mab-expressing CHO line overexpressing the anti-apoptotic Bcl-X1 gene. Associated with increased CHO production capacity. | (Carlage et al. 2012; Harreither et al. 2015) |
| **DNAJC1** | Q61712 (mouse) | - | - | - | Plasma and CHO transcrip-tomic | 2-fold increase in plasma cell RNA abundance. 0.9-fold change in producing CHO cells. Associated with increased CHO production capacity. | (Harreither et al. 2015) |
| **PPIB** | P24369 (mouse) | CHO | dhfr-GFP fusion | - | Plasma transcrip-tomic; CHO trans-/prote-omic | 1.37-fold increase in plasma cell RNA . Both protein and mRNA levels are up-regulated in high-producing CHO cell lines | (Nissom et al. 2006) |
| **[B]HYou1** | Q9JKR6 (mouse) | - | - | - | Plasma transcript-omic; CHO trans-/prote-omic | 2.9-fold increase in plasma cell RNA abundance. 1.9-fold transcript increase in CHO cells during cell growth, but 0.51 fold change in producing cells. 3-fold increase in protein level and 2-fold increase in RNA transcript level in high-producing cell line | (Nissom et al. 2006) |
| **[C]TMEM59** | Q9QY73 (mouse) | HEK293 | Endogen ous APP | Transient over-expression | - | Inhibition of glycosylation and decrease in Amyloid Precursor Protein processing and secretion. Does not affect secretion of soluble proteins. | (Ullrich et al. 2010) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **Transcription Factors and Transcriptional control** | | | | | | | |
| [B]**TFE3** | Q64092 (mouse) | - | - | - | Secretory pathway mapping | Transcription factor activating Golgi stress response – up-regulates Golgi structural proteins, glycosylation enzymes and transport elements | (Sasaki and Yoshida 2015; Taniguchi et al. 2015) |
| [B]**CREB3L2 (CREB family)** | Q8BH52 (mouse) | - | - | - | Transcrip-tomic | 6.1-fold increase in plasma cell RNA abundance. Up-regulates secretory pathway in non-secretory cell types. Binds enhancers of secretory genes. | (Barbosa et al. 2013; Fox et al. 2010) |
| **MLLT3** | A2AM29 (mouse) | Plasma | - | - | Transcrip-tomic | 2-fold increase in plasma cell RNA abundance. Small increase in RNA abundance in producing cells (1.13-fold). Potential role in increasing transcription speed. | - |
| [A]**XBP1s** | O35426 (mouse) | CHO-K1 | SEAP, SAMY, VEGF | Transient overexpre-ssion | Transcrip-tomic | 5.2-fold increase in plasma cell RNA abundance and associated with increased CHO production capacity. Increase in production of SEAP, SAMY and VEGF of between 2-6-fold. Co-expression with Munc18 and Sly1 further enhanced the effect of these genes. 15-85% increase in Mab production at depending on gene dose. An increase in ER/Golgi volume also seen. | (Becker et al. 2008; Harreither et al. 2015; Rajendra et al. 2015; Tigges and Fussenegger 2006) |
| | | CHO-K1 SV GS knockout | Mab | | | | |
| | | CHO-K1 | SEAP, SAMY | | | | |
| **PRDM1/ BLIMP1** | Q60636 (mouse) | Plasma | - | - | Transcrip-tomic | 7.5-fold increase in plasma cell abundance. Master regulator in B lymphocyte maturation and terminal differentiation – likely not suitable for CHO engineering. | - |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **ELL2** | Q3UKU1 (mouse) | Plasma | Endoge-nous Mab | Knock-down | Transcript omic | 4.3-fold increase in plasma cell RNA abundance. Enhances polyadenylation of Ig HC gene in plasma cells. Knockdown reduces secretory-specific forms of Ig heavy chain. | (Martincic et al. 2009) |
| **YY1** | P25490 (human) | Human and CHO | SEAP/SA MY; IgG | Over-expression | - | Transcription factor and chromatin regulator. Expression resulted in 6-fold Mab titre increase | (Tastanova et al. 2016) |
| **Vesicle formation: Cargo loading/sorting** | | | | | | | |
| **[B]Sar1** | P36536 (mouse) | - | - | - | Secretory pathway mapping | 1.7-fold increase in plasma cell RNA abundance. Active (GTP-bound) version involved in cargo selection at ERES. Associated with increased CHO production capacity. | (Harreither et al. 2015) |
| **[B]Sec12 (PREB)** | Q9WUQ2 (mouse) | HeLa, | | | Secretory pathway mapping | 1.9-fold increase in plasma cell RNA abundance. GEF responsible for activating Sar1. Associated with increased CHO production capacity. | (Harreither et al. 2015) |
| **Sec16** | O15027 (human) | Human, insect, yeast | | Knock-down; Overexpr-ession | Secretory pathway mapping | Mediates ERES formation.<br>Knockdown disrupted ERES organisation in 3 cell types. Overexpression and knockdown both inhibited ER-Golgi transport. | (Ivan et al. 2008; Watson et al. 2006) |
| **GBF1** | Q92538 (human) | Rat liver/ CHO | - | - | Proteomic Transcrip-tomic | 1.2-fold increase in producing CHO cells.<br>Presence in rat liver secretory analysis.<br>Arf-GEF with role in Golgi vesicle formation. | (Gilchrist et al. 2006) |

| Protein name | Uniprot ID | Cell Line | Product | Engineering | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| [B]**TMED2 (p24β$_1$)** | Q9R0Q3 (mouse) | Mouse, yeast, insect | Endogenous protein; VSV-G | Knockout (RNAi) | Transcriptomic | 3.6-fold increase in plasma cell RNA abundance.<br>Knockdown in *Drosophila* resulted in reduction in endogenous secretion.<br>Inhibition in yeast blocked secretion of some cargo loads but not others;.<br>Inhibition in mammalian cells inhibited VSV-G transport.<br>Mutations impair anterograde transport, triggering UPR via XBP1 splicing. | (Jerome-Majewska et al. 2010; Schuiki and Volchuk 2012; Strating and Martens 2009) |
| [A]**SRP family (14, 19, 54, SRPR, Translocon)** | P16254; Q9D7A6; P14576 (mouse) | CHO-K1 | IgG (ETE; DTE) | Transient/ Stable Over-expression | - | Engineering of CHO cells with SRP/SR proteins increases of DTE Qp by 20-80% depending upon combinations used.<br>SRPR has positive correlation with stable CHO productivity.<br>SRP19 (1.3-fold), SRP54a (5.2-fold) RNA transcripts both increased in plasma cells. Associated with increased CHO production capacity. | (Harreither et al. 2015; Kang et al. 2014; Le Fourn et al. 2014) |
| **Clec3b** | P43025 (mouse) | CHO | - | - | Transcriptomic | 4-fold increase in producing CHO cells. | - |
| **HID1** | Q8R1F6 (mouse) | Plasma | - | - | Transcriptomic | 9.4-fold increase in plasma cell RNA abundance.<br>Potential role in sorting of cargo at Golgi | (Yu et al. 2011) |
| **ASAP2** | Q7SIG6 (mouse) | CHO-K1 | - | - | Transcriptomic | 4.25-fold increase in producing CHO cells compared to parental.<br>Arf-GEF with role in Golgi vesicle formation. | (Seuter et al. 2014) |

| Protein name | Uniprot ID | Cell Line | Product | Engineering | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **AP3S1** | Q9DCR2 (mouse) | Plasma | - | - | Transcriptomics | 3-fold increase in plasma cell RNA abundance. Mammalian AP3 promotes synaptic vesicle budding. | (Boehm and Bonifacino 2002) |
| **[c]Emp47** | P43555 (yeast) | *Aspergillus oryzae* | Hetrologous cargo | Knockdown/ Overexpression | - | Gene deletion of both (independently of each other) increased secretion of cargo, whilst overexpression of each (individually) decreased protein secretion. Lectin proteins – related to LMAN2. | (Hoang et al. 2015; Schalen et al. 2016) |
| **[c]VIP36** | Q9DBH5 (mouse) | | | | | | |
| **Sec23** | Q15436 (human) | - | - | - | Transcriptomic; pathway mapping | Sec23: 1.5-fold increase in plasma cell RNA abundance. Sec24: 2.6-5-fold increase in plasma cell RNA abundance, depending on isoform. Together they form COPII coat then interacts with ERES. Both associated with increased production capacity in CHO cells. | (Harreither et al. 2015; Jensen and Schekman 2011) |
| **Sec24(D)** | O94855 (human) | | | | | | |
| **COPE1** | O89079 (mouse) | Zebrafish | - | - | Transcriptomic | 2.8-fold increase in plasma cell RNA abundance. Upregulated in Zebrafish when XBP1s is overexpressed | (Hu et al. 2007) |
| **RAB40B** | Q8VHP8 (mouse) | Breast cancer | Endogenous MMPs | - | - | Involed in sorting and regulation of secretion of Matrix metaloproteases (MMPs) in breast cancer cells. | (Jacob et al. 2013) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **Vesicle transport: Tethering and docking** | | | | | | | |
| [A]**Sly1**<br>[A]**Munc18c** | Q62991 | CHO-K1 | SEAP, SAMY, VEGF, IgG1 | Repression Stable and transient Over-expression | - | Overexpression results in 2-3-fold increase in SEAP, SAMY and VEGF production.<br>Co-expression of Sly1 and Munc18c saw a further increase in production levels to close-to 10-fold, which was further increased when XBP1 was also co-expressed | (Peng and Fussenegger 2009b) |
| [B]**Stx5** | Q8K1E0 | Plasma, HeLa | - | Knock-down (siRNA) | Transcrip-tomic | 1.6-fold increase in plasma cell RNA abundance. Associated with increased CHO production capacity. Golgi SNARE mediating vesicle docking at the Golgi. Knockdown showed essential role in mammalian constitutive secretion. | (Gordon et al. 2010; Harreither et al. 2015) |
| [B]**VTI1a** | Q9JI51 | Rat liver, Vero | - | Protein inhibition (antibody mediated) | Transcrip-tomic/Pro-temoic | 1.7-fold increase in producing CHO cells.<br>Found in Golgi or rat liver cells.<br>Inhibition of protein halts VSVG transport to cell surface. Associated with increased CHO production capacity. | (Gilchrist et al. 2006; Harreither et al. 2015; Kreykenbohm et al. 2002) |
| **SNAP29** | Q9ERB0 | Rat liver, HeLa | - | Knock-down (siRNA) | Proteomic | Found in rat liver secretory pathway, but downregulated (-1.2-fold) in plasma cells. Knock-down by siRNA showed essential role in mammalian constitutive secretion. Associated with increased CHO production capacity. | (Gilchrist et al. 2006; Gordon et al. 2010; Harreither et al. 2015) |

| Protein name | Uniprot ID | Cell Line | Product | Engineering | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **YKT6** | Q9CQW1 | Rat liver, CHO-K1 | - | - | Transcriptomic/Protemoic | Present in rat liver secretory pathway. Knockdown showed essential role in mammalian constitutive secretion. Associated with increased production capacity in CHO cells. | |
| **[B]Rab1** | P62821 | Plasma cells | - | - | Transcriptomic; mapping. | 1.8-fold increase in plasma cell RNA abundance. Interacts with ERES and recruits tether molecules, role in vesicle tethering at Golgi. | (Sztul and Lupashin 2009) |
| **RAB11a** | P62492 | - | - | - | Secretory pathway mapping | Mediates interaction between vesicle and cytoskeletal-associated motor proteins. | (Welz et al. 2014) |
| **Rab8** | P55258 | - | - | - | Secretory pathway mapping | Interacts with the Exocyst and Rab11 functions in Golgi-PM transport. Activated by the GEF Rabin8. | - (Go terms) |
| **Rab26** | Q504M8 | HEK293/ MCF-7 | Endogenous proteins | Knockdown | Secretory pathway mapping | Involved in α2-adrenergic receptors export to cell surface. Rab26 depletion/mutation reduces transport to cell surface and formation of secretory granules. | (Li et al. 2012; Tian et al. 2010) |
| **RabD** | Q9ZRE2 (*A. thaliana*) | *A. nidulans* | mRFP | Stable overexpression | - | RabD overexpression resulted in an increase of 25% (shake-flask) and 40% (bio-reactor) in expression of an RFP model cargo protein. | (Schalen et al. 2016) |

| Protein name | Uniprot ID | Cell Line | Product | Engineering | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **Sec3/Exoc1** | Q8R3S^ | CHO-K1; yeast | - | Overexpression | Transcriptomic | Exocyst component at the plasma membrane (PM), marking polarised exocytosis sites. Associated with increased production capacity in CHO cells. Overexpression in budding yeast does not change Sec3 localisation or cell growth/morphology. No information on impact upon secretion levels. | (Finger et al. 1998; Harreither et al. 2015) |
| **SCAMP2** | Q9ERN0 | Plant; PC12 (rat adrenal) | - | Inhibition; overexpression | Transcriptomic. | Small increase in Plasma cells RNA abundance (1.14-fold). Role in Golgi-PM transport in plant cells with roles in membrane fusion and vesicle recycling back to TGN. SCAMP mutation in PC12 cells inhibited $Ca^{2+}$-stimulated secretion. | (Guo et al. 2002; Liu et al. 2002) |
| **SytI** | P46096 | Mouse synapse | - | Mutation | Secretory pathway mapping | Synaptotagmins have role in vesicle binding at the PM, binding syntaxins in a $Ca^{2+}$-independent manner. SytI/SytII mutations in mouse cells decrease controlled secretion, but increased non-controlled secretion events. | (Pang et al. 2006) |
| **SytII** | P46097 | | | | | | |
| **[B]Bet1** | O35153 | Rat liver, CHO-K1, plasma | VSV-G | Inhibition | Transcriptomic; proteomic | 2.5-fold increase in plasma cell RNA abundance. Associated with increased production capacity in CHO cells. Present in rat liver secretory pathway. Facilitates fusion of COPII vesicles to each other to form transport intermediates, and binding at *cis*-Golgi. Interacts with P115. Inhibition (antibody-mediated) inhibits fusion of COPII vesicles at the Golgi. | (Allan et al. 2000b; Gilchrist et al. 2006; Harreither et al. 2015; Sztul and Lupashin 2006) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **[B]Bet3** | O55013 | HeLa | YFP-tagged VSV-G | Inactiva-tion | Secretory pathway mapping | Inactivation leads to 70% decrease in YFP-VSV-G transport cargo, with accumulation at the Golgi. Subunit of TRAPPI tether complex. Role in COPII vesicle tethering to each other then to the Golgi, as well as Rab1 activation. | (Sztul and Lupashin 2006; Yu et al. 2006) |
