Characterisation of Cell Lines with Engineered Golgi Organisation

Hannah Spencer

PhD

University of York

Biology

January 2022

Abstract

N-linked glycosylation is an important modulator of protein structure and function, relevant for regulating cellular characteristics and properties of recombinantly expressed glycoprotein biopharmaceuticals, including monoclonal antibodies (mAbs). Biopharmaceutical glycosylation is tightly regulated, and glycan engineering is attractive to the biopharmaceutical industry. With no molecular template, glycosylation is controlled, inter alia, by localisation of enzymes within the Golgi apparatus. This requires tightly regulated membrane trafficking, involving specific vesicle tethering interactions, which are coordinated by the heterooctameric conserved oligomeric Golgi (COG) complex. COG mutations alter glycosylation, exemplified in congenital disorders of glycosylation, and in COG-mutant cell lines. Different COG subunit mutations cause unique glycosylation changes, highlighting the potential of targeting COG for generating glycan-engineered biotherapeutics.

The Ungar group previously identified mutations in COG subunits that altered specific interactions with Rabs, which are small GTPases with regulatory roles in membrane trafficking. A Cog4L36P mutation was identified with impaired Rab30 binding. In this work, a CHO-K1 cell line expressing Cog4L36P was engineered, characterised, and tested as a host for production of the mAb, Herceptin. The mutation was also engineered into an ex-production host CHO cell line from GlaxoSmithKline for testing Herceptin production in cells optimised for high-level mAb production. The goal was to dissect the function of the Cog4-Rab30 interaction in controlling glycosylation and how this affects cell properties and recombinant biotherapeutic expression. Results show that Cog4L36P cells have altered whole-cell glycosylation, impaired adhesion, slower proliferation and altered metabolic activity. Glycan profiling of WT, ΔCog4 and Cog4L36P cells cultured in monolayer and suspension also revealed cell line specific glycosylation changes in response to suspension adaptation. Cog4L36P did not affect Herceptin production in ex-production CHO cells. However, suspension-adapted Cog4L36P CHO-K1 cells have enhanced Herceptin production with similar glycoforms to Herceptin produced by WT cells, demonstrating the potential of the Cog4L36P mutation for aiding mAb production.

Table of Contents

List of Figures

List of Tables

Abbreviations

2AB	2-Aminobenzamide
ACN	Acetonitrile
ADCC	Antibody-dependant cellular cytotoxicity
alg	Asparagine-linked glycosylation
ASGPR	Asialoglycoprotein receptor
ATP	Adenosine triphosphate
Bmp2	Bone morphogenic protein 2
CATCHR	Complexes associated with tethering containing helical rods
CDC	Complement-dependent cytotoxicity
CDG	Congenital disorder of glycosylation
CHO	Chinese hamster ovarian
CMP	Cytidine monophosphate
CNX	Calnexin
COG	Conserved oligomeric Golgi
COP	Coatamer protein
CRISPR	Clustered regularly interspaced short palindromic repeats
CRT	Calreticulin
DC	Direct current
DHB	2,5-Dihydrobenzoic acid
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dol-P	Dolichol phosphate
Dol-PP	Dolichol pyrophosphate
ECM	Extracellular matrix
EDEM	ER degradation-enhancing mannosidase-like
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
ELISA	Enzyme linked immunosorbent assay
EPO	Erythropoietin
ER	Endoplasmic reticulum

ERAD	ER associated degradation
ERGIC	ER-Golgi intermediate compartment
ESI	Electrospray ionisation
Fab	Antigen binding fragment
FANGS	Filter-aided *N*-glycan separation
FASP	Filter-aided sample preparation
FBS	Foetal bovine serum
Fc	Fragment crystallisable
FDA	Food and drug agency
FcγR	Fc gamma receptor
FS	Forward scatter
FSH	Follicle stimulating hormone
FT	Flowthrough
Fut8	Fucosyltransferase 8
FWHM	Full width half maximum
GAD	Gal4 activation domain
GAG	Glycosaminoglycan
Gal	Galactose
GalT	Galactosyltransferase
GBD	Gal4 binding domain
GDP	Guanidine diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
Glc	Glucose
GlcNAc	*N*-acetylglucosamine
GST	Glutathione S-transferase
GTC	Golgi transport complex
GTP	Guanidine triphosphate
HC	Heavy chain
HDR	Homology directed repair
HEK	Human embryonic kidney
HER	Human epidermal growth factor receptor

HILIC	Hydrophilic interaction chromatography
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
HS	Heparan sulfate
IGFIR	Insulin-like growth factor I receptor
IgG	Immunoglobulin G
IVF	*In vitro* fertility
LacNAc	Galβ1-4GlcNAc
LAMMA	Laser-micro-mass analysis
LDI	Laser desorption/ionisation
LDLR	Low-density lipoprotein receptor
mAb	Monoclonal antibody
MALDI	Matrix-assisted laser desorption/ionisation
Man	Mannose
ManI	Mannosidase I
ManII	Mannosidase II
MGAT1	Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
MGAT2	Alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase
MGAT3	Beta-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase
MGAT4	Alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase
MGAT5	Alpha-1,6-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase
MS	Mass-spectrometry
MSC	Mesenchymal stem cell
MTC	Multi-subunit tethering complex
MTT	3- [4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide
NB	Nucleotide Binding
NE	Nucleotide Exchange

Neu5Ac	*N*-acetylneuraminic acid
NSF	N-ethylmaleimide–sensitive factor
oa	Orthogonal acceleration
OST	Oligosaccharyl transferase
PAGE	Polyacrylamide gel elecrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
PNGase F	Peptide N-glycosidase F
RF	Radio frequency
rh	Recombinant human
RNase	Ribonuclease
RTK	Receptor tyrosine kinases
SDS	Sodium dodecyl sulfate
SiaT	Sialyltransferase
SM	Sec1/munc19
S:N	Signal to noise
SNARE	Soluble NSF attachment protein receptor
SS	Side scatter
ssODN	Single-stranded oligodeoxynucleotides
SWS	Saul Wilson Syndrome
TAE	Tris-acetate-EDTA
TFA	Trifluoroacetic acid
TGF-β	Transforming growth factor beta
TGN	*Trans*-Golgi network
TNFα	Tumour necrosis factor alpha
TOF	Time-of-flight
UDP	Uridine diphosphate
UGGT	UDP-glucose glycoprotein glucosyltransferase
UHPLC	Ultra-high-performance liquid chromatography
VSVG	Viral G-protein of vesicular stomatitis virus
WT	Wild-type

Acknowledgements

I would first like to thank my academic supervisors, Daniel Ungar, Jane Thomas-Oates and Kirsty Skeene for their constant support, guidance, and encouragement throughout my PhD. Their academic input was invaluable, and their personal support got me through many difficult times over the years, not least a global pandemic. I owe thanks to my TAP members, Bob White, and Gareth Evans, for providing welcome advice and ensuring I looked forward to every TAP and progression meeting. I would also like to thank Nia Bryant for her continued support in my project and for hosting many engaging data discussions in her office. Thanks, are also due to BBSRC and GlaxoSmithKline for funding this work.

I owe a great debt of thanks to members of the Technology Facility for their training and support on the wealth of amazing equipment in the Department of Biology, particularly Graeme Park and Karen Hogg for their assistance in flow cytometry. Many thanks are likewise due to members of the Centre of Excellence in Mass Spectrometry, Ed Bergström, Adam Dowle, and Chris Taylor, for teaching and supporting me for many hours on mass spectrometers. I would like to thank Mitul Patel, Robyn Emmins, and everyone I worked with in Biopharm R&D, for their supervision during my time at GSK. Special thanks go to Shawal Spencer at GSK, who went above and beyond to keep my cells alive while site access was heavily restricted by the pandemic.

To everyone in York who I've had the pleasure of working with, thank you for making my time here so enjoyable. To Peter, Elizabeth, Lizzie, Katherine, Dave, Hannah, Dimi, Kamilla, Caroline, Richard, John, David, Ben, Becca, Simeon, Savvas, Bella, Brett, and Andy: thank you for your friendship and many lab-related favours over the years. I've also worked with some excellent undergraduate and master's students, particularly Emily, Maya, and Andrea, who've made valuable contributions to this PhD.

Finally, many thanks are due to my family and friends for their friendship and personal support, which has kept me sane throughout my PhD. To my housemates, Hannah, and Sophie, our many house adventures provided welcome relief from the stress of thesis writing and national lockdowns. Huge thanks must also go to my partner John, who has loved, supported, and fed me in the most difficult final stages of finishing my PhD.

Author's Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. The work presented in this thesis was performed between October 2017 and November 2021. Experiments were carried out in the Department of Biology, University of York, in the laboratory of Professor Dani Ungar, in the Department of Chemistry, University of York, in the laboratory of Professor Jane Thomas-Oates and in the BioPharm Research and Development labs at GlaxoSmithKline, Stevenage, supervised by Dr Kirsty Skeene. All experiments were performed by the author, with the exception of a few cell-line characterisation experiments, performed by other members of the academic group under supervision of the author. These are clearly indicated as such in the thesis. Results of a metabolic activity experiment (figure 2.13) were previously submitted by Emily Priest in her undergraduate dissertation for acceptance of a bachelor's degree. Results of adhesion assays (figure 2.15) were also previously submitted in a different format by Maya Trivedy in her dissertation for acceptance of an integrated master's degree. Figures that contain data from their work are clearly marked. All other sources are acknowledged as references. With the aforementioned exceptions, neither this thesis nor any other part of it has previously been submitted for acceptance of a higher degree.

Introduction

1.1 Glycobiology

All living organisms ultimately consist of small molecule metabolites and a few broad classes of biological macromolecules, including nucleic acids, proteins, and lipids. Deoxyribonucleic acid (DNA) is the molecular blueprint determining the sequences of amino acids that fold into proteins: the central dogma driving the whole field of molecular biology (Crick, 1958). Proteins perform a huge range of functions, from catalysing biochemical reactions to critical structural roles. Lipids are essential components of the membranes encasing every living cell and their intracellular components. However, none of these macromolecules function alone, with covalent modifications altering their structures and properties.

One such modification is the addition of carbohydrates (glycans) to other organic molecules. In conjunction with other macromolecules, notably proteins and lipids, glycans are involved in numerous processes including biomolecular interactions, cell-cell signalling, and adhesion of cells to their extracellular matrix (reviewed in Varki, 2016). Glycans are ubiquitous throughout biology: all cells and most proteins are decorated with an array of covalently attached carbohydrates. Glycans are so prevalent that they can even be seen as a molecular fuzz (glycocalyx) on the surface of many cells (Bhattacharyya et al., 1989).

Along with DNA and protein, carbohydrates are considered the third main class of biopolymer. They consist of monosaccharide subunits, which can form a very large range of branched oligosaccharide structures. However, unlike the template driven synthesis of proteins from DNA, carbohydrate biosynthesis is not controlled by a DNA or other biomolecular blueprint.

1.1.1 Carbohydrate Structure

The definition of a carbohydrate is a biomolecule consisting predominantly of carbon, hydrogen, and oxygen atoms. The simplest carbohydrates are monosaccharides, which are the basic building blocks of oligo- and polysaccharides. Monosaccharides commonly consist of three or more carbon atoms, a carbonyl group and two or more hydroxyl group. Structurally, monosaccharides generally exist in equilibrium between a straight chain structure and a closed ring form which arises from a reversible reaction between the carbonyl group and one of the hydroxyl groups. The resulting ring is a hemiacetal or hemiketal, depending on whether the open chain form has an aldehyde (figure 1.1) or ketone functional group. The carbon atoms in the open structure are numbered, based on the open linear (Fischer) projection of the monosaccharide, with position one assigned to the first carbon from the top, which is always drawn closest to the carbonyl carbon. For aldoses, C-1 is the carbonyl carbon but for ketoses, the carbonyl carbon is at least C-2. In the case of glucose ($C_6H_{12}O_6$), the hydroxyl group on C-5 is typically involved in the ring formation, with the hemiacetal oxygen atom connecting C-1 and C-5 (figure 1.1). However, it can also form a five membered ring if the hydroxyl on C-4 reacts with the carbonyl oxygen. Rings made from five atoms are called furanose and those with six are called pyranose (figure 1.1). On ring formation, the carbonyl carbon becomes chiral, so that it has two possible anomers, with a hydroxyl group sitting either above (β) or below (α) the plane of the ring relative to the substituent attached to the other carbon atom flanking the ring oxygen (the CH_2OH group attached to C-5 in the case of gluco-pyranose) (figure 1.1).
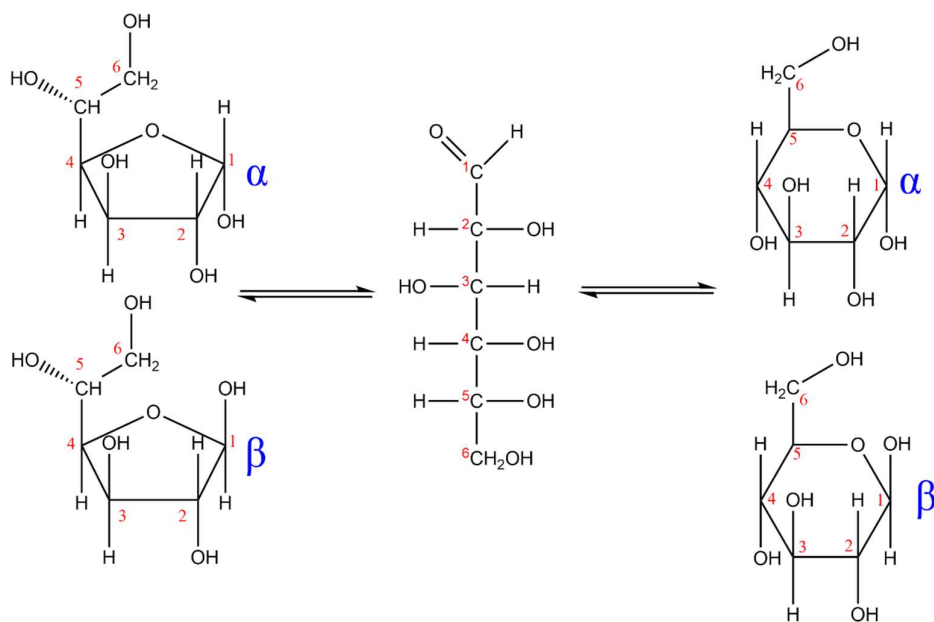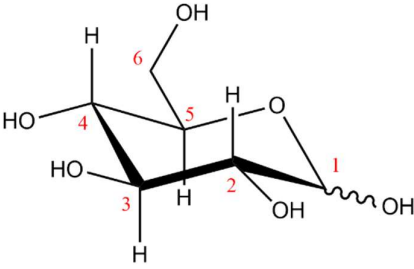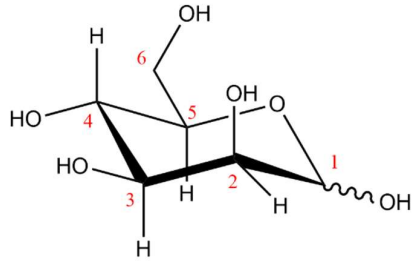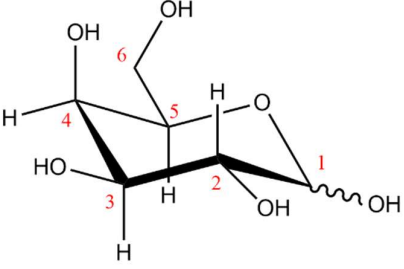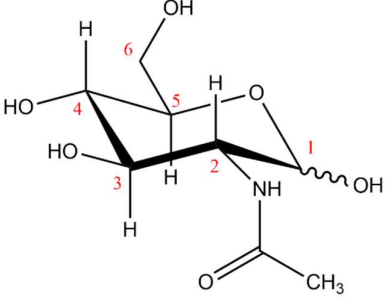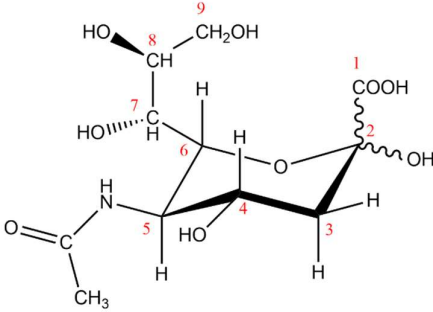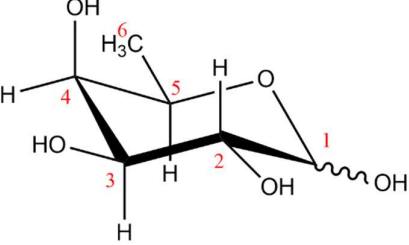
Figure 1.1 | Structure of glucose. Glucose is represented in its linear form as a Fischer projection in the centre, showing reversible ring closure reactions to form a five-membered furanose ring (left) or a six-membered pyranose ring (right). Ring structures are shown as Haworth projections (Haworth, 1925) Carbon positions are numbered in red. For each cyclic structure, both α and β anomers are shown.