| **SCFD2** | Q8BTY8 | - | - | - | - | Sec1/Munc18 family member – role in vesicle tethering and fusion at Golgi membrane. Maintained in animals and plants, but not insects/ yeast. | (Iordanova and Fasshauer 2016) |
| **[B]P115/USO1** | Q9Z1Z0 mouse | Rat liver, plasma | - | - | Proteomic /Transcri-ptomic | 2.6-fold increase in plasma cell RNA abundance. Protein present in rat liver secretory pathway. Tether involved in Golgi transport vesicle tethering at *cis*- and medial Golgi. | (Gilchrist et al. 2006) |
| **VPS45** | P97390 | Rat Liver, yeast | - | - | Proteomic | Present in rat liver secretory pathway. Sec1/Munc18 family member | (Gilchrist et al. 2006) |
| **Gorasp2** | Q99JX3 | Rat liver, plasma, amoebe | - | - | Transcrip-tomic/ proteomic | 1.4-fold increase in plasma cell RNA abundance. 1.6-fold increase in CREB-expressing cells. In amoebe is not required for conventional protein secretion but has role in unconventional secretion during cell development. | (Barbosa et al. 2013; Gilchrist et al. 2006; Kinseth et al. 2007) |
| **Golph3** | Q9CRA5 | Rat liver | - | - | - | 1.36-fold increase in plasma cell RNA abundance. Protein present in rat liver secretory pathway. PI4-kinase effector, which regulate mediated export from Golgi via vesicle budding and lipid dynamics | (Gilchrist et al. 2006; Graham and Burd 2011) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **VAMP4** | O70480 | CHO | - | - | Transcrip-tomic | Associated with increased production capacity in CHO cells. 1.5-fold increase in gene expression in higher-producing CHO cells at 32°C | (Yee et al. 2009) |
| **Miscellaneous plasma cell hits** | | | | | | | |
| **CRELD2** | Q9CYA0 | Plasma cell | - | - | Transcript-omic | 4.9-fold increase in plasma cell RNA abundance. High levels of CRELD2 expression in endocrine tissues. ER_stress related, being downstream of ATF6. Interacts with chaperones and has some PDI-like activities. | (Hartley et al. 2013; Maslen et al. 2006; Oh-hashi et al. 2009) |
| **FNDC3A** | Q8BX90 | Plasma cell, odonto-blasts | - | - | Transcript-omic | 3.3-fold increase in plasma cell RNA abundance. 10-fold increase in transcript abundance as pulpal precursor cells mature to dentin synthesising and secreting cells. | (Carrouel et al. 2008) |
| **RRBP1** | Q99PL5 | Plasma yeast, fibroblast | - | - | Transcript-omic | 3.0-fold increase in plasma cell RNA abundance. Activated in XBP1-overexpressing fibroblasts which see a >2-fold increase in ER abundance. Has been used to increase Mab production in yeast. | (Chiba et al. 2012; Sriburi et al. 2007) |
| **FNDC3B** | Q6NWW9 | Plasma cell | - | - | Transcript-omic | 1.4-fold increase in plasma cell RNA abundance. Amplified in over 20% of cancers. FNDC3B overexpression induces several cancer pathways. | (Cai et al. 2012; Lin et al. 2016) |
| **TBL2** | Q9R099 | Plasma cell, human | - | Knock-down | Transcript-omic | 1.5-fold increase in plasma cell RNA abundance. TBL2 knockdown in human cells can impair ATF4 induction under ER stress, having similar effects to knockdown of PERK pathway components. | (Tsukumo et al. 2014) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **NBAS** | A2RRP1 | Plasma cell | - | - | Transcript-omic | 2.2-fold increase in plasma cell RNA abundance. Role in retrograde transport. | - |
| **CAB45/Sdf4** | Q61112 | Plasma cell. Rat cell | - | - | Transcript-omic | 1.5-fold increase in plasma cell RNA abundance. Inhibition with antibodies reduced secretion of amylase in rat cells. Interacts with SM proteins and required for $Ca^{2+}$-dependent secretory cargo sorting at TGN. | (von Blume et al. 2012) |
| **Lman1/ ERGIC53** | Q9D0F3 | Plasma cell, CHO, human | Blood Factors V and VIII | Mutation | Transcript-omic | LMAN1: 4.3-fold increase in plasma cell RNA abundance; associated with increased production capacity in CHO cells.<br>MCFD2: 2.4-fold increase in plasma cell RNA abundance.<br>Together LMAN1/MCFD2 complex (1:1) play role at ERGIC. Roll in transport and secretion of blood factors. Mutation leads to coagulation of blood factors | (Harreither et al. 2015; Zhang et al. 2005) |
| **MCFD2** | Q8K5B2 | Plasma cell, human | | | | | |
| **IRF4** | Q64287 | Plasma cell | - | - | Transcript-omic | 3.0-fold increase in plasma cell RNA abundance. Transcription factor required for terminal differentiation of plasma cells to B cells, acting upstream of XBP1 | (Klein et al. 2006) |
| **ARID3A** | Q62431 | Plasma cell | - | - | Transcript-omic | 2.7-fold increase in plasma cell RNA abundance. Activator of Ig HC transcription. | (Puissegur et al. 2012) |
| **CopZ2** | Q9jHH9 | Plasma cell | - | - | Transcript-omic | 8.8-fold increase in plasma cell RNA abundance. COPII subunit likely involved in regulating coat assembly by acting as an adaptor protein. | - |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **Lipid biosynthesis** | | | | | | | |
| [B]**PRKD1** | Q62101 | HeLa | - | Inhibition | Lipid synthesis mapping | PRKD1 activity inhibition in HeLa leads to Golgi tubulation with vesicles not separating from the TGN. Recruited to Golgi by DAG. | (Liljedahl et al. 2001) |
| [A]**CERT (S132A mutation)** | Q9EQG9 | CHO 1-15(500) | tPA | Stable over-expression | Lipid synthesis mapping | 29-45% increase in tPA Qp, 26% increase in IgG Qp. CERT traffics ceramide (DAG pre-cursor) from the ER to the Golgi. Phosphorylation of CERT at S132 reduces its lipid transfer ability. | (Florin et al. 2009; Rahimpour et al. 2013) |
| | | CHO DG44 | IgG (HAS) | | | | |
| [B]**SGMS1** | Q8VCQ6 | - | - | - | Lipid synthesis mapping | Decrease in plasma cell RNA abundance. Converts ceramide to DAG through a reversible process. | (Bard and Malhotra 2006) |
| **FADS2** | Q9Z0R9 | Plasma cell | - | - | Transcrip-tomic | 7-fold increase in plasma cell RNA abundance. ER-located oxidoreductase with role in lipid metabolism. Increased levels of saturated FAs results in ER stress.. | (Jump 2011; Leamy et al. 2014) |
| **PItp** | P55065 | CHO | IgG | - | Transcrip-tomic | 1.6-fold increase in higher-producing CHO cells at 32 °C. Facilitates exchange of PI and PC between lipid bilayers, leading to high PI/PC ratio in Golgi, promoting budding from TGN. | (Cleves et al. 1991; Yee et al. 2009) |
| **PLD3** | O35405 | MAK | | | Transcrip-tomic | 2.3-fold increase in higher-producing butyrate-treated cells | (Yee et al. 2009) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **ER stress response** | | | | | | | |
| **[A]Ero1Lα** | Q8R180 | CHO-S | Mabs | Stable co-expression with XBP1 | Transcrip-tomic | Co-expression with XBP1 resulted in a 5-6-fold increase in Mab productivity (greater than just XBP1 expression), increase in ER volume and better recovery from ER stress. Associated with increased production capacity in CHO cells. | (Cain et al. 2013; Harreither et al. 2015) |
| **MZB1** | Q9D8I1 | MZ B/B1; NSO; Plasma | IgM | Transient overexp-ression; deletion | Transcrip-tomic | 3.4-fold increase in plasma cell RNA abundance. Overexpression in MZ B cells resulted in 5-fold increase in antibody secreting cells. Deletion in plasma cells reduces antibody secretion. Possibly Plasma cell specific involved in IgM folding and secretion. | (Flach et al. 2010; Rosenbaum et al. 2014) |
| **PERK/ EIF2AK3** | Q9Z2B5 | Mouse embryo-nic stem; NIH-3T3 | - | Mutation; Over-expression | - | PERK mutation in mouse embryonic stem cells leads to increased protein synthesis levels but also higher ER stress levels. Over-expression in NIH-3T3 cells induced G1 phase cell cycle arrest. | (Brewer and Diehl 2000; Harding et al. 2000) |
| **ATF6** | F6VAN0 | CHO-K1; HeLa | - | Activation | Transcrip-tomic | 2.3-fold increase in plasma cell RNA abundance. Activation of ATF6 with small molecules reduces secretion of miss-folded/destabilised proteins. Over-expression in HeLa cells saw increased transcription from ERSE. Expression in CHO cells increased product titre, though this was product dependent. | (Johari et al. 2015; Li et al. 2000; Plate et al. 2016) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **IRE1/Ern1** | Q9EQY0 | CHO-K1; COS-1 | - | Over-expression | Transcrip-tomic | Associated with increased production capacity in CHO cells. Ire1 overexpression activates BiP and CHOP in mammalian cells. | (Wang et al. 1998) |
| **eIF2α/eiF2s1** | Q6ZWX6 | CHO-K1; PC12 | - | Activation | Transcrip-tomic | 1.8-fold increase in plasma cell RNA abundance. eIF family members. Activation, helps protect cells from ER stress. Associated with increased production capacity in CHO cells. | (Boyce et al. 2005; Harreither et al. 2015) |
| **ATF4** | Q06507 | CHO-K1SV GS KO | Mab | Transient over-expression | Transcrip-tomic | 1.45-fold increase in plasma cell RNA abundance. Transient overexpression in CHO resulted in a decrease in Mab production. Stable overexpression in CHO resulted in 2-fold increase in human antithrombin III productivity. Associated with increased production capacity in CHO cells. | (Harreither et al. 2015; Ohya et al. 2008; Rajendra et al. 2015) |
| | | CHO 13D-15D | AT-III | Stable over expression | | | |
| **[C]BiP (GRP78/ HSPA5)** | P20029 | CHO2F5 | Mab | Stable over expression | - | Stable overexpression resulted 34% decrease in specific secretion. | (Borth et al. 2005; Dorner and Kaufman 1994; Dorner et al. 1988; Harreither et al. 2015; Nissom et al. 2006) |
| | | CHO DUKX | tPA | Repression /over-expression | - | Recombinant protein expression decreased upon BiP overexpression and increased upon BiP reduction. | |
| | | CHO DUKX | tPA | Stable repression | - | Reduction of endogenous BiP levels improved secretion of recombinant tPA in CHO cells | |
| | | CHO, CHO-K1 | - | - | Transcrip-tomic, proteomic | Downregulated in high-producing CHO cells, but also associated with increased production capacity in CHO cells. | |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| **CHOP/DDIT3** | P35639 | CHO-K1 SV | Mab Mab | Transient/ stable Over-expression | - | Transient overexpression resulted in a decrease in Mab expression. | (Nishimiya et al. 2013; Rajendra et al. 2015) |
| | | CHO/ COS1 | | | | Transient overexpression in COS-1 saw a 2-fold increase in IgG production. Overexpression alongside BiP, ERdj3/5 resulted in a 3-fold increase in IgG. Overexpression alongside PDI and ERO1-Lβ saw a 2-fold increase. | |
| **TRIB3/TRB3** | Q8K4K2 | HEK293 | - | Knock-down | - | Role in UPR-mediated apoptosis. Knockdown lead to reduced levels of ER stress-related cell death. | (Ohoka et al. 2005) |
| **[C]TMEM214** | Q8BM55 | | - | KD/Over-expression | - | 2.8- fold increase in plasma cell RNA abundance. Overexpression induces apoptosis via CASP4. | (Li et al. 2013) |
| **mTOR (c1)** | Q9JLN9 | - | - | - | - | Cell growth regulator Activated in early stage of UPR activation. | (Hussain et al. 2014) |
| **NRF1/NRF2** | Q9WU00 | mouse | - | Knock-down | - | eIF2α-activating transcription factor. Deficiency results in an increase in oxidative stress and apoptosis | (Leung et al. 2003) |
| **[A]GCLM** | O09172 (mouse) | Epi-CHO (K1 line) | Transient IgG1 | Stable over-expression | - | Antioxidant role in glutathione (GSH) synthesis. GSH increase linked to increased CHO productivity levels. Stable expression of GCLM in CHO transiently expressing IgG1 saw Qp increase between 1.37-1.95-fold depending on growth phase, with a 1.71-2.0-fold increase in titre over 15 days culture and 75% more high producer clones compared to wild type. | (Orellana et al. 2017; Orellana et al. 2015) |

| Protein name | Uniprot ID | Cell Line | Product | Engineer-ing | Other info | Effect | References |
|---|---|---|---|---|---|---|---|
| [A]**KIT (onco-D816 mutant)** | P05532 (mouse) | CHO-K1 | Stable GFP-Fc | Expression | - | Expression of mutated KIT resulted in increased global protein synthesis, cell proliferation, stress resilience and a 2-fold increase in productivity. | (Mahameed and Tirosh 2017) |
| [B]**PRDX4** | Q13162 (human) | Plasma cell | - | - | - | 5-fold increase in plasma cell RNA abundance. Role in hydrolysis of hydrogen peroxide, protecting against oxidative stress. | (Shaffer et al. 2004) |
| **Other miscellaneous targets** | | | | | | | |
| **Largen (PRR16)** | Q569H4 (human) | Human cells | - | Over-expression | - | Molecular regulator of mammalian cell size. Overexpression increased cell size. | (Yamamoto et al. 2014) |

## Appendix 2: Gene sequences for CHO cell engineering.

Genes selected for engineering of the CHO cell biosynthetic and secretory pathways were designed with 5' and 3' flanking regions of DNA enabling them to be inserted into a MedImmune-specific single gene expression vector or multi gene expression vector (MGEV). This appendix shows the cDNA sequence for each accessory gene used in the project, from the ATG start codon to the TAA/TGA stop codon. Flanking regions are not included.