Monosaccharides are classified based on the position of the carbonyl group and the number of carbon atoms, for example glucose is an aldohexose because it has six carbon atoms (suffixed hexose) and an aldehyde group (prefixed aldo) (figure 1.1). Those with ketone groups are classed as ketoses and prefixed keto and those with three, four or five carbon atoms are suffixed triose, tetrose, and pentose respectively. The most common monosaccharides found in glycoproteins are hexoses, including glucose (Glc), galactose (Gal) and mannose (Man), which are all isomers with the monosaccharide chemical formula $C_6H_{12}O_6$ (table 1.1). Many carbohydrates have a hydrogen-oxygen atom ratio of 2:1 and a chemical formula of $C_m(H_2O)_n$, however there are multiple monosaccharide substituents of glycoconjugates that do not conform to this rule. For example, some monosaccharides have deoxy modifications, including fucose, $C_6H_{12}O_5$, which is an aldohexose like galactose but with four hydroxyl groups instead of five. Other exceptions include those with amide modifications, including *N*-acetylglucosamine (GlcNAc), $C_8H_{15}NO_6$, and sialic acids, such as *N*-acetylneuraminic

acid (Neu5Ac), $C_{11}H_{19}NO_9$. Neu5Ac is the most common sialic acid found in eukaryotes, consisting of a nine-carbon backbone, which forms a six-membered ring with a carboxylate group at C1, a hemiketal between C2 and C6 and an *N*-acetyl functional group on C5. All the monosaccharide components of glycoprotein glycans in eukaryotes that are relevant to this thesis are summarised in table 1.1.

Table 1.1 | Summary of all the monosaccharide components of glycoproteins discussed in this thesis. For each monosaccharide, the full and abbreviated names are displayed, alongside the chemical formula, structure, and the symbol nomenclature for glycans (SNFG) symbol. Structures shown are the pyranose forms, with carbon atom numbering in red and bonds representing two different anomers shown with a wavy line.

Name	Formula	Structure	Symbol
Glucose (Glc)	$C_6H_{12}O_6$		
Mannose (Man)	$C_6H_{12}O_6$		

Galactose (Gal)	$C_6H_{12}O_6$		
N-acetylglucosamine (GlcNAc)	$C_8H_{15}NO_6$		
N-acetylneuraminic acid (Neu5Ac)	$C_{11}H_{19}NO_9$		
Fucose (Fuc)	$C_6H_{12}O_5$		

1.1.2 Glycans: disaccharide and oligosaccharide formation

Glycans, which are oligosaccharides, are formed when monosaccharides are covalently joined by glycosidic linkages. Disaccharides are the simplest glycans, consisting of two monosaccharides joined by a covalent glycosidic bond between the hemiacetal or hemiketal carbon and a hydroxyl group from another monosaccharide. Glycosidic bond formation is a condensation reaction, which is generally catalysed by glycosyl transferase enzymes in the cell. The substrates are monosaccharides bound to nucleoside phosphates, which act as donor molecules for the monosaccharides, from which they are transferred onto other monosaccharides or oligosaccharides. For example, glucose, galactose and GlcNAc are donated from uridine diphosphate (UDP), fucose and mannose from guanidine diphosphate (GDP) and sialic acid is donated from cytidine monophosphate (CMP).

Since almost every carbon atom in most monosaccharides is bound to a hydroxyl group, there are multiple places for a glycosidic linkage to occur (figure 1.2). Glycosidic linkages are therefore classified based on the carbon numbering system, with the position of the carbonyl carbon written first, and the position of the carbon bound to the reacting hydroxyl group of the bound monosaccharide written next (figure 1.2). These linkages are further classified based on whether the glycosidic bond is in the α or β anomeric position relative to the carbonyl carbon (figure 1.2).

Oligosaccharides (glycans) are formed from three or more monosaccharides joined by glycosidic linkages. If multiple monosaccharides are joined by the same glycosidic linkage, a linear glycan chain is formed. However, unlike polypeptide chains, which are always linear, many glycans in nature have branched structures owing to the ability of a single monosaccharide to form glycosidic linkages from each of its hydroxyl groups. In the example shown in figure 1.2, C-1 of a β-mannose residue forms a glycosidic linkage with the hydroxyl group on C-4 of *N*-acetylglucosamine (a β-1,4 linkage). Carbons C-3 and C-6 of that same mannose residue also form α-1,3 and α-1,6 glycosidic bonds to α-mannose residues (figure 1.2). These other mannose residues can then form further glycosidic linkages, creating a glycan with multiple branches, termed bi-, tri- or tetra-antennary, depending on the number of branches. It is also possible for the central mannose residue (figure 1.2) to link to a second GlcNAc residue via a 1,4 linkage

with carbon 1 of the additional GlcNAc, forming an intersecting glycan. Branching can thus give rise to huge structural diversity in glycans.
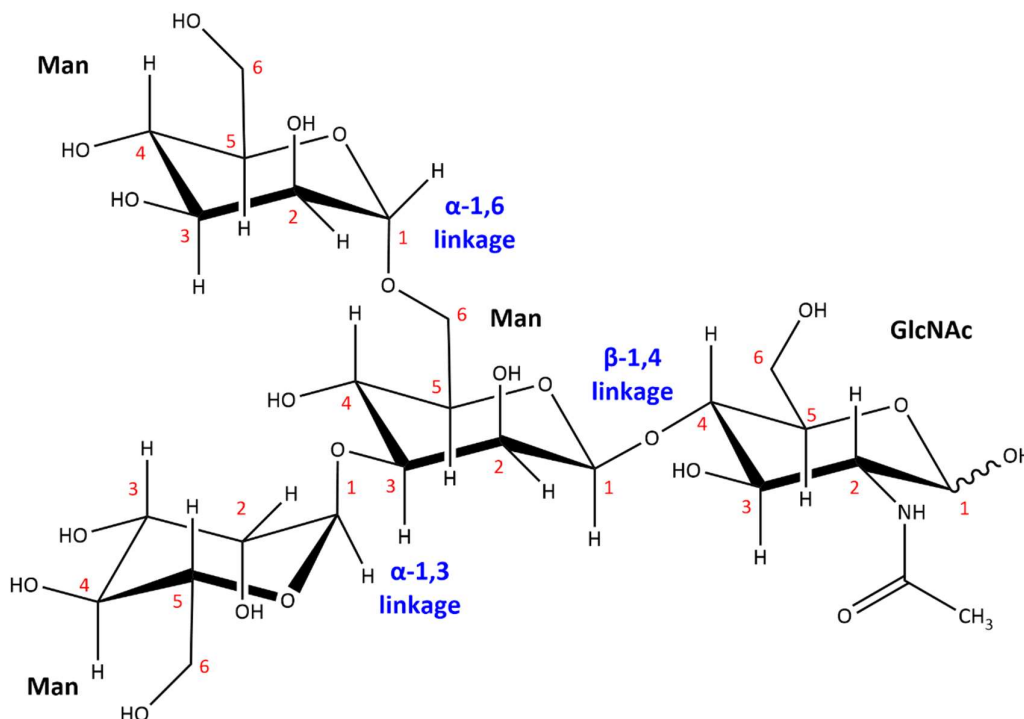
Figure 1.2 | Glycosidic Linkages. A glycan structure, consisting of three mannoses and N-acetylglucosamine, labelled with three different glycosidic linkages. Carbon atom positions within each pyranose monosaccharide are labelled in red.

1.1.3 Protein Glycosylation

The work presented in this thesis only concerns glycoproteins, specifically in mammalian cells, so the aim of the following sections is to introduce the process of protein glycosylation and some of the important functions of glycoproteins. Protein glycosylation is the process of adding glycan chains to newly synthesised proteins in the endoplasmic reticulum (ER) and Golgi apparatus. There are two main types of protein glycosylation, *N*-linked and O-linked, which are classified based on the atoms involved in the covalent bond between the glycan and the amino acid (Figure 1.3). Many glycosylation sites can be predicted as glycans are only added at a specific consensus sequence or onto particular amino acids. In the case of *N*-linked glycosylation, at the Asn-X-Ser/Thr consensus sequences (where X is any amino acid

except proline) (Marshall, 1972), the glycan chain is bound by an amide linkage between C-1 of a GlcNAc residue and the amide nitrogen atom of the asparagine side chain. O-linked glycosylation occurs between a glycan and the hydroxyl group of, generally, either a serine or threonine side chain. Unlike *N*-linked glycosylation, there is no consensus sequence for O-linked glycosylation and predicting it remains enigmatic.
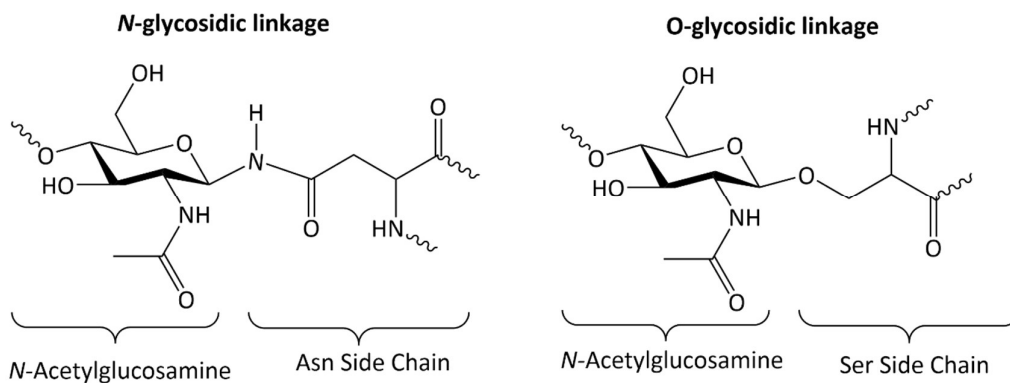
Figure 1.3 | Structures of N- *and O-linked glycosidic bonds.* N-*glycans always form an amide linkage with the amide group of an asparagine side chain (left). O-glycans can link glycosidically to hydroxyl groups of either serine (right) or threonine sidechains, though only serine is presented in this figure for simplicity.*

Both types of protein glycosylation are highly heterogenous. Not all potential glycosylation sites on a protein become glycosylated, in different molecules of the same glycoprotein, specific glycosylation sites are not always glycosylated, and different glycosylation sites can have different glycans attached to them. Micro-heterogeneity arises from the propensity for multiple different glycans to occupy the same glycosylation site on different molecules of the same protein. Glycan processing is also a stochastic process, meaning that each reaction in the biosynthetic pathway has a random probability distribution and cannot be predicted precisely. Since glycans influence both the structure and stability of proteins, this unpredictability and heterogeneity complicates the central dogma that protein structure and function are encoded entirely by a DNA template. Although glycosylation enzymes and other machinery are encoded by DNA, external factors such as nutrient availability can also

influence glycosylation, thus impacting glycoprotein structure and function in a DNA independent manner.

1.1.4 *N*-glycan Biosynthesis in the Endoplasmic Reticulum

The work presented in this thesis focuses on *N*-linked glycosylation, so the following sections all relate to *N*-glycans. The biosynthesis of *N*-glycans begins at the cytosolic face of the ER membrane with the transfer of UDP-GlcNAc onto the ER-resident lipid, dolichol phosphate (Dol-P), to generate a GlcNAc-conjugated dolichol pyrophosphate (Dol-PP). This is achieved by the transmembrane glycosyltransferase, *N*-acetylglucosaminyltransferase (Kukuruzinska et al., 1994), which also adds a second GlcNAc residue onto Dol-PP-GlcNAc. The addition of five mannose residues from GDP-Man is then catalysed by a series of transmembrane mannosyltransferases, which are members of the ALG protein family due to their original identification in yeast *alg* (asparagine-linked glycosylation) mutants (Ballou et al., 1980). The resulting Dol-PP-$GlcNAc_2Man_5$ is then flipped across the membrane into the ER lumen by an as yet unidentified flippase enzyme (Rush, 2015). Inside the ER, a further four mannoses are added, followed by the addition of three glucose residues, catalysed by glucosyltransferase enzymes. The substrates required for these reactions are Dol-P-Man and Dol-P-Glc respectively, which are also flipped across the ER membrane by other flippase enzymes (Rush, 2015).