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **PRKD1**<br>CHO<br>XM_016974047.1<br>2522 bp | atggcttgctccatcgtggaccagaagttccctgaatgtggtttctatggactgtatgataagatcctgcttttccgtcatgatcctacctctgaaaacatccttcagctggtgaaaatggcaagtgatatccagg<br>aaggagatcttattgaagtggtcctgtcagcttcagccacctttgaagactttcaaatccggccccatgctctcttcgttcattcctacagagcccccgctttctgtgaccactgtggagaaatgctctggggact<br>ggtgcgccaaggccttaaatgtgaaggatgtggtctgaattaccataagagatgtgcatttaaaatccctaacaattgcagtggtgtgagaaggagaaggctctcaaacgtttccctcactgggcttagcact<br>gtccgcacttcatctgctgagttctccaccagtgcccctgatgagcctctactgtctcctgtaagccctggctttgagcaaaagtcaccatcagagtcatttattggccgtgagaagaggtcaaattcgcaatcg<br>tatattggacggccaattcaacttgacaagattttgatgtcaaaggtgaaggtgccgcatacttttgtcatccactcctacactcgacccacagtatgccagttctgcaaaaagctcctcaagggcctcttccgg<br>cagggcttgcagtgcaaagattgcagattcaactgtcacaaacgctgtgcaccaaaagtaccaaacaactgcttgggtgaagtgaccatcaatggagatttgcttagccctggagcagaatctgatattgtta<br>tggaagaaggaagtgatgacaatgacagtgagcggaacagcggactcatggatgacatggatgaggccatggtccagaatgctgagatggctatggctgagggccagggtgatggtgcggaaatgcagg<br>atccagatgcagaccaggaggactctaacaggaccatcagcccttcgacaagcaacaacatccccctcatgcgggtggtgcagtctgtcaagcacacaaagcggaggagcagcactgtgatgaaggaag<br>gatggatggtccattacaccagcaaagacacactgcggaaaagacattactggagattggacagcaagtgcatcacactcttccaaaatgacacaggaagccggtactacaaggaaattccctttatcagaa<br>attttatgtctggaacctgcgaaaccttcagcattaattcccattggagctaacccccactgttttgaaatcactacagcaaatgtagtgtattacgttggagaaaatgtggtcaatccttcaagtcccccaccaa<br>acaacagtgttttcaccagcggcattggtgcagatgtggccagaatgtgggaggtggccattcagcatgctctcatgcctgtcatccccaagggctcctctgtggggtcaggaaccaatttacataatgtttct<br>gtgagcatttcagtttcaaattgccagatccaggaaaatgtggacatcagcaccatctatcagatttttcctgatgaagttttgggttctggacagtttggaattgtttatggaggtaaacatcgtaaaacagga<br>agagatgtagctattaagattattgacaaattaagatttccaacaaaacaagaaagtcagcttcgtaatgaggttgcaattttacagaaccttcatcaccctggtgttgtaaatttggagtgtatgtttgagacg<br>cctgaaagagtgtttgttgttatggaaaaactccatggagacatgctggagatgatcttgtcaagtgaaaagggcaggttgccagaacacataacgaagtttttaattactcagatactagtggctttgcggca<br>tcttcatttcaaaaacatcgttcactgtgacctcaagccagaaaatgtgttgctggcatctgccgatcctttccctcaggtgaaactttgtgattttggttttgcccggatcattggagagaagtctttccggaggt<br>cagtggtgggtaccccagcctacctggcacctgaggttctgaggaacaagggctataatcgctctctagacatgtggtctgttggggtcatcatctatgtgagcctcagtggcaccttccctttttaatgaagatg<br>aagacatccatgatcagatccagaatgcagccttcatgtatccgcccaatccgtggaaggagatttctcatgaagccattgatcttatcaataacttgctgcaagtgaaaatgagaaagcgctacagtgtgga<br>caaaaccttgagtcacccgtggctacaggactatcagacctggttagatttacgagagctggaatgcaagattggagaacgctatattacccacgaaagtgatgactcgaggtgggaacagtatgccggtg<br>accaggggctgcagtatcccgcacacctgatcagcctgagcgctagccacagcgacagtcctgaggctgaagagagagagatgaaagccctcagtgagcgtgtcagcatcctctga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **SGMS1**<br>CHO<br>XM_007648117.2<br>1244 bp | atgaaggaagtggtttattggtcacccaagaaggtggcagactggctgctggagaatgctatgcctgaatactgtgagcctctagaacatttcacaggccgggacttgatcaatctaacccaagaggatttca<br>caaaaccccactgtgccgagtctcatctgacaacgggcagcggctcttagacatgatcgagaccctgaagatggagcaccatatggaagcgcacaagaacggccacgccaatgggcacctcagcgttgg<br>cgtcgacattcccaaccctgatggcggcagcttcagcatcaagatgaaacccaatggaatgccaaatgggtataggaaagagatgatcaagatccccatgccagaaccagagcgctctcagtaccccatgg<br>agtggggcaagactcttctggccttttctttatgcacttttcctgttttgttctcactacagtgatgatctcggtcgtccatgaacgagtacctcctaaggaggtgcagcctccactaccggacacgtttttgaccatt<br>ttaaccgggtgcagtgggccttttctatttgcgaaattaacggcatgatccttgtaggactctggttaattcagtggctggtcttaaaatacaagtctattattagcagaagatttttctgcatagttggcacgctg<br>tacctgtatcggtgtattacaatgtatgtaactacactcccagtacctggcatgcatttcaactgttctccgaagctctttggagactgggaagctcaagtgcggagaataatgaagctcattgctggaggtggc<br>ttgtccatcacaggctctcacaacatgtgtggggactatctatacagtggccacacggtcatgctaacgctcacctatttatttatcaaagagtattccccgaggcgactctggtggtaccactggatttgctgg<br>ctccttagcgtcgttggaatcttctgtattctcttagcgcatgaccactacactgtggacgtggtggtggcctactacatcaccacaaggctcttctggtggtatcacacaatggccaatcagcaagtgctgaag<br>gaagcttcccagatgaacctcctggccagggtgtggtggtacaggccatttcagtactttgaaaagaatgtccaaggaattgtacctcgatcttaccattggcccttcccctggccagtagtccaccttggtag<br>gcaagttaaatatagccggttggtaaacgacacataa |
| **CERT**<br>**S132A mutation (TCA -> GCT)**<br>CHO<br>XM_007648610.1<br>1796 bp | atgtcggataaccagagctggaactcgtcgggctcggaggaggatccggagacggagtccgggccgcctgtggagcgctgcggggtcctcagcaagtggacaaactatattcatgggtggcaggatcgtt<br>gggtagttttgaaaaataatactttgagttactacaaatctgaagatgagacagagtacggttgcaggggatccatctgtcttagcaaggctgtgatcacacctcatgattttgatgaatgtcggtttgatatca<br>gtgttaatgatagtgtttggtatcttcgtgctcaggacccagatcacagacagcagtggatagatgccattgaacagcacaagactgaatcaggatatggatctgagtccagcttacgtagacatggc**<u>GCT</u>**a<br>tggtgtcactggtgtctggagcaagtgggtactctgctacatccacatcttcattcaagaaaggacacagtttacgtgagaaattggctgaaatggaaacttttagagacatcttatgtagacaagttgacact<br>ctccaaaagtactttgatgtctgtgctgatgctgtctccaaggatgaacttcaaagggataaagtggtagaagatgatgaagatgacttccctacaactcgttctgatggagacttttgcacaataccaatggt<br>aataaggaaaaattatttccacatgtaacccccaaaggaattaatggtatagactttaaaggggaagcaataacttttaaagcaactactgctggaatccttgctacactttctcattgtattgaattaatggta<br>aaacgggaagagagctggcaaaaaagacatgataaggaaatggagaagagaagacgattagaggaagcatacaagaatgcaatggcagagcttaagaagaaaccccgttttggagggcctgattatga<br>agaaggtccgaacagtctgattaatgaggaggagttctttgatgctgttgaagctgctcttgacagacaagataaaatagaggaacagtcacagagcgagaaggtcaggttacactggcctacacctttgcc<br>atctggagatgccttttcttctgttgggacccatagatttgtacaaaaggttgaagagatggtacagaaccacatgacttactcattacaggatgtaggtggtgatgcgaattggcaactagttgtagaagaag<br>gagaaatgaaggtatacagaagagaagtcgaagaaaatggaattgttctggatcctttgaaagctacccatgcagttaaaggtgttacaggacacgaagtctgcaattacttttggagtgttgatgttcgcaa<br>tgactgggaaactactatagaaaacttccatgtagtggaaacattagctgataatgcaatcatcatttatcaaacgcacaagagagtgtggcctgcttctcagagagatgtactgtatctttctgctattcgaaa<br>gatcccagccttgactgagaacgaccctgagacttggatagtttgtaatttttctgtggatcatgacagcgctcctctgaacaatcgatgtgtccgtgccaaaatcaatgttgctatgatttgtcaaaccttagta<br>agcccaccagagggaaaccaggaaataagcagagacaacattctgtgcaagattacatatgtagctaatgtgaacccaggaggatgggcaccagcctcggtgttaagagcagtggcaaaacgagaatat<br>cctaaatttctaaaacgttttacttcttacgtccaagaaaaaactgcaggaaaaccaattttgttttag |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **Bet1**<br>CHO<br>XM_003496956.3<br>357 bp | atgaggcgtgcaggcctgggtgaaggaggacctcctggcaactatgggaactatggctatgctggtagtggctataatgcctgtgaagaagaaaacgacagactcactgaaagtctgagaagcaaagtga<br>ctgccatcaaatctctttccattgaaataggccatgaagttaaaaatcaaaacaaactactagctgaaatggattcacagtttgattctacaactggatttctaggtaaaaccatgggaagactgaagattttgt<br>ccagaggaagccaaacaaaattgctatgctatatgatgttgttttcattgtttgtctttttgtcatttactggattattaaactgaggtga |
| **Rab1a**<br>CHO<br>XM_003498103.2<br>619 bp | atgtccagcatgaatcccgaatatgattatttattcaagttacttctgattggcgattctggggttggaaagtcttgcctcctcctcaggtttgcagatgacacgtacacggaaagctacatcagcacgattggtg<br>tggatttcaagatccgaactatagagttagacgggaagacaatcaagcttcagatatgggacacagcaggccaggaaagatttcgaacaatcacctccagttactacagaggagcccatggcatcatagtt<br>gtgtatgatgtgacagatcaggagtccttcaataatgttaaacagtggctgcaagagatagatcgctatgccagtgaaaatgtcaacaagttgttggtagggaacaaatgtgacctgaccacaaagaaagta<br>gtagattacacaacagcgaaggaatttgcagattcccttggaattccatttttggaaaccagtgctaagaatgcaacgaatgtagaacagtctttcatgacgatggcagccgagattaaaaagcgaatgggc<br>cctggagcaacagctggtggtgcagagaagtccaatgttaaaatccagagcactccagtcaagcagtcaggtggaggttgctgctaa |
| **Rab11a**<br>CHO<br>XM_003508776.2<br>651 bp | atgggcacccgtgacgacgagtacgactacctctttaaagttgtccttattggagattctggtgttggaaagagtaacctcctgtctcgatttactcggaatgagttcaatctggaaagcaagagtaccattgga<br>gtagagtttgcaacaagaagcatccaggttgatgggaaaacaataaaggcgcagatatgggacacagcagggcaggagcggtacagggctataacatctgcatactatcgtggagcagtaggcgccttac<br>tggtttatgacattgctaagcatctcacatatgaaaatgtagagcgatggctgaaagaactgagagaccatgctgatagcaacattgttatcatgcttgtgggtaataagagtgatttacgccatctcagggca<br>gttcctacagatgaagcaagagcttttgcagagaagaatggtttgtcgttcattgagacatctgctctagattctacaaatgttgaagctgcttttcagacaattctaacagagatataccgcattgtttctcaga<br>agcaaatgtcagacagacgtgaaaatgacatgtctccaagcaacaatgtggttcctattcacgttccacctaccactgaaaacaagccaaaggtgcagtgctgtcagaacatctaa |