Meanwhile, proteins are synthesised at ribosomes and translocated through translocons into the ER lumen. In the case of integral membrane proteins, these are embedded into the RER membrane shortly following translation and from then, only the side of the protein facing the ER lumen is glycosylated. At Asn-X-Ser/Thr consensus sequences to be glycosylated (Marshall, 1972), *N*-glycans are added by the *en bloc* transfer of $GlcNAc_2Man_9Glc_3$ from Dol-PP to the amide group of the asparagine residue. This process is catalysed by the multi-subunit transmembrane protein oligosaccharyl transferase (OST) and it typically occurs co-translationally (Kornfeld and Kornfeld, 1985). Once attached to the protein, the immature glycan structure undergoes trimming and processing, whilst the protein assumes its correct fold (figure 1.4).
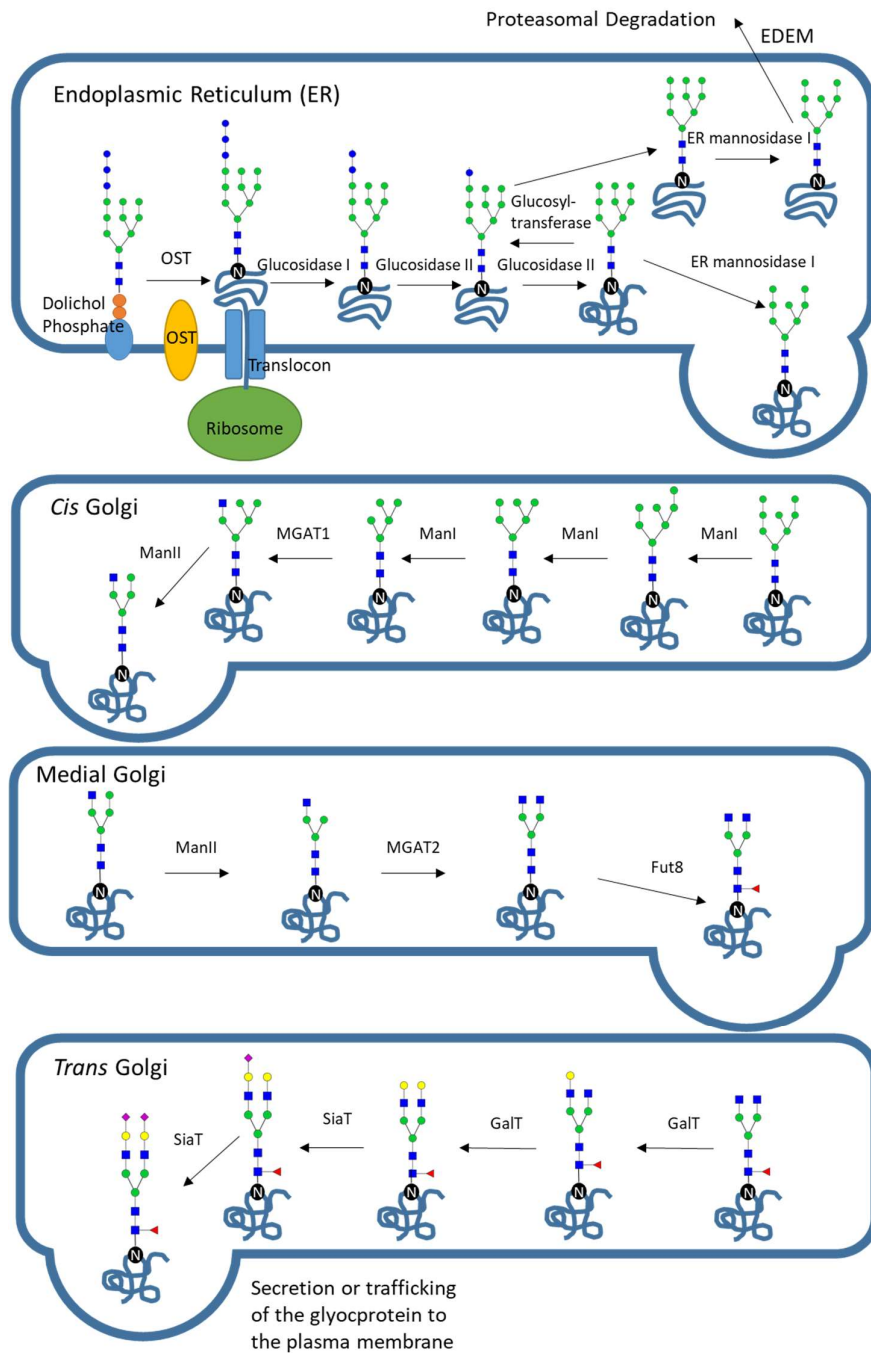
Figure 1.4 | Schematic representation of glycan processing in the ER and Golgi apparatus. A simplified summary of key glycan processing steps in the ER and Golgi, starting with en bloc transfer of GlcNAc$_2$Man$_9$Glc$_3$ from dolichol phosphate to a newly synthesised protein, catalysed by OST. A summary of the key stages of the unfolded protein response in the ER is shown. The Golgi apparatus is then simplified into three cisternae: cis, *medial, and* trans-*Golgi. An exemplar glycan biosynthesis pathway is shown, whereby GlcNAc$_2$Man$_8$ is processed into the complex, fucosylated, biantennary glycan, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$NeuAc$_2$.*

Glycosylation can also occur post-translationally (Compton and Courtney, 1984) and there is evidence to suggest that OST isoforms with different catalytic subunits and enzymatic properties, STT3A and STT3B (Kelleher et al., 2003), are involved in glycosylation at different stages of translation. Both STT3A and STT3B are required for co-translational glycosylation, but STT3B also mediates post-translational glycosylation of unfolded proteins, whilst STT3B is primarily involved co-translationally (Ruiz-Canada et al., 2009).

1.1.5 ER Glycosylation and the Unfolded Protein Response

In the ER, all three glucose residues and a mannose residue are removed by glucosidases and a mannosidase respectively, steps which occur co-translationally (Atkinson and Lee, 1984). These steps are intricately linked to protein folding (figure 1.4). The chaperone proteins calnexin (CNX) and calreticulin (CRT) recognise and bind to mono-glucosylated *N*-glycans, retaining them in the ER until proper folding is complete (Hammond et al., 1994). This is achieved by a cycle of de-glucosylation by glucosidases and re-glucosylation (Parodi et al., 1984) by UDP-glucose glycoprotein glucosyltransferase (UGGT), releasing the chaperone proteins intermittently whilst the protein undergoes the process of folding. If mis-folding occurs, re-glucosylation is prevented and an ER mannosidase enzyme trims a terminal mannose residue from the middle branch, generating an isomer of $GlcNAc_2Man_8$ (figure 1.8). This structure is recognised by lectins in the ER degradation-enhancing mannosidase-like (EDEM) family, which recognise the mis-folded glycoproteins bearing $GlcNAc_2Man_8$ and target them to the ER associated degradation (ERAD) pathway (Ruggiano et al., 2014). It is still unclear how exactly the enzymes and chaperones involved distinguish between mis-folded and correctly folded proteins, however the process is essential for glycoprotein folding and is highly conserved across eukaryotes. So vital is this process that mice deficient in CRT show embryonic lethality (Mesaeli et al., 1999), and mice lacking CNX die within days of birth (Denzel et al., 2002). Loss of UGGT in mice also causes embryonic lethality (Molinari et al., 2005). Quality control in protein folding is therefore an important function of *N*-glycosylation (Caramelo and Parodi, 2015).

1.1.6 *N*-Glycan Processing in the Golgi

All *N*-glycans consist of the same GlcNAc$_2$Man$_3$ chitobiosyl core and further glycan processing in the Golgi gives rise to numerous possible glycan structures (Kornfeld and Kornfeld, 1985). Glycoproteins exit the endoplasmic reticulum with predominantly oligomannose *N*-glycans consisting of eight or nine mannose residues, sometimes still with a single glucose residue (reviewed by Stanley et al., 2015). Immature oligommanose glycans subsequently undergo trimming and processing in Golgi cisternae to form a broad range of glycans, including oligomannose, complex and hybrid type (figure 1.5). These consist of the monosaccharides mannose, *N*-acetylglucosamine, fucose, sialic acid, and galactose (table 1.1), which can be arranged with varying numbers of branches (figure 1.6) and in different orders (reviewed by Stanley et al., 2015, Ruiz-Canada et al., 2009, Kornfeld and Kornfeld, 1985).
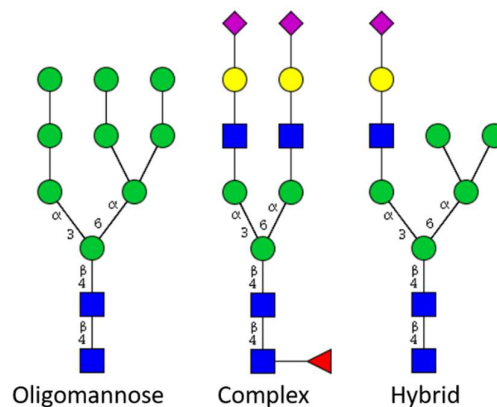
Figure 1.5 | Types of N-*glycan.* N-*glycans consist of a common Man$_3$GlcNAc$_2$ chitobiosyl core and are categorised based on the types of monosaccharide on each branch. The branches of* N-*glycans consist of either all mannose residues (oligomannose, left structure), a combination of GlcNAc, galactose and sialic acid and fucose residues (complex glycans, middle structure) or a mixture of these with one branch being complex, the other oligomannose (right structure). Figure adapted from figure 1.1 in (Fisher and Ungar, 2016).*

Oligomannose type glycans have only mannose residues attached to their core (figure 1.5). The total number of mannose residues ranges from five to nine, inclusive of the core mannoses. Complex glycans have all but the three core mannose residues removed and between two to four antennae decorated with GlcNAc residues, which

can be adorned further with galactose and sialic acid residues (figure 1.5). Hybrid glycans retain two or more mannose residues on the α-1,6 arm of their core. The α-1,3 arm of a hybrid glycan can be decorated with GlcNAc, galactose and sialic acid residues in the same way as a complex glycan (figure 1.5).

The first stages of glycan processing in the *cis* Golgi are catalysed by mannosidase I (ManI), which trims large oligomannose *N*-glycans down to $GlcNAc_2Man_5$ (figure 1.4). Oligomannose *N*-glycans can exit the Golgi without any further processing. Alternatively, a GlcNAc residue can be added to the α-1,3 mannose residue by the enzyme α-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (MGAT1), resulting in the formation of a hybrid type glycan. Mannosidase II (ManII) then trims the two remaining mannoses from the core, converting the glycan from hybrid to complex type (figure 1.4). The simplest complex glycan is therefore $GlcNAc_2Man_3GlcNAc_1$. In the medial Golgi, alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase (MGAT2) can then catalyse the addition of a second GlcNAc onto the α-1,6 mannose to form $GlcNAc_2Man_3GlcNAc_2$ (figure 1.4). Further branches can be added by alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase (MGAT4) and alpha-1,6-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase (MGAT5). MGAT4 and MGAT5 catalyse the β-1,4 linkage of GlcNAc residues to the outermost core mannose residues to yield tri- and tetra-antennary glycans (figure 1.6a). An additional GlcNAc residue can also be added by a β-1,4 linkage to the innermost core mannose residue of complex or hybrid glycans by the enzyme beta-1,4-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase (MGAT3), to yield an intersecting *N*-glycan (figure 1.6b). These stages of the glycan biosynthesis pathway are well-understood and reviewed in (Kornfeld and Kornfeld, 1985).

All GlcNAc residues bound to the outer core mannoses can be further decorated in the *trans*-Golgi by the addition of galactose and sialic acid residues, which are added by galactosyltransferase (GalT) and sialyltransferase (SiaT) enzymes respectively (figure 1.4). Galactose residues can either be capped with sialic acid (figure 1.6a) or extended by repeating units of the disaccharide Galβ1-4GlcNAc (LacNAc), known as poly-LacNAc repeats. Poly-LacNAc repeats were first discovered on erythrocytes and are

preferentially added to multiantennary glycans, most commonly to the β1-6GlcNAc branch synthesised by MGAT5 (figure 1.6c). The biosynthesis of poly-LacNAc is catalysed by the enzyme, β-1,3 N-acetylglucosaminyltransferase (β3gnt), which adds GlcNAc to *N*-glycan Gal termini. Poly-LacNAc glycans are the largest form of complex glycans as they can have upwards of 20 poly-LacNAc repeats (Krusius et al., 1978).
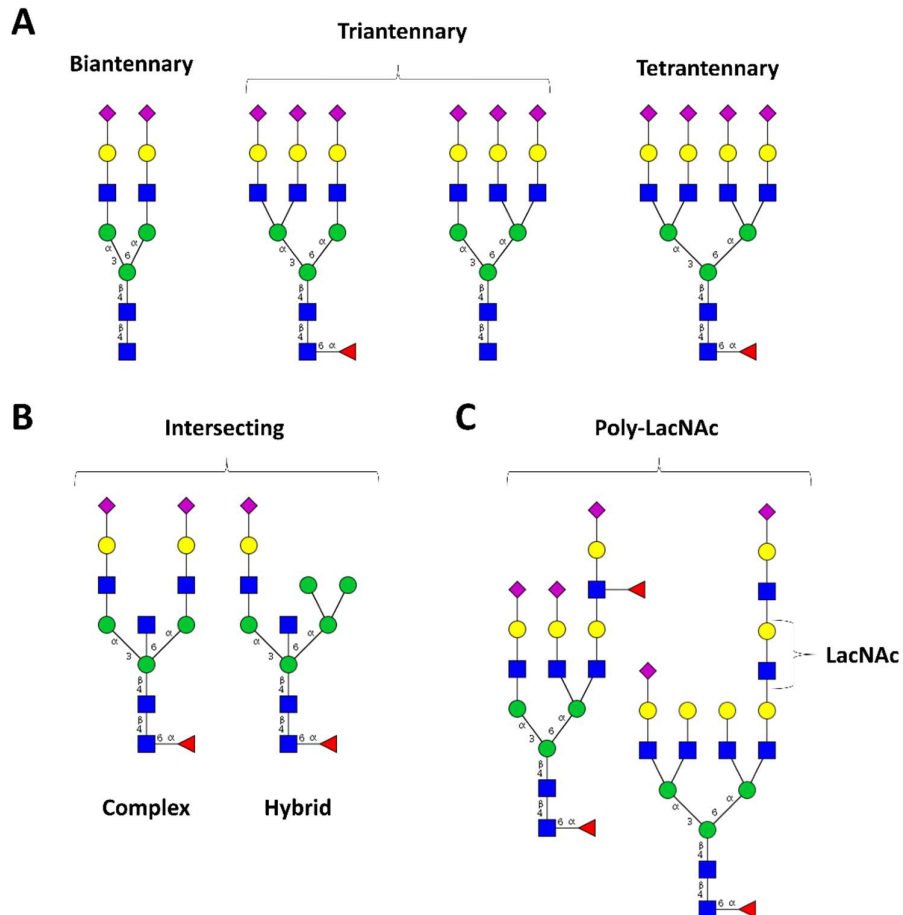
Figure 1.6 | Examples of branched complex and hybrid N-*glycan structures. A) Multiply branched* N-*glycans B) Intersecting* N-*glycans C) poly-LacNAc type* N-*glycans. These are only representative examples of each* N-*glycan type. In nature,* N-*glycan core structures can either be fucosylated or not. Branches can terminate in either GlcNAc, Gal or NeuAc. GlcNAc and Gal residues within the branches can also be fucosylated. Antenna GlcNAcs can also be sialylated and occasionally other monosaccharides, such as GalNAc can be added.*