| **CRELD2**<br>CHO<br>NM_001244136.1<br>1048 bp | atgcacctgccgcccgctgccgcagtcgggctgctactgctgctgctgccgcctcccgcgcgcgtggcctcccggaagccgacaatgtgccagaggtgccgggcgctggtggacaagttcaaccaggggat<br>ggccaacacggccaggaagaatttcggcggcggcaacacggcgtgggaggagaagagtctgtccaagtacgaattcagtgagattcggctcctggagattatggagggcctgtgtgacagcaacgacttt<br>gaatgcaaccaactcttggaacagcatgaggagcagctagaggcctggtggcagacactgaagaaggagtgccctaacctatttgagtggttctgtgtacacacactgaaagcatgctgtcttccaggcacc<br>tatgggccagactgtcaggaatgccagggtgggtctcagaggccttgtagcgggaatggccactgcgacggagatggcagcagacagggcgacgggtcctgccagtgtcacgtaggatacaaggggccg<br>ctgtgtatcgactgcatggatggctacttcagcttgctgaggaacgagacccacagcttctgcacagcctgtgatgagtcctgcaagacatgctcaggtccaaccaacaaaggctgtgtggagtgcgaagtg<br>ggctggacacgtgtggaggatgcctgtgtggatgttgacgagtgtgcagcagagaccccaccctgcagcaatgtacagtactgtgaaaatgtcaacggctcctacacatgtgaagagtgtgattctacctgtg<br>tgggctgcacaggaaaaggcccagccaattgtaaagagtgtatctctggctacagcaagcagaaaggagagtgtgcagatatagatgaatgctcattagaaacaaaggtgtgtaagaaggaaaatgaga<br>actgctacaatactccagggagctttgtctgcgtgtgtccggaaggtttcgaggaagacagaagatgcttgtgtacagacagcagaaggcgaagtggcagaggaaagtcccacacagccaccctcccatga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **CREB3L2 (cytosolic domain)** CHO XM_003500972.3 948 bp | atggaggtggagccatctccaacatcaccagcgcctctcatccaggctgaacacagctactccctgagcgaggagccccgggctcagtcaccattcacccatgtggctaccagcgatggcttcaatgacgag gaagtcgagagtgagaagtggtacctgtctccagactttccttcggccaccgtcaagacagagcccatcacagaggagcagcccccagggcttgttccctctgtcactctgaccatcacagccatttctactcc ttttgaaaaagaagagtcccctctggatatgaatgctggggtctttggtctccaacaaagcatttatcctaggatctcttgttattcttctgggatggatggaatgcatatgttctcttcatcctcagcctccgtgga tcagctgcacttaccaccaacaccacctagtagccacagcagtgactccgaggggagcctgagccctaacccccgcctgcatcccttcagcctgtctcaggcccacagccctgccagagccatgccccgggg cccctctgcgttgtccacatctcctctcctcacagctccacataagctacagggatctggtccactggtcctgacggaggaagagaagaggaccttgattgctgagggctatcctattcctaccaaactgcccct gacaaaatctgaggagaaggccctgaagaaaatccggagaaaaatcaagaataagatttctgcccaagaaagcaggagaaagaagaaagaatatatggacagcctggagaaaaaagtggagtcatgtt caactgagaacttggaacttcggaagaaggtggaggtgctagagaacaccaataggactctgcttcagcaacttcagaagcttcaggctctggtgatgggcaaggtgtcccgatcctgcaagttagctggca cccagactggcacctaa |
| **PREB** CHO XM_003497042.2 1254 bp | atgggcaggcgccggggtgtggagctgtaccgggccccgttcccgttgtacgcgctccaggtggaccccaagaacgggctgctcattgctgcgggcggaggaggagccgccaagaccggcatcaagaatg gtgtgcactttctgcagctagagcaaatcaatggctgcctgagcgcctccttgctacactctcatgacacagagacacgggccactatgaacttggcactggctggtgacatccttgctgccggacaggatgc ccgatgtcagcttctacgttttcatgtccatcaacagaagcgcaataaaacagagaagtcaggttccaaggagcagggacctcgacaaaggaaggggggcagctccagcagagaagaaatctggagcaga agtacaccctgaaggggtagaactcaaagtaaaaaatttggaggcagtacagacagactttagcactgaaccactgcagaaagttgtgtgcttcaaccacgataataccctgcttgctactggagggactga tggccatgttcgtgtctggaaggtacctagcctggagaaagttctggatttcaaagcacatgaaggagagattggagatttggctttgggtcctgatggcaagttggtgactgtaggctgggacttaaaggcct ctgtgtggcagaaggagcaactagtgacacagctacagtggcaagagaatggacccgcctcttctgacacaccataccgctaccaggcctgcaggtttgggaaggttccagatcaacctgggggctgcga ctcttcacagtgcagatccccacaagcgcctgcgtcagcccccaccctgctacctcacagcctgggacagttccacattcttgcctcttcggaccaggccctgtggccatgaagtcatctcctgcctcagtgtc agtgaatccggtaccttcctaggcttaggcacggtcactggctctgtcgccatctatatagctttctctctccagcgcctctattatgtgaaggaggcccatggcattgtggtgacagatgtaacctttctacctg agaagggttgtggcccagagcttcttggaccccatgaaactgctcttttctctgtggctgtggatagtcgttgccagttgcacctgctgccctcacggcggagtgttccagtatggctactgctgctgatgtgtgt cggccttattattgtgaccatcctgctgctccagagtgcctttccaggttttctttag |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **HSP70/ HSPA1a** CHO XM_003508473.3 1926 bp | atggccaagagcacggcgatcggcatcgacctgggcaccacctactcgtgcgtgggcgtgttccagcacggcaaggtggagatcatcgccaacgaccagggcaaccgcacgacccccagctacgtggcct tcaccgacaccgagcggctgatcggagacgccgccaagaaccaggtggcgctgaacccgcagaacaccgtgttcgacgcgaagcggctgatcggccgcaagttcggcgacgcggtggtgcaggcggac atgaagcactggcccttccaggtggtgaacgacggcgacaagcccaaggtgcaggtgagctacaagggcgagacgcgggccttctaccccgaggagatctcgtccatggtgctgacgaagatgaaggag gtcgccgaggcgtacctgggccacccggtgaccaacgcggtgatcacggtgcccgcctacttcaacgactcgcagcggcaggccaccaaggacgcgggcgtgatcgcgggtctgaacgtgctgcggatca tcaacgagcccacggcggccgccatcgcctacgggctggaccgctcgggcaagggcgagcgcaacgtgctcatcttcgacctgggggggcggcacgttcgacgtgtccatcctgacgatcgacgacggcat cttcgaggtgaaggccacggcgggggacacccacctgggcggcgaggacttcgacaaccggctggtgagccacttcgtggaggagttcaagaggaagcacaagaaggacatcagccagaacaagcgcg cggtgcggcggctgcgcacggcctgcgagcgcgccaagaggaccctgtcgtccagcacccaggccagcctggagatcgactccctgttcgagggcatcgacttctacacgtccatcacgcgggcgcggttc gaggagctgtgctcggacctgttccgcggcacgctggagcccgtggagaaggcgctgcgcgacgccaagctggacaaggcgcagatccacgacctggtgctggtgggcggctccacgcgcatccccaag gtgcagaagctgctgcaggacttcttcaacgggcgcgacctcaacaagagcatcaacccggacgaggcggtggcctacggggcggcggtgcaggcggccatcctgatgggggacaagtctgagaacgtg caggacctgctgctgctggacgtggcgccgctgtcgctgggtctggagacggcgggcggcgtgatgacggcgctcatcaagcgcaactccaccatccccaccaagcagacgcagaccttcaccacctactc ggacaaccagcccggcgtgctgatccaggtgtacgcgggcgagcgggccatgacgcgcgacaacaacctgctgggccgcttcgagctcagcggcatcccgccggcgcccaggggcgtgccccagatcga ggtgaccttcgacatcgatgccaacggcatcctgaacgtcacggccaccgacaagagcaccggcaaggccaacaagatcaccatcaccaacgacaagggccggctgaggaaggaggagatcgagcgca tggtgcaggaggccgagaggtacaaggccgaggacgaggtgcagcgcgagagggtggcggccaagaacgcgctcgagtcctacgccttcaacatgaagagcgccgtggaggacgagggcctcaaggg caagatcagcgaggccgacaggaagaaggtgctggacgggtgtcaggaggtcatctcctggctggacgccaacacgctggccgacaaggaggagttcgtgcacaagcggcaggagctggagagggtgt gcggccccatcgtcagcgggctctaccagggtgcgggtgctcccggggcgggcggcttcggggcccaggcgcccaagggaggctccgggtcggggcccaccatcgaggaggtggattag |
| **Stx5a** CHO XM_003514181.3 435 bp | atgatcccgcggaaacgctacggatctaagaacacagatcagggtgtctacctgggtctctcaaagacacaggttctgtcccctgcaactgctgtcagtagcagcagcgacatcactcctttgcccaccccgg tggccctggtcccttcccctcccgacaccatgtcctgccgggatcggacccaggagttcctgtctgcctgtaagtcgctgcagagccgtcagaatggaatccaaacaaataaaccagctctacgggctgccag acaacgcagtgaatttacccttatggccaagcgcattggaaaagatctcagcaatacatttgccaagctggagaagctaacaatcttggcaaagcgcaagtccctctttgatgataaagcagtagaaattga ggagctaacatacatcatcaaacaggtgagctgctag |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **P115**<br>CHO<br>XM_007652162.1<br>2901 bp | atgaatttcctccgcggggtgatgggggccagagtgccggaccccagcacacagaagccgagacgattcagaagctctgtgaccgagtagcttcatcaactttactggatgaccgaagaaatgctgtgcg<br>tgctcttaaatcattatctaagaaataccgcttggaagtaggaatccaagctatggaacatcttatccatgttttacaaacagatcgttcggattctgaaataatagcttatgctttggatacactctataatataa<br>tatctaatgatgaagaggaagaaatagaagaaaattctgcaagacagactgaggatttgggaagccagttcacagaaattttcatcaagcagccagaaaatgtcactctcctgttgtctctgttggaggaatt<br>tgatttccatgtccgttggcctggcgtgagactcctgacttctcttttaaaacaattagggcctccagtgcagcagattatcttagtcagtcccatgggtgtttcaagactgatggacttactggcagattccaga<br>gaaattatacgtaatgatggtgtcctactgctgcaggccttaacaaggagcaatggagcaatccagaaaattgtggcttttgaaaatgcttttgagaggctgctggacatcattacagatgaggggaacagtg<br>atggaggtatagtagtagaagattgtttgatcttacttcaaaacttgttaaagaacaacaattccaatcaaaacttttttaaggaaggctcatatattcaacgtatgaaaccttggtttgaagttgcagatgaaa<br>attctggttggtcagcacagaaagtgaccaatctacatttgatgctacagcttgtccgagtactagtgtctcccaccaaccctcctggtgcgaccagcagctgccagaaggccatgtttcagtgcgggttactg<br>cagcaactttgtaccatcctgatggctacaggaattcctgctgatatcctgaccgagaccataaatactgtatctgaagttattcgaggctgccaagtcaaccaggactactttgcttctgtgaatgcaccttca<br>aatcccccaagaccagcaattgttgtacttctcatgtccatggtgaacgaaaggcagccatttgtgttacgctgtgctgtcctctactgtttccagtgctttttatataaaaaccagaaaggacagggagaaatt<br>gtggcaacacttctaccgtccactattgatgcaacaggtaactcagtctcagctggtcagctgctctgtggaggcctcttttctacagattctctttcaaactggtgtgctgctgtggcccttgcccatgcactgca<br>agggaatgctacccagaaggagcagctgctcagggttcagctggccacgagcattggtaacccccagtgtccctgctgcagcaatgcaccaacatcctctcccagggtgataagatcgacagacgggga<br>agcaaaatacagacaagagttgggttgttaatgttgctttgtacctggctaagcagctgtcccattgcagtaacacattttcttcacaattcagccaatgttccatttcttacaggacaaattgcagaaaatctcg<br>gagaagaagagcagttagtccaaggcctatgtgcccttcttttgggcatttcaatttatttcaatgataattcactagaaaactacatgaaagagaaactaaagcaactaatagagaagagaattggcaagg<br>agaatttcatagagaaactaggattcattagcaaacatgagttatactccagagcctcacagaaaccccagccaaacttcccgagtccagaatacatgatatttgatcatgagtttacaaaactggtgaaaga<br>acttgaaggtgttattactaaggctatttataaatccagtgaagaagataagaaagaggaagaggtaaagaagacactagaacagcatgacaacattgtgactcactataagaatatgattcgtgagcaag<br>acctacagctggaagaactgaaacagcaagtttctacactgaaatgtcaaaatgaacagctgcaaacagcagtcacacaacaggtgtctcagattcaacagcacaaggatcagtataatctccttaaagtc<br>cagcttggaaaagacagtcaccatcaaggctcccacagtgaaggggcccaggtgaacggcattcagctagaagaggtcagtcggctgagggaagaggtagaagagctgaaaagtcaccaggggctctta<br>caaagccagctagccgaaaaggactctttgatagaaaatttgaaatcttcacaagcgtctgccctgaatgaacaggcttcagtaacagcttcacccagagattcagagcaagttgttgaattaaaacaggag<br>ctgacagcattaaagtcccagttgaactcacaggccctggagatcaccagactcaaggcagaaaatcatgagctgctccagagagcagaagccttggcaaagtcagtccctgtacaaggagagagtgaga<br>atgtgacaactgctaaaactactgatgtagaaggaaggctgtctgccttgcttcaggaaacgaaggaactaaagaatgaaattaaagcgttatctgaggaaaaaactgccattaaaaaacaattggataca<br>tctaacagcaccattgccattctacaaacggagaaagacaagttagatttggaagtgacagattctaagaaagagcaagatgatctcttggtactgttggccgatcaggatcagaaaatcctgtcactgaag<br>agtaaactcaagaacctcggtcatccagttgaagaagaagacgaatctggagaccaagaagatgatgatgatgaaatggatgatggtgacaaggaccaggatatctag |
| **Bet3**<br>CHO<br>XM_003500861.3<br>543 bp | atgtcgaggcaggcgaaccgtggcaccgagagcaagaaaatgagttctgagctatttacgctgacgtatggagcgcttgtcacccagctgtgcaaggactatgaaaacgatgaagacgtgaataagcagc<br>tggacagaatgggctataacattggagtccgactaattgaagatttttggcacgttcaaatgttggaagatgtcatgattttcgggaaactgcggatgtcattgctaaggtggcattcaagatgtacttaggca<br>tcactccaagcatcacgaattggagcccagctggcgatgaattctccctcatcttggaaaataacccgttggtggactttgtggagcttcccgatagccactcatcccttatttactccaacctcctgtgtggggt<br>gctgcggggagccttggagatggtccagatggctgtggaagccaagtttgtccaggacactctgaagggagatggtgtgacagagatcaggatgaggttcctcaggcggattgaagacaatcttccagctg<br>gagaagaataa |

| Gene name;<br>Human/CHO;<br>NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **TFE3**<br>**S108A mutation (TCA -> GCT)**<br>CHO<br>XM_007645617.2<br>1743 bp | atgtctcatgcagccgagccagctcgggacggcgtagaggccagcgtggagggccctcgagccgtgttcgtgctgttggaagagcgcaggccggccgactcggcccagctgctcagcctgaattctttgctt<br>ccggaatccgggattgttgctgacatcgaattagaaaacatccttgatcctgacagcttctacgagctcaaaagccaacctctacccctccgctccagcctcccaatatctctgcaggccacaccaaccacccc<br>agctacactctctgcatcgtcttctgcaggggctccaggacccctgccatgtca**GCT**tcatcttcatcgcgggtcttgctgcgtcagcagctgatgcgggcccaggcacaggagcaggagaggcgtgagc<br>gacgggaacaggcagcagctgctcccttccccaatcctgcacctgcctcaccagccatctctgtgattggcgtctctgctggtggccacacactgggtcgtccacccccctgctcaggtgcccagggaggtgct<br>caaggttcagacacacctggagaaccccacacgctaccacctgcagcaagctcgccggcagcaggtgaaacagtacttgtctaccacacttgggcccaagctggcttcccaggccctcacccccccaccgg<br>ggcctgccagtgcccagccacttcctgcccctgaaactgcccatgccactggccctacaggcagtgctcctaacagccccatggcgctgctcaccattggatccagctcagagaaggagtatttacttttttccc<br>cagattgatgatgtcattgatgagatcatcagcctagagtccagttacaacgatgagatgctcagctatcttcctggaggcactgcagggctgcagctccccagcacgctgcctgtgtctggcaacctgcttga<br>tgtgtacagcaaccaaggagtggccacaccagccatcactgtcagcaattcctgtcctgctgagctgcctaacatcaaacgggagatttctgaaaccgaggcaaaggccctttgaaggagcgacagaaga<br>aagacaatcacaacctaattgaacgacgcaggcgattcaacattaacgataggatcaaagagctgggcaccctcatccccaagtccaatgatccggagatgcgctggaacaagggcaccatcctgaaggc<br>atctgtggattacatccgcaaattacaaaaggaacagcaacgctccaaagacctggagagccggcagcgatccctggagcaagccaaccgaagtctgcagctccgaattcaggagctagaactgcaggcc<br>cagatccatggtctgccagtacctcctaccccaggactgctctccctagccactagttccgtctctgacagcctcaagccagagcagctggacattgaggaggagggcaggccaagcacgacatcatttcatg<br>tatcaggggacctatccagaatgcacctcagcagcagcctccagcaccaccctcggatgcccttctggacctgcactttcccagcgaccacttgggggacctgggggacccccttccacctggggctagagg<br>acattctgatggaggaggaggggtggtgggaggactgtcaggggtaccctgtcccctctgcgggctgcctctgaccccctttctttcttcagtatccccggctgtgtccaaggccagcagccgtcgcagcag<br>cttcagcatggaggaggagtcctga |
| **ATF6αc**<br>Human<br>1134 bp | atgggagaacctgctggcgtggccggcaccatggaaagccctttagccctggcctgttccaccggctggacgaggattgggactctgccctgtttgccgagctgggctacttcaccgacaccgacgaactcc<br>agctggaagccgccaacgagacatacgagaacaacttcgacaacctggacttcgacctggacctgatgccctgggagtccgacatctgggacatcaacaaccagatctgcaccgtgaaggacatcaaggc<br>cgagccccagcctctgtcccctgcctccagctcttactccgtgtcctcccccagatccgtggactcctactcctccacccagcacgtgcccgaggaactggacctgtcctccagcagccagatgtccccccctgtc<br>tctgtacggcgagaactccaactccctgtccagcgccgagcctctgaaagaggacaagcctgtgaccggccctcggaacaagaccgagaatggcctgacccccaagaaaaagatccaagtgaactccaag<br>ccctccatccagcccaagcccctgctgctgcctgctgctcctaagacccagaccaactcctccgtgcccgccaagaccatcatcatccagaccgtgcccaccctgatgcccctggctaagcagcagcccatca<br>tctccctccagcctgcccctaccaagggccagacagtgctgctgtctcagcccaccgtggtgcagctccaggctcctggcgtgctgccatctgctcagcctgtgctggctgtggctggcggagtgacccagctg<br>cctaatcacgtcgtgaacgtggtgcctgccccctccgccaactctcctgtgaacggcaagctgtccgtgaccaagccagtgctccagtccaccatgcggaacgtgggctccgatatcgccgtgctgagaagg<br>cagcagagaatgatcaagaacagagagtccgcctgccagtcccggaaaaagaaaaaagagtacatgctgggcctggaagcccggctgaaggccgccctgtctgagaacgagcagctgaagaaagaga<br>acggcaccctgaagcggcagctggatgaggtggtgtccgagaaccagagactgaaggtgcccagccccaagcggagatga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **BiP (Grp78)**<br>Human<br>1965 bp | atgaagctgtctctggtggccgccatgctcctgctgctgtctgccgctagagccgaggaagaggacaagaaagaggacgtgggcaccgtcgtgggcatcgacctgggcaccacctactcttgcgtgggcgt<br>gttcaagaacggccgggtggaaatcattgccaacgaccagggcaaccggatcacccccctcttacgtggcctttacccccgagggcgagagactgatcggcgacgccgctaagaaccagctgacctccaac<br>cccgagaacaccgtgttcgacgccaagcggctgatcggcagaacctggaacgacccctccgtgcagcaggatatcaagttcctgcccttcaaggtggtggaaaaaaagaccaagccctacatccaggtgg<br>acatcggcggaggccagaccaagaccttcgcccccgaagagatctccgccatggtgctgaccaagatgaaggaaaccgccgaggcctacctgggcaagaaagtgacccacgccgtcgtgaccgtgcccg<br>cctacttcaacgacgctcagagacaggccaccaaggacgccggcacaatcgccggcctgaacgtgatgcggatcatcaacgagcccaccgccgctgctatcgcctacggcctggataagcgggaaggcga<br>gaagaacatcctggtgttcgatctgggcggaggcaccttcgacgtgtccctgctgaccatcgacaacggcgtgttcgaggtggtggctaccaacggcgatacccacctgggcggcgaggacttcgaccaga<br>gagtgatggaacacttcatcaagctgtacaagaagaaaaccggcaaggacgtgcggaaggacaaccgggccgtgcagaaactgcggagagaggtggaaaaggccaagcgggccctgtcctctcagcac<br>caggccagaatcgagatcgagtccttctatgagggcgaggatttctccgagacactgacccgggccaagttcgaggaactgaacatggacctgttccggtccacaatgaagcctgtgcagaaggtgctgga<br>agattccgacctgaagaagtccgacatcgacgagatcgtgctcgtgggcggctccaccagaatccccaagatccagcagctcgtgaaagagttcttcaacggcaaagagccctcccggggcatcaaccctg<br>atgaggctgtggcttacggcgctgctgtgcaggctggcgtgctgtctggcgatcaggataccggcgatctggtgctgctggacgtgtgtcctctgaccctgggaatcgagacagtgggcggcgtgatgacca<br>agctgatccccagaaacaccgtggtgcctacaaagaagtcccagatcttctccaccgcctccgacaaccagcctaccgtgaccatcaaggtgtacgagggcgaacggcccctgaccaaggataaccatctg<br>ctgggcacctttgacctgaccggcatccctcctgctcctagaggcgtgccacagatcgaagtgacctttgagatcgacgtgaacggcatcctgagagtgaccgctgaggacaagggcaccggcaacaagaa<br>caagatcaccatcaccaacgatcagaacaggctgacacctgaagagatcgagcggatggtcaacgatgccgagaagttcgccgaagaagataagaagctgaaagagcggatcgacacccggaacgagc<br>tggaatcctacgcctactccctgaagaaccagatcggagacaaagagaagctgggcggcaagctgtccagcgaggacaaagaaaccatggaaaaagccgtggaagagaagatcgagtggctggaatct<br>caccaggacgctgacatcgaggactttaaggctaagaagaaagaactggaagaaatcgtgcagcccatcatctctaagctgtacggctctgccggccctccacctaccggcgaggaagataccgccgaga<br>aggacgagctgtga |
| **Calr**<br>Human<br>1254 bp | atgctgctgtctgttcctctgctgctgggactgcttggactggctgttgctgagcccgccgtgtacttcaaagagcagtttttggacggcgacggctggacctccagatggatcgagtctaagcacaagtccga<br>cttcggcaagttcgtgctgtcctctggcaagttctacggcgacgaggaaaaggacaagggcctccagacctctcaggacgccagattctacgccctgtctgcctctttcgagcccttcagcaacaagggccag<br>acactggtggtgcagttcaccgtgaagcacgagcagaacatcgactgtggcggcggatacgtgaagctgttccctaactctctggaccagaccgacatgcacggcgactccgagtacaacatcatgttcggc<br>cctgacatctgcggccctggcaccaagaaagtgcacgtgatcttcaactacaagggcaagaacgtgctgatcaacaaggacatccggtgcaaggacgacgagttcacccacctgtataccctgatcgtgcg<br>gcccgacaacacctacgaagtgaagatcgacaactcccaggtggaatccggcagcctggaagatgactgggacttcctgcctcctaagaagatcaaggaccccgacgcctctaagcctgaggattgggat<br>gagagagccaagatcgacgaccccaccgactccaaaccagaggactgggacaagcccgagcacatccctgatcctgatgccaaaaagcccgaagattgggacgaagagatggatggcgagtgggagcc<br>tcctgtgatccagaatcctgagtacaaaggcgaatggaagccccggcagatcgataaccccgactacaaaggcacctggattcaccccgagattgacaaccccgagtactctcccgatccttccatctacgc<br>ctacgacaactttggcgtgctgggcctcgatctgtggcaagtgaagtccggaaccatcttcgataactttctgatcaccaacgacgaggcctacgccgaagagttcggcaatgagacatggggcgtgaccaa<br>ggccgctgagaagcagatgaaggacaagcaggatgaggaacagcggctgaaagaggaagaagaggacaagaagcgcaaagaagaggaagaggccgaggacaaagaggacgacgaagataagga<br>cgaggatgaagaagatgaagaggacaaagaagaggacgaggaagaggacgtcccaggccaggctaaggatgagctgtaa |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **CANX** Human 1779 bp | atggaaggcaagtggctgctgtgtatgctgctggtgctgggcacagctatcgtggaagctcacgacggccacgacgacgacgtgatcgacatcgaggacgacctggatgatgtgatcgaggaagtggaag atagcaagcccgacaccaccgctcctccatctagccccaaagtgacctacaaggctcctgtgcctaccggcgaggtgtacttcgccgattctttcgacagaggcaccctgtccggctggattctgtctaaggcc aagaaggacgacacagacgacgagatcgctaagtacgacggcaagtgggaagtcgaggaaatgaaggaatccaagctgcccggcgacaaaggcctggtgctgatgtctagagccaagcaccacgccat ctccgccaagctgaacaagcccttcctgttcgacaccaagcctctgatcgtgcagtacgaagtgaacttccagaacggcatcgagtgtggcggcgcttacgtgaagctgctgtctaagacccctgagctgaac ctggaccagttccacgacaagacccttacaccatcatgttcggccctgacaagtgcggcgaggactacaagctgcacttcatcttccggcacaagaaccccaagaccggcatctacgaggaaaagcacgc caagaggcctgacgccgacctgaaaacctacttcaccgacaagaaaacccacctgtacaccctgatcctgaatcctgacaactccttcgagatcctggtggaccagtccgtggtcaactctggcaacctgct gaacgatatgacccctcctgtgaacccctccagagagatcgaggaccctgaggacagaaagcctgaggactgggacgagcggcctaagattcctgatccagaggccgtgaagcccgacgactgggatga agatgctcccgctaagatccccgacgaggaagctacaaagcctgaaggctggctggacgacgagcctgagtatgtgccagatcctgatgccgagaagccagaagattgggacgaagatatggatggcga gtgggaagcccctcagatcgccaatcctagatgcgagtctgctcctggatgtggcgtgtggcagaggcctgtgatcgataaccccaactacaaaggcaaatggaagcctcctatgatcgacaacccttcctа ccaaggcatctggaagccccggaagatccctaatcctgacttcttcgaggacctggaacctttccggatgacccctttttccgccatcggcctggaactgtggtctatgacctccgacatcttcttcgacaacttc atcatctgcgccgaccggcggatcgtggatgattgggctaatgatggctggggcctgaagaaagctgctgatggtgcagctgaaccaggcgtcgtgggacagatgattgaggccgccgaagaaagacctt ggctgtgggtcgtgtacatcctgacagtggccctgcctgtgttcctggtcatcctgttctgctgctccggcaagaaacagacctccggcatggagtacaaaaagaccgacgctccccagcctgacgtgaaaga ggaagaagaggaaaaagaagaagagaaggacaagggcgacgaagaggaagagggcgaagagaagctggaagaaaagcagaagtccgacgccgaagaggatggcggcacagtgtcccaagaaga agaagatcggaagcccaaggccgaggaagatgagatcctgaacagatcccctcggaaccggaagcctcggagagaatga |