Fucosylation adds another level of complexity to *N*-glycans. In the medial Golgi, a fucose residue can be added to the first GlcNAc within the core, catalysed by the enzyme fucosyltransferase 8 (Fut8) (figure 1.4). GlcNAc residues in *N*-glycan antennae can also be fucosylated, catalysed by fucosyltransferases 3 and 4. This can occur on antennal GlcNAc residues bound directly to the core mannoses or within poly-LacNAc repeats. If a fucosylated GlcNAc is also bound to a galactose residue, the resulting structure is called a Lewis-x epitope, or a sialyl Lewis-x epitope if the galactose residue is sialylated.

The process of glycosylation is highly heterogeneous, and glycoproteins can follow many different *N*-glycosylation reaction pathways in the Golgi. Each cisterna hosts multiple competing glycan processing enzymes, which can act in different orders and on numerous different glycan structures, owing to their substrate specificity often being focussed on a specific residue, regardless of the rest of the glycan. Whilst the number of possible *N*-glycan structures that can be produced in the Golgi by a given organism is very large, the number of possible reaction pathways is larger, with multiple pathways producing the same *N*-glycan. This makes predicting *N*-glycosylation difficult. Therefore, the aim of this section was not to introduce every possible *N*-glycan biosynthesis pathway or every possible structure, but to paint a picture of the complexity inherent in the field of glycobiology.

1.1.7 Biological Functions of Mammalian Glycoprotein *N*-glycans

Broadening understanding of the factors affecting glycosylation is of great importance considering the numerous functional roles played by *N*-glycans. On top of the well-characterised involvement of *N*-glycans in protein folding in the ER and Golgi apparatus (figure 1.4) (Hammond et al., 1994), they can also impact the stability of glycoproteins (reviewed by Wormald and Dwek, 1999). This is important, for example in immunity, where glycosylation of immunoglobulin G (IgG) antibody can affect its quaternary structure (Corper et al., 1997), and alter its serum half-life. Glycoprotein *N*-glycans are also directly involved in numerous biological processes including cell-cell recognition and signalling, cell-cell adhesion and adhesion to extracellular matrices (ECM) (reviewed in Varki, 2016). Note that *N*-glycan functions are difficult to categorise as, for example, influence on protein stability or folding can have knock-on effects on cell

signalling or adhesion if the glycoprotein to which they are attached is involved in either function. The following sections are not extensive reviews of all the functions of glycoprotein *N*-glycans, which would be beyond the scope of this thesis, but rather they are summaries of key functions relevant to the work discussed in this thesis.

1.1.7.1 Protein folding and stability

Early studies aiming to understand the functions of *N*-glycosylation focussed on inhibiting steps in the glycan biosynthetic pathway and studying the effects (Berger et al., 1982). For example, treatment with tunicamycin prevents all *N*-glycosylation by inhibiting the enzyme N-acetylglucosamine phosphotransferase, which transfers N-acetylglucosaminyl-1-phosphate to phosphorylated dolichol (Duksin and Mahoney, 1982). Tunicamycin was used to study the synthesis of the viral G-protein of vesicular stomatitis virus (VSVG) (Gibson et al., 1979), which is synthesised and glycosylated in host-cells during the course of infection. It was found that treatment with the drug inhibited virion production. VSVG failed to reach the plasma membrane despite being synthesised in normal amounts, due to improper folding and aggregation of the protein (Gibson et al., 1979). Treatment of cancer cells with tunicamycin also reduces levels of various receptor tyrosine kinases (RTKs), including insulin-like growth factor I receptor (IGFIR) and epidermal growth factor receptor (EGFR), which are both normally heavily *N*-glycosylated, through retention in the ER and Golgi (Contessa et al., 2008). Tunicamycin treatment has also previously been shown to reduce IGFIR expression at the cell surface, resulting in apoptosis (Dricu et al., 1997). However, inhibiting glycosylation completely has mixed results due to its variable and unpredictable effects on different proteins and cell types (Varki, 2016, Berger et al., 1982).

It is also difficult to study the effects of *N*-glycans on protein structure due to their inherent flexibility, which makes structural determination of glycoproteins by crystallography ambiguous. Nevertheless, there is spectroscopic evidence that *N*-glycosylation can alter the stability of some proteins (Wormald and Dwek, 1999, Hanson et al., 2009). For example, bovine pancreatic ribonuclease (RNase) exists in two forms: the aglycosylated RNase A, and the monoglyocosylated RNAse B. There is no difference in the peptide sequence of each form (Williams et al., 1987). However, studying the dynamics of each RNAse form revealed that glycosylation of RNase B

reduced solvent access to regions of the protein backbone, likely due to enhanced stability, reducing fluctuations of the peptide backbone (Joao and Dwek, 1993). In more recent times, bioinformatics approaches have been employed to simulate the effects of glycosylation on protein structures, also revealing a decrease in protein dynamics and enhanced stability caused by *N*-glycosylation (Lee et al., 2015a).

1.1.7.2 Glycoprotein steady-state levels

In addition to their roles in protein folding and stabilising protein structure, *N*-glycans can also influence the steady-state levels of protein through their role in glycoprotein localisation, with mis-localisation often leading to degradation. For example, in glycosylation-deficient, mutant Chinese hamster ovarian (CHO) cell lines, identified on the basis of altered low-density lipoprotein receptor (LDLR) activity, LDLRs were synthesised to essentially normal levels and were able to reach the cell surface, but were degraded much more rapidly than in wild-type cells (Kingsley et al., 1986). This was likely due endocytosis and degradation in the endolysosomal system. These studies not only highlight the importance of *N*-glycosylation for maintaining steady-state protein levels but also for protein localisation.

Another mechanism by which *N*-glycans can influence glycoprotein steady-state levels is via interactions with asialoglycoprotein receptors (ASGPRs) (Ashwell and Morell, 1974, Ashwell and Harford, 1982) or mannose receptors (Lee et al., 2002), which remove proteins from serum by endocytosis. ASGPRs bind preferentially to asialoglycans with exposed galactose residues, mainly in hepatocytes (Ashwell and Morell, 1974). Mannose receptors bind to fucose and GlcNAc, as well as mannose (Stahl, 1990, Taylor and Drickamer, 1993). The presence of these receptors means that *N*-glycans can influence the serum half-life of glycoproteins, such as IgG (figure 1.7). IgGs are often administered as biotherapeutics and the presence of high-mannose Fc *N*-glycans on therapeutic IgG has been shown to result in increased serum clearance in humans due to endocytosis upon binding to mannose receptors (Goetze et al., 2011).

1.1.7.3 Signalling

In some of the previous examples, *N*-glycosylation affects signalling through altering the expression of signalling receptors. *N*-glycans on glycoprotein receptors or their ligands are also directly involved in signalling through modulating signalling

interactions. This is important in immunology, where IgG mediates pro- and anti-inflammatory activities through interaction of its Fc fragment (Fc) (figure 1.7) with distinct Fcgamma receptors (FcγRs). *N*-glycosylation of the Fc region can modulate antibody effector function and different types of *N*-glycans can alter functions in different ways. For example, sialylation of Fc *N*-glycans correlates with anti-inflammatory activity in the steady state, and in response to an antigen-specific immune response, sialylation decreases, and the antibody confers pro-inflammatory properties (Kaneko et al., 2006). This effect is prominent in autoimmune and inflammatory diseases, such as in rheumatoid arthritis, where Fc *N*-glycans are often deficient in galactosylation and sialylation (Parekh et al., 1985). It was found that serum levels of the antibody were unaffected by removing sialylation, suggesting the mechanism was independent of ASGPRs. Instead, sialylation enhanced binding to an anti-inflammatory subclass of Fcγ receptors, FcγRIIB (Kaneko et al., 2006), which cross-link with pro-inflammatory FcγRIII and inhibit a pro-inflammatory signalling cascade (reviewed in Smith and Clatworthy, 2010).

Another well-documented relationship between antibody Fc *N*-glycans and their effector function is the effect of core fucosylation on antibody dependent cellular cytotoxicity (ADCC), which is discussed in more detail in section 1.3.3.1. Briefly, a lack of core fucosylation has been shown to improve IgG binding to human FcγRIII, triggering ADCC (Shields et al., 2002, Shinkawa et al., 2003). This effect is likely due to steric hindrance between core fucose on the Fc *N*-glycan and *N*-glycans on FcγRIII, preventing interaction between Fc and the receptor (Mizushima et al., 2011, Ferrara et al., 2011).
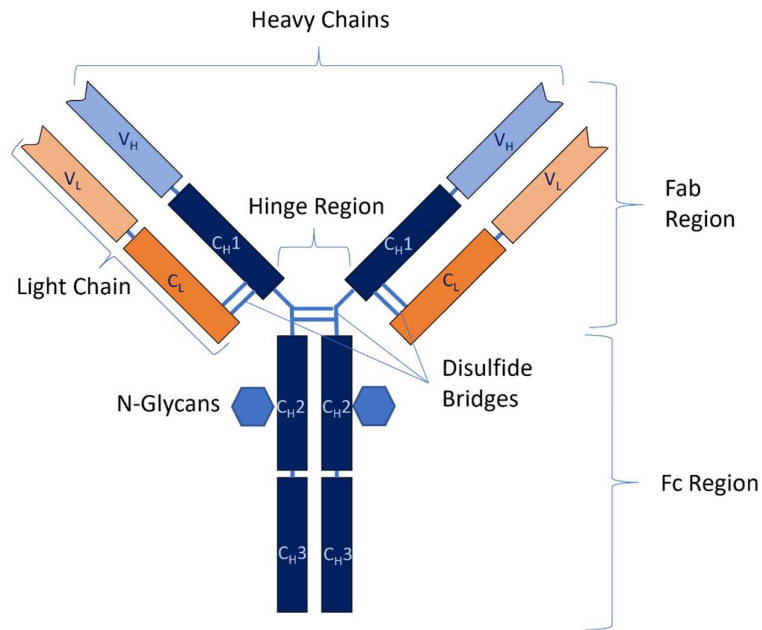
Figure 1.7 | Structure of Immunoglobulin G (IgG) antibody with N-*glycosylation sites labelled. IgG antibodies consist of pairs of heavy and light chains joined by disulphide bridges. Heavy chains consist of three constant regions (C_H) and a variable region (V_H). Light chains consist of a single constant region (C_L) and a variable region (C_V). The tail end of an IgG antibody (below the hinge region), consisting of constant heavy chains, C_H2 and C_H3, is called the fragment crystallisable (Fc) region, which binds to cell surface Fc receptors. The remainder of the antibody is the antigen-binding fragment (Fab) region. Both C_H2 fragments have* N-*glycosylation consensus sequences, which are normally* N-*glycosylated.*

1.1.7.4 Adhesion and Migration

In addition to their role in mediating antibody effector function, *N*-glycans also contribute to other immune responses through their involvement in cell adhesion and migration (Marth and Grewal, 2008). For example, during the process of leukocyte rolling along the vascular wall, a cell-cell adhesion system operates between leukocytes and endothelial cells, involving glycan binding proteins called selectins (Rosen, 2004). P-, L-, and E-Selectins are well-studied members of the lectin family of proteins, which bind specific *N*-glycan motifs. P- and E-selectins bind sialyl-Lewis-X epitopes and L-selectin binds a 6-sulfo-sialyl-Lewis-X epitope (Marth and Grewal, 2008). Expression of selectins and of the glycoprotein carriers of selectin ligands is tightly regulated and restricted to certain tissues, but dysregulation of their expression can contribute to

inflammatory disorders or cancer (McEver, 1997). Targeting selectins or the *N*-glycan selectin ligands has therapeutic potential for treating a range of autoimmune and inflammatory diseases and limiting tumour growth and metastasis (Romano, 2005, Uchimura and Rosen, 2006, Woollard and Chin-Dusting, 2007, Kim et al., 1998).

Other cell surface glycoproteins, such as E-cadherins (Liwosz et al., 2006) and integrins (Isaji et al., 2009, Isaji et al., 2006) are heavily *N*-glycosylated and the types of glycans present on each can influence their interactions, affecting adhesion and migration (Isaji et al., 2004, Pinho et al., 2013). E-cadherin is a glycoprotein involved in cell-cell adhesion that has been linked to the metastasis of several types of cancer (Guilford et al., 1998, Richards et al., 1999, Shiozaki et al., 1996). One of the first observations implicating E-cadherin in cancer was that decreased expression correlated with poor prostate cancer prognosis (Giroldi and Schalken, 1993, Umbas et al., 1994, Umbas et al., 1992), leading to the hypothesis that cell-cell adhesion is important for limiting migration of cancer cells. This has led to studies of the effects of E-cadherin *N*-glycosylation in cancer cells. For example, treating mouse melanoma cells with MGAT3 to increase intersecting complex *N*-glycosylation on E-cadherin reduced metastasis due to enhanced cell-cell adhesion (Yoshimura et al., 1995, Yoshimura et al., 1996, Kitada et al., 2001). Conversely, overexpression of MGAT5 has been linked to impaired cell-cell adhesion (Pinho et al., 2013) and reducing its levels can suppress cancer metastasis in mice (Granovsky et al., 2000).