| **CypB** Human 651 bp | atgctgcggctgtctgagcggaacatgaaggtgctgctggccgctgctctgatcgccggctctgtgttcttcctgctgctgcctggcccttctgccgccgacgagaagaaaaagggccccaaagtgaccgtga aggtgtacttcgacctgcggatcggcgacgaggacgtgggcagagtgatcttcggcctgttcggcaagaccgtgcctaagaccgtggacaacttcgtggccctggccaccggcgagaagggcttcggctac aagaactccaagttccatcgagtgatcaaggacttcatgatccagggcggcgacttcaccagaggcgacggaacaggcggcaagtccatctacggcgagcggttccctgacgagaacttcaagctgaagc actacggccctggctgggtgtccatggccaatgccggcaaggacaccaacggctcccagttcttcatcaccaccgtgaaaaccgcctggctggacggcaagcacgtggtgtttggcaaggtgctggaaggc atggaagtcgtgcggaaggtggaatccaccaagaccgactcccgggacaagcccctgaaggacgtgatcattgccgactgcggcaagatcgaggtggaaaagcccttcgctatcgccaaagagtga |

| Gene name;<br>Human/CHO;<br>NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **Largen**<br>Human<br>915 bp | atgtccgccaagtccaagggcaacccctccagctcttgtcctgccgagggacctcctgccgcctccaagaccaaagtgaaagagcagatcaagatcatcgtggaagatctggaactggtgctgggcgacct<br>gaaggacgtggccaaagaactgaaagaagtggtggaccagatcgacaccctgacctccgacctccagctggaagatgagatgaccgactccagcaagaccgataccctgaactcctcctcctctggcacc<br>accgccagctccctggaaaagatcaaggtgcaggccaacgccccctgatcaagcctcctgctcacccttctgccatcctgaccgtgctgcggaagcccaaccctcctccaccacctcctcgactgacccccg<br>tgaaatgcgaggaccctaagcgggtggtgcccaccgccaaccctgtgaaaacaaacggcaccctgctgagaaacggcggcctgcctggcggccctaacaagatccctaacggcgacatctgctgcatccc<br>caactccaacctggacaaggccccgtgcagctgctgatgcaccggcctgagaaggacagatgtccccaggctggccctcgggaaagagtgcggttcaacgagaaggtgcagtaccacggctactgcccc<br>gactgcgacacccggtacaacatcaagaaccgggaagtgcatctgcactccgagcccgtgcatccccctggcaagatcccacatcagggccctccactgccccccacacctcatctgcctcccttccctctgg<br>aaaatggcggcatgggcatctcccacagcaacagcttcccccctatccggcctgctaccgtgcctccacctaccgctcccaagccccagaaaaccatcctgcggaagtctaccaccaccacagtgtga |
| **ERGIC53**<br>**(LMAN1)**<br>CHO<br>1554 bp | atggctggctctagaagaaagggaacacaggctggcgctcggccttttatctgtgccctgctgctgtccttcagcagattcgtgggctctgatggcatgggcggagatgctgctgctcctggtgctgcttctact<br>cctgctgagctgccccacagaagattcgagtacaagtactccttcaagggccctcacctggtgcagtctgatggaaccgtgccttttttgggcccacgccggcaatgctatcccttccagcgatcagatcagaat<br>cgcccctagcctgaagtctcagaggggctctgtgtggaccaagaccaaggccgccttcgagaactgggaagtcgaagtgaccttcagagtgaccggcagaggcagaatcggagctgatggcctggctatct<br>ggtacaccgagaatcagggactcgatggccctgtgttcggctctgctgatatgtggaacggcgtgggaatcttcttcgactccttcgacaacgacggcaagaagaacaaccccgccatcgtgatcatcggca<br>acaacggccagatcaactacgaccaccagaacgatggcacaacccaggctctggcctcttgccagagggacttccggaacaaaccttatcctgtgcgggccaagatcacctactaccagaaaaccctgacc<br>gtgatgatcaacaatggcttcacccctgacaagaacgactacgagttctgcgccaaggtcgagaacatgatcatccctactcagggccacttcggcatctctgctgctactggcggactggctgacgaccacg<br>atgtgctgtcttttctgaccttccagctgaccgagcctggcaaagagcctcctacactggaaaaggacatctccgagaaagagaaagaaaagtaccaagaagagttcgagcacttccagcaagagctggac<br>aagaagaaagaagaatttcagaagggacaccccgacctccagggccagcctgctgatgatgttttcgagtccgtgggcgaccgagagctgagacaggtgttcgagggccagaacagaatccacctggaa<br>atcaagcagctgaaccggcagctggacatgatcctggacgagcagagaagatacgtgtcctctctgaccgaggaaatctctcggagaggcgctggaactcctggacagcctggacaagtgtcccagcaag<br>aactcgacaccgtggtcaagacccagcacgagatcctgcggcaagtgaacgagatgaagaactccatgtccgagacagtgcggctggtgtctggtatccagcatcctggatctgctggcgtgtacgagaca<br>acccagcacttcatggacatcaaagaacacctccacgtcgtgaagcgggacatcgacaatctggcccagcggaacatgccctccaacgagaagcctaagtgtcccgagctgcctccatttcctagctgcctg<br>tctaccatccacttcgtgatcttcgtggtggtgcagaccgtgctgttcatcggctacatcatgtaccggacacagcaagaggctgccgccaagaagttcttctag |

| Gene name;<br>Human/CHO;<br>NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **ERo1Lα**<br>Human<br>1407 bp | atgggaagaggctggggcttcctgtttggcctgctgggagctgtgtggctgctgtcctctggccatggcgaggaacagcctccagagacagccgcccagagatgcttctgtcaggtgtccggctacctggac<br>gactgcacctgtgacgtggaaaccatcgacaggttcaacaactaccggctgttcccccggctccagaagctgctggaatccgactacttccggtactacaaagtgaacctgaagcggccctgccccttctgga<br>acgacatctcccagtgcggcagacgggactgcgccgtgaagccttgccagtctgatgaggtgcccgacggaatcaagtccgcctcctacaagtactccgaggaagccaacaacctgatcgaggaatgcga<br>gcaggccgagagactgggcgccgtggatgagtccctgtccgaagagacacagaaagccgtgctccagtggaccaagcacgacgactcctccgacaacttctgcgaggccgacgacatccagagccccga<br>ggccgagtatgtggacctgctgctgaaccccgagcggtacaccggctacaagggccctgatgcctggaagatctggaacgtgatctacgaggaaaactgcttcaagccccagaccatcaagcggcccctga<br>accctctggcttctggccagggcaccagcgaagagaacaccttctacagctggctggaaggcctgtgtgtggaaaagcgggccttttaccggctgatctccggcctgcacgcctccatcaacgtgcacctgtc<br>cgccagatacctgctccaggaaacatggctggaaaagaagtggggccacaacatcaccgagttccagcagagattcgacggcatcctgaccgagggcgagggccctagaaggctgaagaacctgtacttc<br>ctgtatctgatcgagctgcgggccctgtccaaggtgctgccattcttcgagcggcccgacttccagctgttcaccggcaacaagatccaggacgaagaaaacaagatgctgctgctggaaatcctgcacgag<br>atcaagtccttccccctgcacttcgacgagaacagcttctttgccggcgacaagaaagaggcccacaagctgaaagaggacttccggctgcacttccggaacatctcccggatcatggactgcgtgggctgtt<br>tcaagtgccggctgtggggcaagctccagacccagggactgggcaccgctctgaagatcctgttctccgagaagctgatcgccaacatgcccgagtccggcccctcctacgagttccacctgacccggcagg<br>aaatcgtgtccctgttcaacgccttcggccggatctccacctccgtgaaagagctggaaaactttcggaacctgctccagaacatccactga |
| **FKBP11**<br>Human<br>606 bp | atgacactcagacctagcctgctgcctctgcacctgttgctgctgttgctcctctctgccgctgtgtgtagagctgaggctggcctggaaacagagtctcctgtgcgaaccctccaggtggaaactctggtggaa<br>cctcctgagccttgtgccgaacctgctgcttttggcgataccctgcacatccactataccggcagcctggtggacggccggattatcgataccagcctgaccagagatcccctggtcatcgagctgggccaga<br>aacaagtgatccccggcctggaacagtccctgctggatatgtgtgtgggcgagaagcggagagctatcatccccttctcacctggcctacggcaagagaggctttcctccttctgtgcctgctgacgccgtggt<br>gcagtacgatgtggaactgatcgccctgatccgggccaactattggctgaagctggtcaagggaatcctgcctctcgtcggcatggctatggtgcctgctctgctgggcctgatcggctaccacctgtacaga<br>aaggccaaccggcctaaggtgtccaagaagaagctgaaagaggaaaagcgcaacaagtctaagaagaagtga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **GRP94/ HSP90B1 (Endoplasmin)** Human 2412 bp | atgagagctttgtgggtgctgggcctgtgttgcgtgctgctgacctttggctctgtgcgggccgatgacgaagtggatgtggatggcacagtggaagaggacctgggcaagtccagagagggctctagaaca gacgacgaggtggtgcagagagaggaagaggctattcagctggacggcctgaacgcctctcagatcagagagctgagagagaagtccgagaagttcgccttccaagccgaagtgaaccggatgatgaag ctgatcatcaactccctgtacaagaacaaagagatcttcctgcgcgagctgatctccaacgcctctgacgccctggacaagatccggctgatctctctgaccgacgagaacgccctgtctggcaatgaggaa ctgaccgtgaagatcaagtgcgacaaagagaagaacctgctgcacgtgaccgataccggcgtgggcatgacaagagaggaactggtcaagaacctgggcacaatcgccaagtccggcacctccgagttt ctgaacaagatgaccgaggctcaagaggacggccagtccacatctgagctgatcggacagtttggcgtgggcttctactctgccttcctggtggctgacaaagtgatcgtgacctccaagcacaacaacgac acccagcacatctgggagtccgactccaacgagttctccgtgatcgccgatcctagaggcaataccctcggcagaggcaccacaatcaccctggtgctgaaagaagaggcctccgactacctggaactgga caccatcaagaatctcgtgaagaagtactcccagttcatcaacttccccatctacgtgtggtcctccaagaccgagactgtggaagaacccatggaagaagaggaagccgccaaagaggaaaaagaagag tccgacgacgaagccgccgtcgaggaagaggaagaagagaagaagcctaagaccaagaaggtggaaaagaccgtctgggactgggagctgatgaacgacatcaagcctatctggcagcggccctcca aagaggtcgaagaggacgagtacaaggccttctacaagtccttcagcaaagaaagcgacgaccccatggcctacatccatttcacagctgagggcgaagtgaccttcaagtccatcctgttcgtgcctacct ctgctcccagaggcctgttcgatgagtacggctccaagaagtctgactacatcaagctgtacgtgcggcgggtgttcatcaccgacgacttccacgacatgatgcccaagtacctgaacttcgtgaaaggcgt ggtggactccgacgacctgcctctgaatgtgtctagagagacactccagcagcacaagctgctgaaagtgatccggaagaagctcgtcagaaagaccctcgacatgatcaagaagatcgccgacgacaag tacaacgataccttctggaaagagttcggcaccaatatcaagctgggcgtgatcgaggaccactccaacagaaccagactggctaagctgctgcggttccagtcctctcaccatcctaccgacatcaccagc ctggaccagtacgtggaacggatgaaggaaaagcaggataagatctacttcatggccggctccagccggaaagaggccgagagttctccattcgtggaaagactgctgaagaagggctacgaagtcatct acctgaccgagccagtggatgagtactgtatccaggctctgcccgagttcgacggcaagagattccagaacgtggccaaagaaggcgtgaagttcgacgagtccgaaaagaccaaagagagcagagagg ccgttgagaaagagtttgagcccctgctgaattggatgaaggacaaggccctgaaggacaaaatcgagaaggccgtggtgtcccagagactgaccgaatctccttgtgctctggtggcctctcagtacggat ggtccggcaacatggaaagaatcatgaaggcccaggcctaccagaccggcaaggacatctctaccaactactacgccagccagaagaaaaccttcgagatcaaccctagacaccctctgatccgggacat gctgcggcggatcaaagaggatgaggacgacaagaccgtgctggacctggctgtggtgctgtttgagacagctaccctgagaagcggctacctgctgcctgataccaaggcctacggcgacagaatcgag cggatgctgagactgtccctgaacatcgaccctgacgccaaggttgaagaggaacccgaggaagaaccagaggaaaccgctgaggataccaccgaggataccgagcaggacgaggacgaagaaatgg acgtgggcaccgacgaagaagaagaaacagctaaagagtccaccgccgagaaggatgagctgtga |