Overexpression of MGAT3 was later studied in the context of α5β1 integrin (Isaji et al., 2004). Integrins are transmembrane glycoproteins at the cell surface that interact with components of the extracellular matrix (ECM). They act as adhesion receptors, linking components of the cytoskeleton with ECM proteins, enabling cells to adhere to surfaces, spread and migrate. Integrin α5β1 is one of the best characterised integrins, containing 26 *N*-glycosylation sites in total (Gu et al., 2012). Isaji et al. found that introducing intersecting *N*-glycans by overexpressing MGAT3 in HeLa cells inhibited their ability to bind to fibronectin, which attenuated cell spreading and migration (Isaji et al., 2004). The same group later sequentially mutated all 26 glycosylation sites on α5β1 integrin in CHO cells and they found that abolishing *N*-glycosylation sites on the β-propeller domain of the α5 subunit caused it to remain associated with calnexin in

the ER (Isaji et al., 2006). This blocked correct folding and heterodimerisation, preventing expression of α5β1 integrin at the cell surface, resulting in a decrease in CHO cell spreading and migration (Isaji et al., 2006). Integrins also interact with actin filaments to play an important role in CHO cell adhesion, and adaptation of adherent CHO cells to suspension culture has been shown to result in a rearrangement of integrins on the cell surface, coupled with a reorganisation of the actin cytoskeleton (Walther et al., 2016). However, the importance of integrin glycosylation in this process has yet to be studied.

1.2 Membrane Trafficking through the Mammalian Secretory Pathway

The aim of the previous section was to introduce the subject of glycobiology; painting a picture of the intricacies of the *N*-glycan biosynthetic pathway and presenting some of the many crucial biological roles played by *N*-glycans. Given the complexity and significance of *N*-glycosylation, it must be tightly controlled to ensure that glycoprotein structure and function are maintained. In the absence of a molecular template, understanding how *N*-glycan biosynthesis is regulated has been one of biology's greatest enigmas. Although knowledge on the subject has advanced significantly in the last half-century it is not yet possible to predict glycosylation with total accuracy, and science is lacking a complete picture of all the regulatory processes involved. However, it is clear that maintaining levels of glycan processing enzymes and their non-uniform distribution throughout the Golgi apparatus (Rabouille et al., 1995) is one of the major factors controlling *N*-glycosylation (Stanley, 2011). Membrane trafficking through the secretory pathway is critical for transport of both glycoproteins and of glycan processing enzymes. Therefore, the aim of the following sections is to summarise current understanding of membrane trafficking and Golgi organisation in the context of coordinating *N*-glycosylation.

1.2.1 The Mammalian Secretory Pathway

Glycan biosynthesis takes place predominantly in the ER and Golgi apparatus (figure 1.4), both membrane-bound organelles of the mammalian secretory pathway (figure 1.8). The secretory pathway begins in the rough ER (RER), which is an interconnected network of flattened sacs enclosed by a lipid-bilayer. Protein-synthesising ribosomes decorate the surface of the RER, where they deliver growing polypeptide chains through membrane-spanning translocon protein complexes into the RER lumen, where the first stages of *N*-glycosylation usually occur co-translationally.
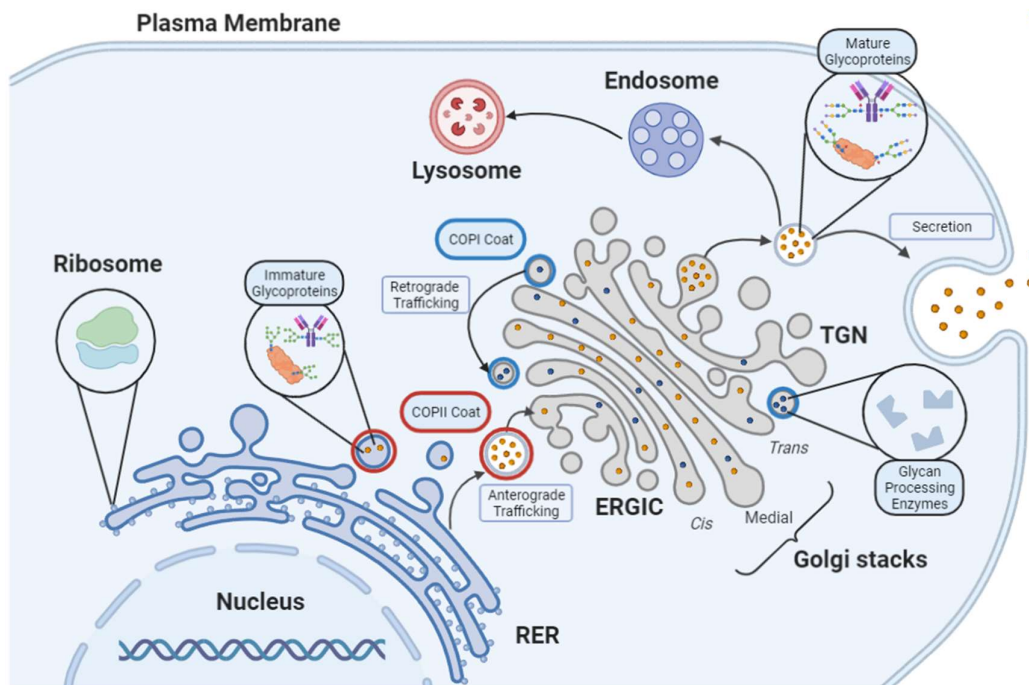
Figure 1.8 | Mammalian Secretory Pathway. A simplified schematic representing different pathways of the mammalian secretory pathway. Proteins are trafficked in the anterograde direction from the ER to the Golgi, via ERGIC, in vesicles coated with COPII protein, before entering the TGN for sorting to their required destination. Some proteins are secreted, whilst others are trafficked to endosomes or lysosomes. Glycan processing enzymes are packaged in COPI coated vesicles and can be trafficked in the retrograde direction from trans *to cis-Golgi. This image was created using BioRender.*

Following protein folding and the initial steps of *N*-glycan biosynthesis, glycoproteins are transported through the secretory pathway by the process of membrane trafficking. This involves enclosing secretory cargo, including glycoproteins, within lipid membrane-bound vesicles, which transport their contents to their required destination. Membrane trafficking can occur in both directions: anterograde trafficking flows from the ER, through the Golgi and onwards, whilst retrograde trafficking flows in the reverse direction (figure 1.8) and retains proteins in the Golgi or returns them to the ER. Anterograde and retrograde trafficking vesicles can be distinguished based on the type of coatamer protein (COP) that surrounds each vesicle. Golgi resident proteins, such as glycan-processing enzymes, can be trafficked in the retrograde direction in vesicles coated in COPI. In the anterograde trafficking pathway, after exiting the ER, COPII coated vesicles deliver glycoproteins to the Golgi apparatus via the ER-Golgi intermediate compartment (ERGIC).

Like the ER, the Golgi consists of an interconnected network of flattened sacs, known as cisternae. These are stacked and generally categorised as *cis* or-*trans* based on functional proximity to the ER or plasma membrane respectively (Farquhar and Palade, 1981). This compartmentalisation of the Golgi enables the localisation and distribution of glycan processing enzymes for successive stages of *N*-glycan processing. Anterograde trafficking through the Golgi culminates at the *trans*-Golgi network, where glycoproteins are sorted to their next destinations, typically, the plasma membrane or the endolysosomal system (figure 1.8).

1.2.2 Models of Trafficking through the Golgi apparatus

Textbook diagrams of the Golgi apparatus tend to depict Golgi ministacks. However, in reality, this is a simplification and the Golgi typically only forms these stacked structures during the process of mitosis (Lucocq and Warren, 1987). In non-mitotic mammalian cells, the Golgi exists as a perinuclear ribbon-like structure, formed from interconnected, adjacent mini-stacks (Wang and Seemann, 2011). This ribbon-like network structure was first observed in the 1890s by the late Italian biologist, Camillo Golgi, who later won the 1906 Nobel prize for his work on developing a staining procedure specific to the organelle that would later bear his name (Dröscher, 1998). Trafficking from the ER to the Golgi and formation of the Golgi ribbon is dependent on

a network of microtubules (Wei and Seemann, 2017). De-polymerisation of microtubules disperses the ribbon structure into mini-stacks naturally during mitosis, but this effect can also be replicated by treatment with the drug Nocodazole, which interferes with microtubule polymerisation (Minin, 1997).

1.1.2.1 The Anterograde Vesicular Shuttling Model

Over the years, several models of protein transport through the Golgi have been proposed (Glick and Luini, 2011). The most well studied models are the anterograde vesicular shuttle (Dunphy et al., 1981, Dunphy and Rothman, 1983) and cisternal maturation (Glick et al., 1997) models. The anterograde vesicular shuttle model hypothesised that the Golgi cisternae are stable structures with distinct enzyme compositions (Dunphy and Rothman, 1983). This model postulates that *N*-glycan processing involves trafficking of glycoproteins to the Golgi in COPII coated vesicles, where glycoproteins are successively processed by enzymes in each cisterna and shuttled in COPII vesicles to the next cisterna (Rothman and Orci, 1990). Dunphy and Rothman were pioneers in the development of this model, which they supported by a series of experiments showing that *N*-linked glycosylation events could be separated into early and late events towards the *cis* and *trans* Golgi faces respectively (Dunphy et al., 1981, Dunphy and Rothman, 1983).

1.1.2.2 The Cisternal Maturation Model

Key pieces of contradictory evidence later shed doubt on the anterograde vesicular shuttle model of protein transport through the Golgi. In 1995, Rabouille and colleagues showed that Golgi enzymes are present in a gradient along the cisternae, as opposed to each cisterna having distinct, predefined enzyme compositions (Rabouille et al., 1995). Another observation that did not fit the anterograde vesicular model was that the secretory protein procollagen is too large for trafficking in COP coated vesicles, yet it is able to transit through the Golgi without leaving the cisternal lumen (Bonfanti et al., 1998). Therefore, an alternative cisternal maturation model was proposed (Glick et al., 1997) (figure 1.9), which is now the more widely accepted model.
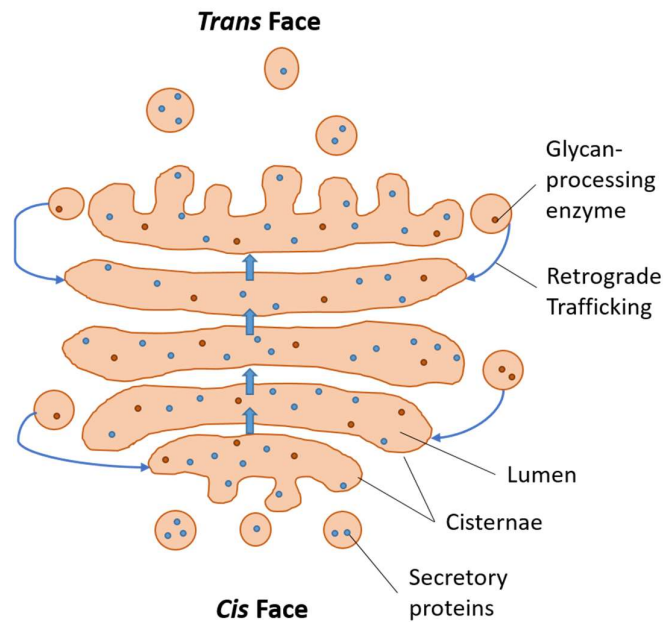
Figure 1.9 | Cisternal Maturation Model of Vesicle Trafficking at the Golgi Apparatus. The Golgi apparatus consists of a series of flattened, stacked, membrane-bound sacks known as cisternae, which have a cis face close to the endoplasmic reticulum (ER) and a trans face close to the plasma membrane. In the cisternal maturation model of vesicle trafficking, proteins are packaged into lipid bound COPII coated vesicles in the ER, which then fuse to form the cis cisternae of the Golgi. Trafficking of secretory proteins occurs through passive movement of the cisternae, which mature to eventually become the trans cisternae through acquisition of enzymes present in COPI coated retrograde trafficking vesicles. In the trans-Golgi network (TGN), proteins are sorted for secretion and new vesicles form to carry specific proteins either to their locations in the cell or to the plasma membrane for secretion. At the same time, Golgi-specific enzymes and trafficking proteins are recycled back towards the cis face in retrograde vesicles (Glick et al., 1997).

Cisternal maturation involves three key steps: the formation of a new *cis*-cisterna by fusion of COPII coated vesicles derived from the ER, dispersion of the old *trans*-cisterna by vesicle formation at the *trans*-Golgi network (TGN), and the recycling of Golgi-resident proteins by retrograde trafficking in COPI coated vesicles (Glick et al., 1997) (figure 1.8, figure 1.9).

1.2.3 Vesicle tethering

Retrograde trafficking is essential to retain glycan processing enzymes within the Golgi against its constant turnover, and to maintain their non-uniform distribution required for normal glycosylation. It involves trafficking glycan processing and other Golgi-resident enzymes in COPI vesicles in a *trans-* to *cis-* direction back to target cisternae, where vesicles tether and fuse with the cisternal membrane to deliver their contents (Cottam and Ungar, 2012, Fisher and Ungar, 2016) (figure 1.10). Targeted vesicle tethering is critical for ensuring delivery of glycosylation enzymes to the right place.

Coordinated interactions between a few different types of tethering proteins ensure that targeting specificity is achieved (Fisher and Ungar, 2016). These proteins include coat proteins, such as COPI, coiled-coil tethers, small GTPases called Rabs, sec1/munc19 (SM) proteins and soluble N-ethylmaleimide–sensitive factor (NSF) attachment protein receptors (SNAREs) and multi-subunit tethering complexes (MTCs) (figure 1.10) (Fisher and Ungar, 2016). In the first stage of retrograde trafficking, COPI coat proteins assemble around a section of membrane to facilitate vesicle budding from cisternae and the TGN, followed by recruitment of Rab proteins. Rabs switch between an inactive, GDP-bound conformation and an activated GTP-bound conformation, facilitated by guanine nucleotide exchange factors (GEFs). Activated Rabs recruit and interact with multiple proteins to drive targeted vesicle tethering (Hammer and Wu, 2002). Some of the first proteins to be recruited to COPI coated vesicles by Rabs are coiled-coil tethers, including the golgin family of proteins (Witkos and Lowe, 2015). Specific golgins are recruited by different Rab GTPases, for example, the golgin GM130 binds specifically to Rab30 (Sinka et al., 2008). Golgins interact directly with Rab proteins decorating retrograde vesicles to form the initial contact linking them to the cisternal membrane (Gillingham and Munro, 2016). Their elongated structures span up to 600 nm in length, making them ideal for long range capture of vesicles (Witkos and Lowe, 2017). It has also been shown that different types of golgins are localised to specific regions of the Golgi, adding another level of specificity to the process of vesicle tethering (Wong and Munro, 2014).
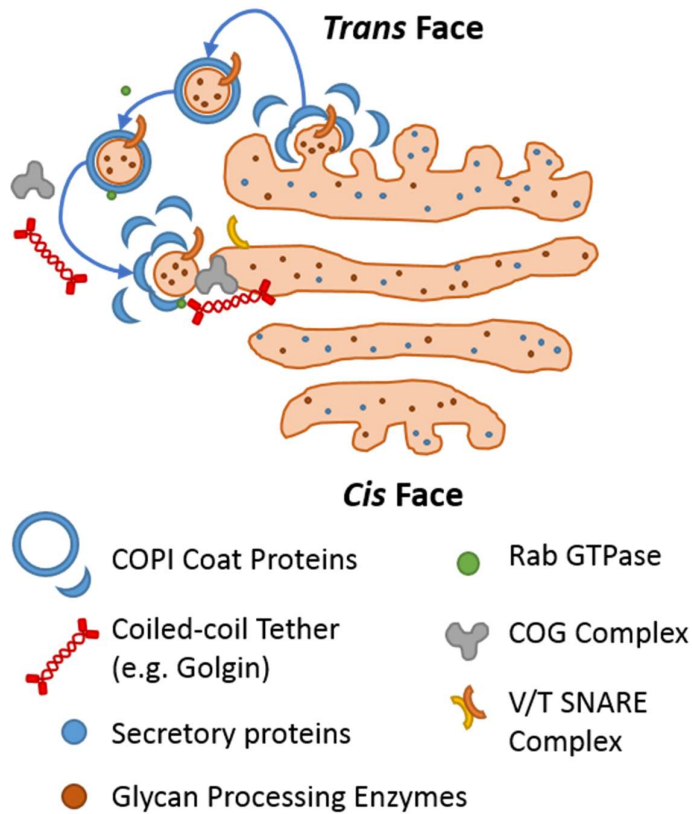
Figure 1.10 | Vesicle Transport and Tethering in the Golgi During Retrograde Trafficking. The steps involved in vesicle tethering are: (1) Budding: The cisternal membrane is deformed at the TGN by COPI coat proteins, causing a new vesicle to form and bud off from the cisterna. (2) Transport: The vesicle is trafficked in the retrograde direction to its target cisterna. An activated Rab GTPase is also recruited to the vesicle membrane, which helps target the vesicle to its specific cisterna. (3) Tethering: As the vesicle nears its target location, the protein coat is shed and the activated Rab GTPase recruits tethering proteins, including coiled-coils and the conserved oligomeric Golgi complex, which draw the membranes and SNAREs closer together. (4) Docking: The v- and t-SNAREs come into contact and form the trans-SNARE complex, which forces the membranes into contact. (5) Fusion: Formation of the trans-SNARE complex and contact between the membranes results in membrane fusion, which causes the vesicle protein content to be released into the cisternal lumen (Fisher and Ungar, 2016). Diagram adapted from Figure 1.1 in (Cottam and Ungar, 2012).

Another important class of molecules involved in tethering is the multi-subunit tethering complexes. Many MTCs involved in retrograde trafficking are classified as complexes associated with tethering containing helical rods (CATCHR) (Bröcker et al., 2010). Members of the CATCHR family include Dsl1p, which is involved in Golgi-ER trafficking (Andag and Schmitt, 2003) and the exocyst complex, which functions in tethering exocytic vesicles at the plasma membrane (TerBush et al., 1996); both of these were identified in yeast and have mammalian orthologues. Trafficking in mammalian cells involves the GARP tethering complex (Liewen et al., 2005), which also has a yeast orthologue (Conibear and Stevens, 2000), and acts in tethering endosome-derived vesicles at the TGN and the conserved oligomeric Golgi (COG) complex, which plays an important role in tethering of retrograde trafficking vesicles within the Golgi (Ungar et al., 2002). Subunits of MTCs interact with a range of Rabs, golgins and SNAREs to facilitate targeted vesicle tethering (Bröcker et al., 2010, Miller and Ungar, 2012, Fisher and Ungar, 2016, Cottam and Ungar, 2012).

The final stage of vesicle tethering involves SM (Aalto et al., 1992) and SNARE proteins (Söllner et al., 1993). Different subclasses of SNARE proteins exist, and each interacts with a range of tethers. A summary of these interactions can be found in (Sztul and Lupashin, 2009). Tethering between the vesicle and cisternal membranes by golgins brings SNARE proteins in close contact, enabling a *trans*-SNARE complex to form between the v-SNARE on the vesicle membrane and the t-SNARE on the target cisternal membrane (Jahn and Scheller, 2006). SNARES form a tight interaction, likened to a molecular zip, which helps dock the vesicle onto the cisternae. SM proteins are also known to be involved in SNARE complex formation as they interact both with individual SNAREs, predominantly from the syntaxin family (Misura et al., 2000), and with *trans*-SNARE complexes (Carr et al., 1999). Formation of the *trans*-SNARE complex forces the vesicle and cisternal membranes into contact, causing them to fuse and enabling the passage of vesicle contents into the cisterna (Weber et al., 1998). SNARE complex formation adds another level of specificity to vesicle trafficking as complexes can only form between cognate pairs of SNARE proteins (McNew et al., 2000).

1.2.4 Identification of tethering proteins in a cell-free system

Many of the proteins involved in membrane trafficking and vesicle tethering were identified following the development of a cell-free Golgi transport assay, pioneered by Fries and Rothman (Fries and Rothman, 1980). The assay involves studying the trafficking of VSVG (see section 1.1.7.1) *in vitro* between two cell extracts containing Golgi membranes (Fries and Rothman, 1980). Golgi membranes are extracted from two different CHO cell clones: 15B, a mutant cell line lacking MGAT1, which is infected with VSV and briefly incubated with $[^{35}S]$-methionine to facilitate labelling of VSVG in the ER, and WT, uninfected CHO cells. In 15B extracts, VSVG glycan-processing of oligomannose to complex glycans is prevented by the lack of MGAT1, making VSVG sensitive to digestion by an enzyme called endoglycosidase H (Endo-H), which can only cleave oligomannose and some hybrid glycans from peptide. When 15B extracts are incubated with WT extracts, the conversion of VSVG to an Endo-H resistant form, which shows no reduction in molecular weight upon Endo-H treatment, is indicative of MGAT1 activity enabled by transport between the Golgi compartments from the two extracts. Rothman et al. concluded that VSVG is trafficked from 15B to WT and MGAT1 in the WT extracts facilitates conversion of oligomannose glycans to Endo-H resistant complex type (Fries and Rothman, 1980). However, this conclusion has since been revised due to acceptance of the cisternal maturation model and it is now understood that the assay involves retrograde trafficking of MGAT1 from WT to 15B Golgi compartments.

Trafficking of MGAT1 cannot take place between the extracts alone, and requires the addition of cytosolic components, including adenosine triphosphate (ATP) (Fries and Rothman, 1980). The assay can therefore be used to identify cytosolic fractions capable of stimulating trafficking, thus guiding the purification of trafficking proteins. Proteins identified using this assay include NSF (Block et al., 1988), a protein involved in SNARE assembly and disassembly, soluble NSF attachment proteins (SNAPs) (Clary et al., 1990) and p115 (Waters et al., 1992), a protein involved in ER-Golgi transport and golgin binding (Alvarez et al., 2001). The identification of these proteins later led to the identification of SNAREs (Söllner et al., 1993), a key discovery in understanding the process of vesicle tethering and membrane fusion.

1.2.5 The Conserved Oligomeric Golgi Complex

The specificity of the vesicle tethering interactions between Rab GTPases, coiled-coil tethers, SM proteins and SNARE complexes drives the targeting of Golgi enzymes to their required destinations for normal glycosylation (Fisher and Ungar, 2016). Co-ordination of these interactions is likely achieved by a hetero-oligomeric multi-subunit tethering complex (MTC) known as the conserved oligomeric Golgi (COG) complex (Laufman et al., 2009, Ungar et al., 2002). COG interacts with a range of Rabs, golgins, SM proteins, SNAREs and the COPI coat protein, making it a hub of interacting tethering proteins (Ungar et al., 2006). COG consists of eight helix-containing subunits, termed COGs 1-8, which are arranged in a tentacular structure (figure 1.11) (Ungar et al., 2002, Ha et al., 2016). These subunits form two distinct lobes: COGs 1-4 form lobe A, which associates with the cisternal membrane, and COGs 5-8 form lobe B, associated with the vesicular membrane. An interaction between COG1 and COG8 bridges the lobes together (figure 1.11) (Ungar et al., 2002, Oka et al., 2005).
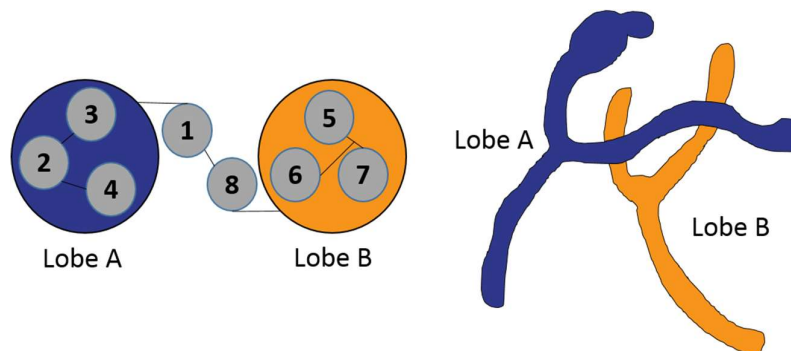
Figure 1.11 | The Conserved Oligomeric Golgi Complex. A schematic representation of the structure (on the left) shows the bi-lobe arrangement of the eight subunits, with lines between each subunit showing interactions. The diagram on the right represents one of the conformations of the COG complex, as identified by electron microscopy. Note that COG has a flexible structure so many other tentacular conformations exist.

1.2.5.1 Identification of the COG complex

COG is a member of the CATCHR family of proteins and was first characterised following multiple separate discoveries of MTCs in the mammalian Golgi. The first, originally termed the Golgi transport complex (GTC), was purified from cow brains after its

identification in a cell-free assay to stimulate Golgi transport (Walter et al., 1998). The authors modified the original assay used by Rothman *et al.* (Fries and Rothman, 1980), by including the previously identified trafficking proteins, NSF, αSNAP and P115 (see section 1.2.4). They found that an additional transport-stimulating protein found in cytosol from cow brains was required to facilitate VSVG trafficking, and they used the assay to guide its purification. They were able to identify that this protein consisted of at least five subunits (Walter et al., 1998).

Another MTC, named ldlCp, was later identified during analysis of two strains of mutant CHO cells with defective low-density lipoprotein receptors (LDLRs), termed ldlB and ldlC cells (Kingsley et al., 1986). Defective LDLRs in these mutants were found to be caused by multiple defects in Golgi-localised reactions. These defects were corrected by introducing a sequence of human cDNA encoding a protein the authors named LDLC protein (ldlCp), which was found, through immunofluorescence staining, to be associated with the Golgi (Podos et al., 1994). Introducing the protein into ldlB cells revealed that both ldlC and ldlB were required for the ldlC to associate with the Golgi (Chatterton et al., 1999, Podos et al., 1994).

A few years later, a human orthologue of the Golgi-localised yeast sec34p tethering protein complex was identified (Suvorova et al., 2001). The yeast sec34/sec35 complex (Whyte and Munro, 2001) is believed to act as a retrograde trafficking tether based on interactions with COPI, Rab GTPases and SNARE proteins (Suvorova et al., 2002). Mutations in sec34/sec35 cause glycosylation defects, reflecting its role in recycling Golgi-resident proteins (Suvorova et al., 2002). Sequence homology between the human orthologue and the yeast complex therefore pointed towards the human complex being involved in vesicle tethering during retrograde trafficking (Suvorova et al., 2001). Another purification and characterisation experiment revealed that the human complex was identical to both the GTC and ldlCp (Ungar et al., 2002) and that it consisted of eight subunits arranged in two lobes (Ungar et al., 2005).

1.2.5.2 Characterisation of the COG complex

Following the revelation that the aforementioned complexes were all identical, the complex was renamed COG (Ungar et al., 2002). Mutagenesis experiments by Oka, Ungar and colleagues later led to the identification of a subset of COG-dependent, Golgi

localised type II transmembrane proteins, termed GEARS, whose expression levels are reduced in COG subunit mutant cells (Oka et al., 2004). These proteins include: the golgins CASP, golgin84 and giantin; the glycosidase enzyme mannosidase II; the phosphorylated glycoprotein GPP130, which is involved in manganese sensing (Mukhopadhyay et al., 2010); and the SNARE proteins GOS-28 and GS15 (Oka et al., 2004). The authors concluded that reduced levels of the GEARS may have been due to mis-localisation in the ER as a result of retrograde trafficking and degradation by proteasomes. The results of this experiment suggested that COG not only directs intra-Golgi transport of GEARS, but that it is also heavily involved in retrograde trafficking in general (Oka et al., 2004, Oka et al., 2005).

Lupashin and colleagues have also worked extensively to characterise COG complex function (Climer et al., 2018, Shestakova et al., 2006, Smith et al., 2009, Willett et al., 2014, Zolov and Lupashin, 2005). Some of their more recent findings, based on immunofluorescence experiments, have revealed that COG lobes rarely detach from vesicular or cisternal membranes (Willett et al., 2014) and that membrane detachment is not essential for COG complex function (Climer et al., 2018). This has led to the most recent model of the COG complex's mode of action, whereby COG lobes A and B remain associated with cisternal and vesicular membranes respectively as they traverse the stack in both anterograde and retrograde directions (Climer et al., 2018).