| **MCFD2** CHO 438 bp | atggcttctctgtggctgctgagagtgcctttcctgtgtgctctgctgtggacctcttgcacccctggaacaagagcacatgagcctggcgctggctctcatcattcctctgtcggcctggacaagtccaccgtgc acgaccaagagcatatcatggaacacctggaaggcgtgatcgacaagcccgagacagagatgtcccctcaagaactccagctgcactacttcaagatgcacgactacgacggcaacgacctgctggatgg cctggaactgtctaccgctatcacccacgtgcacaaagaggaaggctctgatcaggcccctgtgatgtctgaggacgagctgatcaacatcatcgacggcgtgctgcgggacgacgacaagaacaacgac ggctacatcgactacgccgagttcgccaagagcctccagtaa |

| Gene name;<br>Human/CHO;<br>NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **Hyou1**<br>Human<br>3000 bp | atggctgacaaagttcggagacagaggcctcggagaagagtgtgttgggctcttgtggctgtgctgctggctgatctgctggccttgtctgataccctggccgtgatgtctgtggacctgggctctgaatccat<br>gaaggtggccattgtgaagcccggcgtgcccatggaaatcgtgctgaacaaagagtctcggcgcaagacccctgtgatcgtgaccctgaaagagaacgagcggttcttcggcgactctgccgcctctatgg<br>ccatcaagaatcctaaggctaccctgcggtacttccagcatctgctgggcaagcaggccgataatcctcacgtggccctgtaccaggccagatttcctgagcacgagctgacctttgatccccagcggcagac<br>cgtgcacttccagatctctagccagctccagttcagccctgaagaggtgctcggcatggtgctgaactactctagatccctggccgaggatttcgccgagcagcctatcaaggatgccgtgatcaccgtgcctg<br>tgttcttcaatcaagccgagagaagggccgtgctccaggctgctagaatggctggactgaaggtgctccagctgatcaacgataacaccgccaccgctctgtcttacggcgtgttcagacggaaggacatca<br>acaccaccgctcagaacatcatgttttacgacatgggctccggctccaccgtgtgcaccatcgtgacataccagatggtcaagaccaaagaggccggcatgcagccccagttgcagatcagaggcgtgggc<br>tttgatagaaccctcggcggcctggaaatggaactgagactgagagaaagactggccggcctgttcaacgagcagagaaaaggccagagagccaaggacgtccgcgagaatcctagagctatggccaag<br>ctgctgagagaggccaacagactgaaaaccgtgctgtctgccaacgccgaccacatggctcagatcgagggactgatggacgacgtggacttcaaggccaaagtgaccagagtggaatttgaggaactgt<br>gcgccgacctgttcgagagagttcctggacctgttcagcaggccctccagtctgctgagatgtccctggatgagatcgagcaagtgatcctcgtcggcggagctaccagagtgcctagagtgcaagaagtgc<br>tgctgaaggccgtgggcaaagaggaactgggcaagaacatcaacgccgatgaggctgctgctatgggcgctgtttatcaggccgctgctctgtccaaggccttcaaagtgaagcctttcgtcgtgcgggacg<br>ccgtggtgtatcctatcctggtcgagttcaccagagaggtggaagaggaacctggcatccacagcctgaagcacaacaagcgggtgctgttctcccggatgggcccttatcctcagcggaaagtgatcacct<br>tcaaccgctactcccacgacttcaacttccacatcaactacggcgatctgggctttctgggccctgaggatctgagagtgttcggctctcagaacctgaccaccgtgaagctgaaaggcgtgggagactccttc<br>aagaagtaccccgactacgagtccaagggcatcaaggcccacttcaacctggacgagtctggcgtgctgagcctggatagagtggaaagcgtgttcgagacactggtggaagattccgccgaggaagaga<br>gcaccctgaccaagctgggcaacaccatcagttctctgttcggcggaggcaccacacctgacgctaaagagaatggcaccgacaccgtgcaagaggaagaagagtctcctgccgagggctctaaggatga<br>gcctggcgaacaggtggaactgaaagaagaggctgaggcccctgtggaagatggctctcaacctcctccacctgaacctaagggcgacgctacacctgagggcgagaaggccacagagaaagaaaacg<br>gcgacaagagcgaggcccagaagccatctgaaaaggccgaagctggaccagaaggtgtcgctcctgctccagaaggcgagaagaagcagaagcctgccagaaagcggcggatggtcgaagagattgg<br>agtggaactggtggtgctggacctgcctgatctgccagaagataagctggcccagtccgtgcagaagctccaggacctgactctgcgcgacctggaaaagcaagagagagagaaggccgccaacagcct<br>ggaagccttcatcttcgagactcaggacaagctgtaccagcctgagtaccaagaggtgtccaccgaggaacagagagaggaaatctctggcaagctgtccgctgccagcacctggcttgaggatgaaggc<br>gttggcgctacaacagtgatgctgaaagaaaagctggctgagctgagaaagctgtgtcagggcctgttcttcagagtcgaggaacggaagaagtggcccgagagactgtctgccctggacaacctgctgaa<br>tcactcctccatgtttctgaagggcgctagactgatccccgagatggaccagatcttcaccgaggtggaaatgaccacactcgagaaagtcatcaacgagacatgggcctggaagaacgctacactggccg<br>aacaggctaagctgcctgctacagaaaagcccgtgctgctgtctaaggacatcgaggccaagatgatggccctggaccgggaagtgcagtacctgctcaacaaggccaagttcaccaagcctcggcctag<br>acctaaggacaagaacggaaccagagccgagcctcctctgaacgcctctgcttctgatcagggcgaaaaagtgatccctcctgccggccagacagaggatgctgaacctatctctgagcccgagaaggtg<br>gaaaccggctctgaacctggcgataccgaacctttggaactcggaggacctggcgctgagccagagcagaaagaacagtctaccggacagaagcggccccctgaagaacgacgagctgtaa |

| Gene name;<br>Human/CHO;<br>NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **Dad1**<br>CHO<br>342 bp | atgtctgcatccgtggtgtccgtgatctcccggttcctggaagagtacctgagcagcacccctcagagactgaagctgctggatgcctacctgctgtacattctgctgaccggcgctctccagttcggctattgt<br>ctgctcgtgggcacctttcctttcaattccttcctgtccggctttatctcctgcgtgggctctttcatcctggccgtgtgcctgagaatccagatcaaccctcagaacaaggccgacttccagggcatctctcctga<br>gagagccttcgccgatttcctgttcgcctctaccatcctgcacctggtggtcatgaacttcgtgggctga |
| **PDI**<br>Human<br>1527 bp | atgctgagaagggccctgctgtgtctggccgtggctgctctcgtgcgagccgatgctcctgaggaagaggaccacgtgctggtgctgcggaagtccaactttgccgaggctctggccgcccacaagtacctg<br>ctggtggagttctacgccccttggtgcggccactgtaaagccctggcccctgagtacgctaaggccgctggaaagctgaaggccgagggctctgagatccggctggctaaggtggacgccaccgaggaatc<br>tgacctggcccagcagtatggcgtgcggggctaccccaccatcaagttcttccggaacggcgacaccgcctcccccaaagagtataccgccggcagagaagccgacgacatcgtgaactggctgaagaag<br>agaaccggccctgccgccaccacactgcctgatggtgctgctgctgagtccctggtggaatcctctgaggtggccgtgatcggcttcttcaaggacgtggaatccgactccgccaagcagtttctccaggccg<br>ccgaggctatcgacgatatccccttcggcatcacctccaactccgacgtgttctccaagtatcagctggacaaggacggcgtggtgctgttcaagaagttcgacgagggccggaacaacttcgagggcgaag<br>tgaccaaagagaacctgctggacttcatcaagcacaaccagctgcccctcgtgatcgagttcaccgagcagaccgccccaagatcttcggcggcgagatcaagacccacatcctgctgtttctgcccaagt<br>ccgtgtccgactacgacggcaagctgtccaacttcaagaccgccgctgagtccttcaagggcaagatcctgttcatcttcatcgactccgaccacaccgacaaccagcgcatcctggaatttttcggcctgaa<br>gaaagaggaatgccctgccgtgcggctgatcaccctggaagaggaaatgaccaagtacaagcccgagtctgaggaactgaccgccgagcggatcaccgagttctgccacagatttctggaaggcaagatc<br>aagcctcacctgatgtcccaggaactgcccgaggactgggacaagcagcccgtgaaagtgctcgtgggcaagaactttgaggacgtggccttcgacgagaagaaaaacgtgttcgtggagttttatgctcct<br>tggtgtgggcattgcaagcagctggcccccatctgggataagctgggcgagacatacaaggaccacgagaacatcgtgatcgccaagatggactccaccgccaacgaggtggaagccgtgaaggtgcac<br>agcttccccaccctgaagtttttccccgcctctgccgaccggaccgtgatcgattacaacggcgagagaaccctggacggctttaagaagttcctggaatctggcggccaggacggcgctggcgacgatgac<br>gatctggaggacctggaagaagctgaggaacccgacatggaagaagatgacgaccagaaagctgtgaaggacgagctgtga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **P115**<br>CHO<br>2901 bp | atgaactttctgagaggcgtgatgggcggccagtctgctggacctcaacataccgaggccgagacaatccagaaactgtgcgacagagtggccagctctaccctgctggacgatagaagaaatgccgtgcg<br>ggccctgaagtccctgtccaagaagtatagactggaagtcggcatccaggccatggaacacctgatccatgtgctccagaccgacagatccgactccgagatcattgcctacgctctggacaccctgtacaa<br>catcatctccaacgacgaggaagaggaaatcgaggaaaactctgcccggcagaccgaggatctgggctctcagttcaccgagatcttcatcaagcagcccgagaacgtgaccctgctgctgtctctgctgg<br>aagagttcgatttccatgtgcggtggcctggcgtcagactgctgacatctctgctgaaacagctgggccctcctgtgcagcagatcatcctggtgtctccaatgggcgtgtcccggctgatggatctgctggcc<br>gattctagagagatcatccggaacgacggcgtgctgctgttgcaggccttgaccagatctaacggcgccatccagaaaatcgtggccttcgagaacgccttcgagcggctgctggatatcatcaccgacgag<br>ggcaattctgacggcggcatcgtggtggaagattgcctgatcctgctccagaacctgctgaagaacaacaactccaaccagaacttcttcaaagagggctcctacatccagcggatgaagccttggtttgag<br>gtggccgacgagaactctggatggtccgcacagaaagtgaccaacctgcacctgatgctccagctcgtcagagtgctggtgtcccctacaaatcctcctggcgccaccagctcttgccagaaagctatgtttc<br>agtgcggactgctccagcagctgtgcaccattctgatggctaccggcattcctgccgacatcctgaccgagactatcaacaccgtgtccgaagtgatccggggctgccaagtgaaccaggactacttcgcctc<br>tgtgaacgcccctagcaaccctcctagacctgctatcgtggtgctgctgatgtccatggtcaacgagagacagcccttcgtgctgagatgcgccgtgctgtactgcttccagtgcttcctgtacaagaaccaga<br>aaggccagggcgagatcgtggccacactgctgccttctacaatcgacgccaccggcaactctgtttctgctggtcaactgctgtgtggcggcctgttctctaccgacagcctgtctaattggtgcgccgctgttg<br>ctctggctcacgctctccagggcaatgctacccagaaagaacagctgctcagagtgcagctggctacctctatcggcaacccacctgtgtctctgttgcagcagtgcacaaacatcctgtctcagggcgacaa<br>gatcgacagacggggctccaagatccagaccagagttggactgctgatgctgctgtgtacctggctgtctagctgccctatcgctgtgacccacttcctgcacaactctgccaacgtgccattcctgaccggcc<br>agatcgctgagaatctgggcgaagaggaacagctggtgcagggactgtgtgctctgctgctgggcatctccatctacttcaacgacaactccctcgagaactacatgaaggaaaagctgaagcagctgatc<br>gagaagcggatcggcaaagagaacttcattgagaagctgggcttcatcagcaagcacgagctgtactccagagccagccagaagcctcagcctaacttcccatctcctgagtacatgatcttcgatcacgag<br>ttcaccaagctggtcaaagaactggaaggcgtgatcaccaaggccatctacaagtcctccgaagaggacaagaaagaggaagaagtcaagaaaaccctcgagcagcacgacaacatcgtcacccactac<br>aagaacatgatcagagagcaggacctccagttggaggaactcaagcagcaagtgtctaccctgaagtgccagaacgagcagctccagacagccgtgacacaacaggtgtcccagattcagcagcataag<br>gaccagtacaatctgctgaaggtgcagctcggcaaggatagccaccaccagggatctcattctgagggcgcccaagtgaatggcatccagctcgaggaagtgtctcggctgagagaagaggtggaagaac<br>tgaagtctcaccagggcctgttgcagtctcagctggccgagaaggactccctgattgaaaacctgaagtctagccaggcctccgctctgaatgagcaggcttctgtgaccgcctctcctcgggattctgaaca<br>ggtggtggaactgaagcaagaactgacagccctgaaatcccagctgaactcccaggctctggaaatcaccagactgaaggccgagaaccacgaactgttgcagagagccgaagctctggccaagtctgtg<br>ccagtgcagggcgagtctgagaatgtgaccaccgctaagaccaccgacgtggaaggcagattgagcgccctgctgcaagagacaaaagagctgaaaaacgagatcaaggccctgtccgaggaaaagac<br>cgccatcaaaaagcagctggacacctccaactccacaatcgccatcctccagactgagaaggacaagctggacctggaagtgaccgactccaaaaaagaacaggacgacctgctggtgctcctggccgac<br>caggatcagaagattctgtctctgaagtccaagctcaagaacctgggccatccagtggaagaagaggacgagtctggcgaccaagaggatgacgacgacgagatggacgacggcgacaaggaccagga<br>catctag |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **PDIA4**<br>Human<br>1938 bp | atgaggcctagaaaggcttttctgctgctgctcctgctgggacttgtgcagttgttggctgttgctggcgctgagggccctgacgaggattcctctaacagagagaacgccatcgaggacgaggaagaagaa<br>gaggaagaggacgacgacgaagaagaggacgacctggaagtgaaagaagagaacggcgtcctggtgctgaacgacgccaacttcgataacttcgtggccgacaaggacaccgtgctgctcgagttttat<br>gctccttggtgcggccactgcaagcagtttgcccctgagtacgagaagatcgccaacatcctgaaggacaaggaccctcctattcctgtggccaagatcgacgctacctctgcctctgtgctggcctccagatt<br>cgatgtgtctggctaccccaccatcaagatcctgaagaaaggccaggccgtggactacgagggcagcagaacacaagaagagatcgtcgccaaagtgcgcgaggtgtcccaacctgattggacccctcca<br>cctgaagtgaccctggtcctgaccaaagagaacttcgacgaggtcgtgaacgatgccgacatcatcctggtcgagttctacgccccttggtgtggacattgcaagaagctggctcccgagtacgaaaaggcc<br>gccaaagagctgtccaagcggtcccctccaattccactggctaaggtggacgccaccgctgagacagatctggccaagagatttgacgtgtccggctatcctacactcaagatcttcagaaagggcagaccc<br>tacgactacaacggccctcgcgagaagtacggcatcgtggattacatgatcgagcagagcggccctcctagcaaagagatcctgacactgaaacaggtgcaagagtttctgaaggacggcgacgacgtga<br>tcatcatcggagtgttcaagggcgagagcgaccctgcttaccagcagtaccaggatgccgccaacaatctgagagaggactacaagttccaccacaccttctctaccgagatcgctaagttcctgaaagtgt<br>cccagggccagctggtggtcatgcagcctgagaagttccagtctaagtacgagcctcggagccacatgatggacgtgcagggctctacccaggactccgccatcaaggacttcgtgctgaagtatgccctgc<br>ctctcgtgggccacagaaaggtgtccaacgatgccaagaggtacacccgcagacctctggtggtggtgtactactccgtggacttcagcttcgactacagagccgccacacagttctggcggtctaaggtgct<br>ggaagtggccaaggactttcctgagtacaccttcgctatcgccgacgaagaggattacgccggcgaagtgaaggatctgggactgtctgagtctggcgaggatgtgaatgccgccattctggacgagtccg<br>gcaagaaattcgccatggaacccgaagagttcgacagcgacaccctgagagaatttgtgaccgccttcaagaagggcaagctgaagcccgtgatcaagagccagcctgtgcctaagaacaacaagggcc<br>ctgtgaaggtggtcgtgggcaagaccttcgactccatcgtgatggaccccaagaaagacgtgctgattgagttttacgcaccttggtgtgggcactgtaaacagctggaacccgtgtacaactccctcgccaa<br>gaagtacaagggccagaaaggcctcgtgatcgccaagatggatgccaccgccaacgatgtgccctccgacagatacaaggtggaaggcttccctaccatctacttcgccccttccggcgacaagaaaaac<br>cccgtgaagttcgaaggcggcgacagagatctggaacacctgtccaagttcatcgaggaacacgccaccaagctgtcccggacaaaagaggaactgtag |
| **PRDX4**<br>Human<br>816 bp | atggaagctctgcctctgttggccgctaccacacctgatcacggcagacatagaaggctgctgctgctccctctgctgctgtttttgttgcctgctggcgctgtgcaaggctgggagacagaggaaagacccc<br>ggaccagagaagaggaatgccacttttatgctggcggccaggtttaccctggcgaggcttctagagtgtctgtggccgatcactccctgcacctgtccaaggccaagatctccaagcctgctccttattggga<br>gggcaccgctgtgatcgacggcgagtttaaagagctgaagctgaccgactaccggggcaagtacctggtgttcttcttctaccctctggacttcaccttcgtgtgccccaccgagatcattgccttcggcgaca<br>gactggaagagttccggtctatcaataccgaggtggtggcctgctccgtggactctcagtttacccacctggcctggatcaacacccctagaaggcaaggcggactgggccctattagaatcccactgctgtc<br>cgatctgacccaccagatctctaaggactacggcgtgtacctggaagatagcggccatacactgcggggcctgttcatcatcgacgacaagggcatcctgcggcagatcaccctgaatgatctgcctgtggg<br>cagatccgtggacgagacactgagactggtgcaggccttccagtacaccgataagcacggcgaagtctgtcctgctggctggaagcctggctccgagacaatcattcctgatcctgccggcaagctcaagta<br>cttcgacaagctgaactga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **SRP14**<br>CHO<br>333 bp | atggttctgctggaatccgagcagtttctgaccgagctgacccggctgttccagaagtgcagatcctccggctctgtgtacatcaccctgaagaagtacgacggccggaccaagcctacacctagaaagtctg<br>ccgtggaatccgtggaacccgccgagaacaaatgtctgctgagagccaccgacggcaagcggaagatttctaccgtggtgtcctccaaagaagtgaacaagttccagatggcctactccaacctgctgagg<br>gccaacatggatggactgaagaagcgggacaagaagaacaagtccaaaaagaccaagccagctcagtga |
| **Tor1a**<br>Human<br>999 bp | atgaagctgggtagagctgttctgggactgctgctgctggctccttctgtggttcaggccgtggaacctatctctctcggactggctctggctggcgtgctgaccggctatatctaccctagactgtactgcctgt<br>tcgccgagtgctgcggccagaagagatccctgtctagagaggccctccagaaggacctggacgataacctgtttggccagcacctggccaagaagatcatcctgaacgccgtgttcggcttcatcaacaacc<br>ccaagcctaagaagcccctgacactgtctctgcatggatggaccggcaccggcaagaacttcgtgtccaagatcattgccgagaacatctacgaaggcggcctgaatagcgactacgtgcacctgtttgtgg<br>ctaccctgcacttccctcacgcctccaacatcaccctgtacaaggatcagctccagctgtggatcagaggcaacgtgtccgcctgcgccagatccatcttcatcttcgacgagatggacaagatgcacgccgg<br>cctgatcgacgccatcaagccctttctggactactacgacctggtggacggcgtgtcctaccagaaagccatgttcatctttctgtccaacgctggcgccgagagaatcaccgatgtggccctggatttttggc<br>ggagcggcaagcagagagaggacatcaagctgaaggacatcgagcacgccctgtccgtgtccgtgttcaacaacaagaactccggcttctggcactcctctctgatcgaccggaacctgatcgattacttcg<br>tgcccttcctgcctctcgagtacaagcacctgaagatgtgcatcagagtggaaatgcagtcccggggctacgagatcgacgaggacattgtgtccagagtggccgaggaaatgacattcttcccaaaagag<br>gaacgggttttctccgacaagggctgcaagaccgtgttcaccaagctggactattactacgacgactga |
| **XBP1s**<br>CHO<br>1113 bp | atggttgtggtggctgcttctccttctgctgccactgctgctcctaaggtgctgctgttgtctggccaacctgctgctgatggtagagccctgcctctgatggtgcctggatctagagctgccggatctgaggcta<br>atggcgcccctcaggctagaaagagacagcggctgacccacctgtctccagaggaaaaggccctgcggcggaagctgaagaatagagtggctgcccagaccgccagagacagaaagaaagcccggatg<br>tccgagctggaacagcaggtcgtggacctggaagaggaaaaccagaagctgctgctggaaaatcagctgctgagagaaaagacccacggcctggtcatcgagaatcaagagctgagaacccggctcggc<br>atggacgtgctgacaacagaggaagcccctgagacagagtccaaaggcaatggcgttagacctgtggccggctctgctgaaagtgctgctggtgctggacctgtggtcacctctcctgagcatctgcctatg<br>gactccgacaccgtggactcctctgactccgagtctgatatcctgctgggcatcctggacaagctggaccccgtgatgttcttcaagtgcccatctcctgagtccgccaatctggaagaactgcccgaagtgta<br>ccccggaccttcttctctgcctgcctcactgtctctgtccgtgggcacatcttctgccaagctggaagccatcaacgagctgatcagattcgaccacgtgtacaccaagcctctggtgctcgagatcccttccga<br>gactgagtctcagaccaacgtggtggtcaagatcgaagaggcccctctgtcctccagcgaggaagatcaccctgagttcatcgtgtccgtgaagaaagagcccctcgaagaggacttcatccccgagcctg<br>gcatctctaacctgctgtcctctagccactgcctgaagcctagctcttgcctgctggacgcctactctgattgtggctacgagggcagcccttctcctttctccgatatgtcctctcctctgggaatcgaccactcc<br>tgggaagataccttcgccaatgagctgttccctcagctgatctcagtgtga |

| Gene name; Human/CHO; NCBI ref; Gene size | cDNA Sequence |
|---|---|
| **XBP1s**<br>Human<br>1131 bp | atggtggtggtggccgctgctcctaatcctgccgatggcacccctaaggtgctgctgctgtctggccagcctgcttctgctgctggtgctcctgctggacaggccctgcctctgatggtgcctgctcagagaggc<br>gcttctcctgaggctgcttctggcggactgccccaggccagaaagagacagagactgacccacctgtcccccgaggaaaaggccctgcggcggaagctgaagaacagagtggctgcccagaccgccagg<br>gaccggaagaaagccagaatgtccgagctggaacagcaggtggtggacctggaagaggaaaaccagaagctgctgctggaaaaccagctgctgagagaaaagacccacggcctggtggtggaaaatc<br>aggaactgcggcagcggctgggcatggatgctctggtggctgaggaagaggccgaggccaagggcaatgaagtgcgacctgtggccggctctgccgaaagtgctgctggcgctggacctgtcgtgacccc<br>tcctgagcatctgcctatggactccggcggcatcgactcctccgactccgagtctgatatcctgctgggcatcctggacaacctggaccccgtgatgttcttcaagtgccccagccctgagccgcctccctgg<br>aagaactgcctgaggtgtaccctgagggcccctcttctctgcctgcctccctgagtctgtccgtgggcacctcttccgccaagctggaagccatcaacgagctgatcagattcgaccacatctacaccaagccc<br>ctggtgctggaaatcccctccgagacagagtcccaggccaacgtggtcgtgaagatcgaggaagccccccctgtcccctagcgagaacgaccaccctgagttcatcgtgtccgtgaaagaggaacccgtgga<br>agatgacctggtgcccgagctgggaatctccaacctgctgtcctccagccactgccccaagcctagctcctgtctgctggacgcctactccgactgtggctacggcggatctctgtccccccttctccgacatgtc<br>ctccctgctgggagtgaaccactcctgggaggacaccttcgccaatgagctgttcccccagctgatctcagtgtga |
| **YY1**<br>CHO<br>1239 bp | atggcatctggcgataccctgtatatcgccaccgacggctctgagatgcctgccgagattgtggaactgcacgagatcgaggtggaaacaatccccgtggaaaccatcgagacaaccgtcgtgggcgaaga<br>ggaagatgaggacgaagaggatgaagatggtggtggcggtggcggcggaggtggacatggacatgctggacaccatcaccatcatcatcaccaccatcaccatccacctatgatcgccctccagcctctg<br>gtcaccgatgatcctacacaggtgcaccaccaccaagaagtgatcctggtccagaccagagaggaagtcgtcggcggcgacgattctgatggactgagagctgaggacggcttcgaggaccagatcctga<br>ttcctgttcctgctccagccggcggagatgacgattatatcgagcagaccctcgtgaccgtggccgctgctggaaaatctggtggcggaggatctagttcctctggcggaggcagagtgaagaaaggcggcg<br>gaaagaagtccggcaagaagtcttatttgtctggcggtgctggtgcagcaggcggaggcggagctgatcctggaaacaagaagtgggagcagaagcaggtccagatcaagaccctggaaggcgagttct<br>ccgtgaccatgtggtcctccgacgagaagaaggacatcgaccacgagacagtggtggaagaacagatcatcggcgagaactccccacctgactactccgagtacatgacaggcaagaagctgcctcctgg<br>cggcatccctggaatcgacctgtctgatcctaagcagctggccgagttcgcccggatgaagcctagaaagatcaaagaggacgacgcccctcggacaatcgcttgtcctcataagggctgcaccaagatgtt<br>ccgggacaactccgccatgcggaaacatctgcacacccacggacctagagtgcacgtgtgcgctgagtgtggcaaggccttcgtggaatcctctaagctgaagcggcatcagctggtgcataccggcgaga<br>agcctttccagtgcacctttgaaggctgcggcaagcggttctccctggacttcaatctgagaacccacgtgcgcatccacaccggcgatagaccttacgtgtgccctttcgacggctgcaacaagaaattcgc<br>ccagtccaccaacctgaagtcccacatcctgacacacgccaaggccaagaacaaccagtga |