1.2.6 Congenital disorders of glycosylation

Mutations in COG subunits can result in serious consequences for glycan processing (Freeze et al., 2014). This is exemplified in patients with congenital disorders of glycosylation (CDGs) caused by COG mutations (Foulquier et al., 2006, Wu et al., 2004), exhibiting a range of pleiotropic symptoms as a result of altered glycosylation (Jaeken, 2011, Leroy, 2006). For example, CDG-associated COG8 deficiency causes symptoms including seizures and failure to thrive, due to a sialylation deficiency (Kranz et al., 2007). On a molecular level, COG8-deficient CDG cells exhibit impaired SNARE complex formation (Laufman et al., 2013a). Another CDG, caused by an R729W mutation in COG4 resulted in a disrupted Golgi structure and a reduction in galactosylation and sialylation (Reynders et al., 2009). A similar effect is observed in COG7 CDG patients, but to a much greater extent than in COG4-CDG patients, resulting in lethality (Wu et

al., 2004). All COG CDG mutations result in destabilisation of the COG complex, however levels of other subunit expression depend on which lobe the mutation is found in (Foulquier, 2009).

So far, CDG cases have been reported from mutations in seven out of the eight COG subunits, with COG3 being the exception, suggesting that a COG3 mutation would be embryonic lethal. Indeed, other lobe A CDGs are hypothesised to cause more severe phenotypes than lobe B (Haijes et al., 2018). This theory is supported by the observation that CDGs caused by bi-allelic truncating mutations in lobe A subunits are observed far less frequently than those in lobe B, suggesting that those in lobe A are normally non-viable (Haijes et al., 2018). Mutations found in lobe A CDGs tend to be less detrimental to the subunit; typically involving small changes as opposed to large truncations. Furthermore, knock-down of lobe A subunits in yeast (Ram et al., 2002) or mammalian cells (Peanne et al., 2010) causes a more severe phenotype than lobe-B subunit knock-down (Haijes et al., 2018).

In each CDG case, different COG subunit mutations alter glycosylation in different ways, highlighting the different roles played by each subunit and their importance for glycosylation. Since COG plays such a critical role in retrograde Golgi trafficking, the existence of COG-CDGs also emphasises the significance of glycosylation enzyme sorting for controlling the cellular glycan profile.

1.3 Glycoprotein Biotherapeutics

In addition to their role in disease, there is now growing interest in the field of glycobiology from the pharmaceutical industry due to the increasingly evident potential of glycoprotein therapeutics (for simplicity, these will be referred to as biologics or biotherapeutics but note that other non-protein classes of biological therapeutics do exist, though these are beyond the scope of this thesis). Biologics include monoclonal antibodies, hormones, enzymes, cytokines and other glycoproteins. In many diseases, particularly cancers and autoimmune diseases, where pathogenesis involves aberration of naturally occurring processes in the host body as opposed to infection, the use of proteins as therapeutics can be advantageous over chemically synthesised drugs. Unlike many small-molecule synthetic drugs, proteins

can interact with host biomolecules in a manner that closely resembles a natural process. Biotherapeutics therefore offer exquisite biocompatibility and specificity to their natural targets in comparison to alternative treatments, such as radio- or chemotherapy, that can be toxic to healthy tissue as well as cancerous or diseased tissues.

The aims of the following sections are to introduce a few key classes of biologics and some examples of each, with a particular focus on monoclonal antibodies and their mechanism of action. Biologic production methods and challenges arising from glycan heterogeneity will also be introduced.

1.3.1 Hormones, Growth Factors and Cytokines

1.3.1.1 Erythropoeitin

One of the most successful biologics to date is the cytokine hormone, erythropoietin (EPO). In the human body, it is normally secreted by kidneys in response to cellular hypoxia to stimulate the production of red blood cells. Constant turnover of red blood cells means it is constitutively secreted at a low level (reviewed by Goldwasser, 1976, Jelkmann, 2007). Abnormal levels of hypoxia, associated with anaemia or chronic lung disease, result in elevated levels of EPO. However, patients suffering with kidney disease are deficient in EPO secretion and more likely to suffer from anaemia. Therefore, there was significant clinical interest in developing exogenously produced EPO for therapeutic intervention (Goldwasser, 1996). Purification of human EPO in 1977 by Goldwasser *et al.* (Miyake et al., 1977) enabled the amino acid sequence to be determined, leading to the identification of the gene and isolation of complementary DNA (cDNA). This cDNA was cloned into mammalian expression systems, yielding secreted EPO with biological activity (Lin et al., 1985, Jacobs et al., 1985). Less than two years later, phase I and II clinical trials were completed, demonstrating the efficacy of administering recombinantly produced EPO in the treatment of anaemic patients with end-stage renal failure (Eschbach et al., 1987). As well as its effective use in treating pathological hypoxia, EPO is commonly administered following blood loss to stimulate red blood cell production. EPO has also been exploited as a performance enhancing drug in sport, leading to the implementation of anti-doping controls for detecting elevated amounts of red blood cell markers (reviewed in Jelkmann and Lundby, 2011).

EPO is heavily glycosylated, with over 40% of its molecular weight comprising of carbohydrate at four glycosylation sites, of which three have predominantly tetra-antennary *N*-glycans (Sasaki et al., 1987, Sasaki et al., 1988). Removal of these glycans by endoglycosidase treatment abolished *in vivo* activity in mice, suggesting *N*-glycosylation is essential for EPO function, though this did not abolish its interaction with target cells *in vitro,* suggesting *N*-glycans were important for maintaining protein levels in the body (Dordal et al., 1985). Within these glycan structures, sialylation is a particularly important feature, due to the interaction of asialylated glycans with ASGPRs and their subsequent clearance (see section 1.1.7.2). The effect of removing sialic acid from EPO is startling: in rats, fully glycosylated exogenous EPO resides in blood plasma for up to four hours, whereas 96% of desialylated EPO is cleared within two minutes, partially by the action of ASGPRs in the liver but predominantly by filtration from the kidney, independently of ASGPRs (Spivak and Hogans, 1989).

1.3.1.2 Follicle Stimulating Hormone

The relationship between EPO glycosylation and serum half-life is a strong example of how glycans can influence physiology. Conversely, glycosylation of some glycoproteins can change in response to physiological conditions. Follicle stimulating hormone (FSH) is a sex hormone found in both males and females: in females it stimulates ovarian follicle development, an important step in ovulation, whereas in males it is critical for initiating spermatogenesis. The sex-dependent difference in function of FSH is reflected by differential glycosylation in male and female FSH. In pre-menopausal women, FSH has a lower sialic acid content due to decreased expression of sialyltransferase enzymes, resulting in lower serum half-life because of clearance by ASGPRs, and in all men and post-menopausal women, FSH sialic acid content, and therefore serum half-life, is higher (Ulloa-Aguirre et al., 1999).

Glycosylation of FSH is heavily influenced by levels of other sex hormones. In females, high oestrogen levels correlate with less acidic/sialylated FSH glycoforms, whereas the hypoestrogenic state is associated with increased FSH sialylation (Padmanabhan et al., 1988). In males, glycosylation of FSH is controlled by androgens, such as testosterone, with castration resulting in a decrease in FSH sialylation (reviewed in Campo et al.,

2019). Although serum half-life of the less sialylated FSH glycoforms is lower, biological activity is higher due to enhanced binding to FSH receptors (Zambrano et al., 1999).

FSH has been exploited for therapeutic use in females suffering infertility as an important component of *in vitro* fertility (IVF) treatment. Several commercial recombinantly expressed FSH products exist and their glycosylation is subtly different due to differences in production methods, resulting in differences in pharmacological properties (reviewed in Lunenfeld et al., 2019).

1.3.1.3 Bone morphogenic protein 2

Glycosylation of FSH is highly heterogenous and has been studied extensively (Bousfield and Harvey, 2019), with glycosylation being important for serum half-life, efficacy and therefore an important regulatory factor for the development of new FSH drugs (reviewed in Lunenfeld et al., 2019). In contrast, some clinically successful biologics have poorly understood glycosylation. Bone morphogenic protein 2 (Bmp2) is a cytokine and member of the transforming growth factor β (TGF-β) family of proteins, which is involved in osteoblast differentiation and the formation of new bone and cartilage (Wozney et al., 1988, Wozney, 1992). It is first synthesised as a large precursor protein with three glycosylation sites, one of which is essential for proper folding and secretion. This was demonstrated by mutagenesis of glycosylation sites, revealing that mutation at asparagine 135 sequesters Bmp2 precursor in the ER and prevents osteoblast differentiation (Hang et al., 2013). However, the functions of the remaining glycosylation sites and effects of heterogeneity on Bmp2 function are yet to be elucidated.

Recombinant human Bmp2 (rhBmp2) has been approved by the food and drug agency (FDA) for therapeutic use in the repair of broken bones (Khan and Lane, 2004) and in accelerating bone regeneration after bone grafting (Yamamoto et al., 2006). Treatment of open tibial fractures with a collagen sponge containing rhBmp2 was found to significantly improve the outcome of fracture repair, in terms of minimising infection risk, accelerating healing and reducing the risk of failure (Govender et al., 2002). Despite this success, rhBmp2 has side-effects including heterotopic ossification, which is formation of extra bone around the knee after tibial repair surgery (Boraiah et al., 2009). Multiple complications can also result from "off-label" use of rhBmp2 in

treatment of conditions not specifically mentioned in the FDA approval of rhBmp2 (reviewed by Poon et al., 2016). Therefore, more research is needed to understand whether glycosylation of rhBmp2 could influence these side-effects.

1.3.2 Monoclonal Antibodies

Antibodies are defensive molecules produced by B-lymphocytes, which are a class of white blood cells involved in adaptive immunity. Monoclonal antibodies (mAbs) are defined as antibodies made by clones of a single white blood cell. Typically, the term is used to describe the IgG class of antibodies (figure 1.7), which have high affinity for their antigen and are the most common immunoglobulin in serum. Due to their high affinity, mAbs have been exploited for multiple purposes, including immunoblotting and immunofluorescence in molecular and cell biology experiments, diagnostics and for therapeutic treatment of cancer and autoimmune diseases. This thesis focusses on therapeutic use of mAbs, so the following sections relate to mAb biologics.

It was estimated in 2017 that the global biologic market was worth $221 billion and of that, 43% came from mAbs alone (Reportlinker, 2018). Since their first production in continuous cell-culture (Kohler and Milstein, 1975), and further work to humanise them and prevent cross-species immunogenicity (Jones et al., 1986), they have revolutionised treatment of cancer and autoimmune disease.

1.3.3 mAb Function

1.3.3.1 Mechanism of action

IgG antibodies are highly effective molecules of the adaptive immunity because their antigen binding fragment (Fab) regions have huge structural diversity and so can bind to a multitude of antigens with high affinity (Berek and Milstein, 1988). Naturally, they are involved in fighting pathogens using several different mechanisms (figure 1.12). One mechanism is by binding directly to toxins (e.g. tetanus toxin (Ourth and MacDonald, 1977)) produced by microbes, preventing them from causing symptoms of disease. IgGs can also bind to antigens expressed on the surface of viruses, such as haemagglutinin, which results in prevention of the virus from infecting host cells (Knossow et al., 2002). Alternatively, antibodies function by a process called opsonisation, which means marking the pathogen for destruction by other cells of the immune system. IgGs bind the pathogenic antigen at their Fab region, leaving the Fc

region exposed for recognition by Fc receptors on cells capable of destroying the pathogen, such as phagocytes and natural killer cells. Phagocytes, such as macrophages and dendritic cells, destroy pathogens by the process of phagocytosis, a type of endocytosis in which the pathogen is internalised and digested inside a phagosome compartment. Non-phagocytic natural killer T-lymphocytes can also recognise opsonised pathogens or infected cells and destroy them by releasing cytotoxic granules that trigger apoptotic cell death. This process is called antibody-dependant cellular cytotoxicity (ADCC).
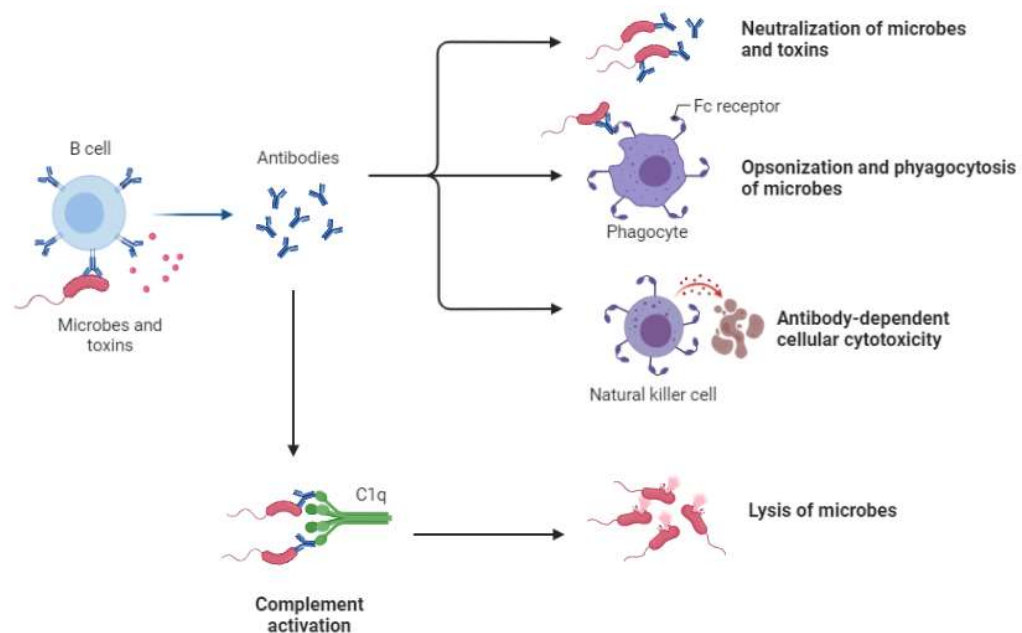
Figure 1.12 | Mechanisms of IgG function in the destruction of microbes. This figure was adapted from a template produced in BioRender.

IgG opsonisation can also trigger complement activation (figure 1.12), which is a type of innate immune response involving a small number soluble circulating "complement" proteins that recognise and bind to either a large number of broad pathogen-associated molecular patterns or to the Fc region of antigen-bound antibodies. Complement proteins circulate as precursor proteins, which become cleaved by proteases in response to various triggers, including Fc binding. The classical pathway involves complement protein C1q binding two IgG antigen-antibody complexes, triggering a cascade of complement protein cleavages, ultimately resulting in formation

of a membrane attack complex that lyses the pathogen or infected cell. This process is known as complement-dependent cytotoxicity (CDC). Antibody effector functions are well-understood and reviewed in (Forthal, 2014).

Naturally, circulating antibodies target pathogens and infected cells, whilst regulation prevents long-lasting formation of auto-reactive antibodies. Developing mAbs to treat autoimmune diseases and cancers where the antigen is derived from a host-cell therefore requires immunisation of a different species with the human antigen, as humans are generally incapable of producing large quantities of antibodies against their own antigens. The resulting antibodies raised against the human antigen can then be "humanised" to prevent cross-species immunoreactivity. This involves cloning the gene encoding the variable domains of the anti-human Fab region into the gene encoding the constant region of a human-produced IgG (Jones et al., 1986). The resulting DNA can then be used for recombinant production of humanised mAbs with reduced chance of rejection by the body.

Humira (adalimumab), the first FDA approved humanised mAb, targets tumour necrosis factor alpha (TNFα) to treat rheumatoid arthritis (Weinblatt et al., 2003) and Crohn's disease (Colombel et al., 2007), and is now the world's best-selling pharmaceutical (Philippidis, 2018). TNFα is a cytokine secreted by macrophages to recruit other immune cells as part of an inflammatory immune response to infection. However, in Crohn's (Van Deventer, 1997) and rheumatoid arthritis, TNFα secretion is elevated, triggering an inflammatory immune response in the intestines and joints respectively. Humira binds to TNFα and prevents the inflammatory immune response, thereby functioning in a similar way to a natural IgG molecule neutralising a toxin (figure 1.12), but in this case the "toxin" is aberrantly expressed TNFα.

Another example, Herceptin, which is a humanised mAb against the HER2 receptor, is an effective breast cancer treatment in patients overexpressing HER2 (Piccart-Gebhart et al., 2005). HER2 is a member of the receptor tyrosine kinase (RTK) protein family (reviewed in Lemmon and Schlessinger, 2010). Specifically, it is a member of a smaller RTK sub-family of human epidermal growth factors (HER) that mediate growth, survival, and differentiation of cells. Certain aggressive breast cancers with poor prognosis involve extreme over-expression of the HER2 receptor and rapid

proliferation of HER2 overexpressing cancer cells (Slamon et al., 1987). Herceptin binds to the HER2 receptor and is thought to function by several different mechanisms, including preventing dimerisation of HER2 receptors, causing inhibition of the signalling cascade they normally induce to promote proliferation, and opsonisation of the HER2 receptor, resulting in destruction of the cancerous cells by ADCC (Valabrega et al., 2007).

1.3.3.2 *N*-glycosylation of mAb Fc regions

Most mAbs are *N*-glycosylated on each of their heavy chain Fc regions at Asn297, and the types of glycans present can influence properties such as their serum half-life, efficacy and immunogenicity (reviewed in Higel et al., 2016). Fortunately, unlike whole-cell glycan profiles, the number of glycans typically found on leukocyte derived IgGs and mAbs recombinantly expressed in mammalian host-cells is largely limited and glycans tend to be much smaller than those found on other glycoproteins. This is because Asn297 glycosylation sites on each C_H2 domain are buried inside the protein structure, meaning larger *N*-glycans and those with negative charges, such as sialic acid, are sterically disfavoured and less common (Flynn et al., 2010). As a result, most IgG *N*-glycans are bi-antennary and further addition of GlcNAc residues tends to be intersecting. Furthermore, most biologic glycans are complex, with hybrid and oligomannose glycans seen less frequently, and of the oligomannose glycans, typically only $GlcNAc_2Man_5$ is observed. A summary of the most common glycans found on the Fc region of leukocyte derived human IgGs is shown in table 1.2, based on analysis by (Stadlmann et al., 2008). These glycan structures are also found on mAbs recombinantly expressed in mammalian host cells, though intersecting GlcNAc residues are seen less commonly than on recombinant mAbs. Certain mammalian expression systems can also produce Fc glycans not found in humans, such as terminal Galα-1,3Gal disaccharides on mAbs expressed in mouse cells (Stadlmann et al., 2008). Although the repertoire of Fc *N*-glycans is limited, each C_H2 domain can contain a different glycan, giving rise to further heterogeneity.

Table 1.2 | Common N-glycan structures on Fc regions of leukocyte derived human IgGs. For simplicity, oligomannose glycans are grouped together and the names and structures are intended to represent glycans with between 5-9 mannose residues attached to the chitobiose region of the core. Bisecting glycans are grouped together too. Structures are based on the most common structures found on human IgGs analysed by (Stadlmann et al., 2008).

Structural Name/s	Oxford Name/s	Structure/s
$GlcNAc_2Man_5$- $GlcNAc_2Man_9$	M5-M9	
$GlcNAc_2Man_5GlcNAc_1$, $FucGlcNAc_2Man_5GlcNAc_1$	M5 + GlcNAc, M5F + GlcNAc	
$GlcNAc_2Man_3GlcNAc_1$, $FucGlcNAc_2Man_3GlcNAc_1$,	A2G0-GlcNAc, FA2G0-GlcNAc	
$GlcNAc_2Man_3GlcNAc_2$, $FucGlcNAc_2Man_3GlcNAc_2$	A2G0, FA2G0	

GlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_1$, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_1$	A2G1, FA2G1	
GlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_1$NeuAc1, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_1$NeuAc1	A2G1S1, FA2G1S1	
GlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$	A2G2, FA2G2	
GlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$NeuAc$_1$, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$ NeuAc$_1$	A2G2S1, FA2G2S1	
GlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$NeuAc$_2$, FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal$_2$ NeuAc$_2$	A2G2S2, FA2G2S2	

GlcNAc$_2$Man$_3$GlcNAc$_3$, GlcNAc$_2$Man$_3$GlcNAc$_3$Gal$_2$NeuAc$_2$ (any of the above from A2G0 with bisecting GlcNAc and afucosylated core)	A2BG0- A2BG2S2	
FucGlcNAc$_2$Man$_3$GlcNAc$_3$, FucGlcNAc$_2$Man$_3$GlcNAc$_3$Gal$_2$NeuAc$_2$ (any of the above from A2G0 with bisecting GlcNAc and fucosylated core)	FA2BG0- FA2BG2S2	

1.3.3.3 Influence of Fc *N*-glycosylation on mAb properties and functions

Some examples of the influence of Fc *N*-glycan type on mAb properties were introduced in section 1.1.7.3 and 1.1.7.4, and they include high mannose glycans reducing serum half-life through clearance by mannose receptors (Goetze et al., 2011, Alessandri et al., 2012), and lack of fucose enhancing ADCC (Shinkawa et al., 2003, Shields et al., 2002). These functions of IgG Fc glycosylation have been well documented and generally, literature agrees that high mannose glycans are detrimental to mAb longevity and that core fucosylation is an important mediator of Fc receptor binding (Liu, 2015). It has also been well-documented that sialylation confers anti-inflammatory properties through enhancing binding to the anti-inflammatory Fc receptor, FcγRIIB (Kaneko et al., 2006, Lin et al., 2015). However, no consensus has yet been reached for the effects of terminal galactosylation on some aspects of mAb function. Studies either report that terminal galactosylation increases binding to pro-inflammatory FcγRIII receptors and enhances ADCC (Kurogochi et al., 2015), no change to ADCC in response to galactosylation (Wright and Morrison, 1998, Peschke et al., 2017, Shinkawa et al., 2003) or downregulation of ADCC (Nimmerjahn et al., 2007). One study also showed that terminal galactosylation enhanced binding to C1q and therefore increased CDC activity (Peschke et al., 2017).

The effects of complete de-glycosylation of IgG have also been evaluated. Studies show that glycosylation confers resistance to proteolytic degradation and increased IgG stability (Zheng et al., 2011) compared to de-glycosylated IgGs (Mimura et al., 2000). Enzymatically de-glycosylated IgGs are both less thermally stable and less resistant to proteolysis, whereas IgGs with high mannose Fc glycosylation were found to be less thermally stable but more resistant to proteolytic degradation (Zheng et al., 2014). Another study used genetic engineering to generate a de-glycosylated mouse monoclonal IgG by mutating the Asn297 *N*-glycosylation site to Ala, which was administered to rats resulting in significantly shorter biological half-life in rat serum compared to the WT IgG (Wawrzynczak et al., 1992). This effect can be partially explained by the observation that glycosylation stabilises the C_H2 domain of the IgG Fc region (Krapp et al., 2003).

1.3.4 Glycoengineering for Biologic Production

Considering the influence of *N*-glycans on protein properties, *N*-glycosylation is classed as a critical quality attribute in biologic production, and this is particularly important for therapeutic mAbs. Unlike synthetic chemical drugs, mAbs can't be synthesised reliably *in vitro*. The complexities of protein synthesis and folding are such that *in vitro* methods are unable to replicate the efficacy of protein biosynthesis machinery found in living cells and even efforts to develop cell-free protein synthesis techniques rely on using crude-cell lysates (Dondapati et al., 2020). Therefore, the vast majority of biologics are synthesised in living cells, and *N*-glycosylation by host-cell machinery occurs during their synthesis. For mAbs, CHO cells are the most common expression system; as mammalian cells, their glycosylation machinery closely matches that found in humans, meaning cross-species immunogenicity is limited (Brooks, 2004). However, glycan heterogeneity still poses a challenge, particularly as glycosylation can be influenced by even small changes to cell culture conditions (Tejwani et al., 2018). Glycoengineering, the controlled biosynthesis of distinct, homogeneous glycan structures, has therefore become attractive to the biopharmaceutical industry (Higel et al., 2016).

1.3.4.1 Bioprocess Control of Recombinant Protein *N*-Glycosylation

Cell culture medium composition is known to influence glycosylation of recombinantly expressed proteins, so numerous studies have focussed on understanding the impact of culture conditions and medium additives on glycosylation (Gramer et al., 2011, Prabhu et al., 2018, Kildegaard et al., 2016). For example, supplementing culture medium with galactose has been shown to increase galactosylated IgG (Kildegaard et al., 2016, Zhang et al., 2019). Adding kifunensine, an inhibitor of α1,2-mannosidase I, can also increase high-mannose glycan species (Ehret et al., 2019). Other culture conditions, such as pH and temperature, can also influence glycosylation (recently reviewed in Edwards et al., 2022). In particular, low pH has often been associated with enhanced production of galactosylated and sialylated mAb Fc *N*-glycans, though productivity decreases in more acidic conditions (Aghamohseni et al., 2014).

Understanding the impact of cell culture conditions on recombinant protein glycosylation is important for developing in-process glycosylation control methods (Zupke et al., 2015, Hossler et al., 2009) and for developing computational models for predicting glycosylation based on culture conditions (e.g. Zhang et al., 2021). In-process glycosylation control involves monitoring mAb glycosylation in as close to real-time as possible whilst measuring cell culture conditions, such as temperature, pH and medium composition, in parallel to enable computational mathematical modelling of the factors influencing glycosylation and adjustment of conditions if necessary (Hossler et al., 2009). Zupke et al. used this approach specifically to monitor high mannose glycans, enabling them to adjust medium mannose concentration to achieve the optimum conditions for growth with minimal oligomannose-type Fc glycosylation (Zupke et al., 2015).

1.3.4.2 Engineering of Host Cells for Production of Glycan-Engineered Biologics

A complementary or alternative approach to glycoengineering involves targeting the host cell line. Methods to genetically engineer mAbs with distinct glycoforms often focus on altering levels of specific glycan-processing enzymes within the host cell. Short-term studies examining the effects of targeting specific host cell proteins on biologic glycosylation commonly exploit non-coding RNAs, such as short interfering RNA (siRNA), for temporarily knocking down or overexpressing specific glycan-

processing proteins (Reviewed in Amann et al., 2019). For example, a short-term study was conducted to examine the effects of knocking down FUT8 on IgG production in CHO cells, using siRNA (Mori et al., 2004). This study demonstrated that 60% defucosylation of mAbs could increase ADCC activity over 100-fold over that of 10% defucosylated mAbs (Mori et al., 2004). SiRNA and short-hairpin RNA (shRNA) has also been used to knock down sialidase enzymes in CHO cells, resulting in a 98% reduction in sialidase function and around a 30% increase in sialic acid content of recombinant human interferon gamma (rhIFNγ), with potential for increased serum half-life due to reduced clearance by ASGPRs (Zhang et al., 2010). Enhanced sialylation of recombinant protein has also been demonstrated for IgG expressed by CHO cells through transient co-expression of human α2,6-sialyltransferase 1 and β1,4-galactosyltransferase 1 (Raymond et al., 2015). In this study, the aim was to overexpress enzymes involved in α2-6 sialic acid formation to increase the presence of α2-6 sialic acids on IgG Fc *N*-glycans, promoting anti-inflammatory properties (Raymond et al., 2015). Using this method, the authors were able to successfully produce highly sialylated mAbs, demonstrating the potential to alter recombinant protein glycosylation by transient overexpression of glycan-processing enzymes (Raymond et al., 2015) .

Temporary gene knockdown and overexpression is a useful tool for studying the effects of altering glycan-processing enzyme levels on recombinant protein production, however permanent genetic modification is desirable for long-term production. Permanent genetic engineering of hosts for biologic production, including CHO cells, has been achieved using zinc-finger nuclease (ZFN)-mediated gene editing (Santiago et al., 2008). ZFNs are engineered fusion proteins that have a zinc finger protein DNA-binding domain, which specifically binds to a target DNA sequence, fused to an endonuclease catalytic domain called FokI (Urnov et al., 2005). Endonuclease activity of the engineered protein is dependent on dimerization of the FokI domain, so DNA binding-induced cleavage requires a pair of ZFNs designed to target DNA at precise relative locations that place the FokI domains in contact with each other (Urnov et al., 2005). ZFNs have been exploited to disrupt the Mgat1 gene in a high producing CHO K1 cell line, yielding recombinantly expressed IgG with predominantly $GlcNAc_2Man_5$ glycan in the Fc region, designed to interact with mannose receptors of target cells (Sealover et al., 2013). The relationship between Fc *N*-glycan core fucosylation and

ADCC has also been exploited by the pharmaceutical industry and complete afucosylation of mAbs has been achieved by ZFN-mediated knockout of the FUT8 gene in CHO cells, resulting in mAbs with enhanced ADCC activity (Malphettes et al., 2010, Yamane-Ohnuki et al., 2004).

A disadvantage to ZFN-mediated gene editing is that designing and producing the protein constructs is time-consuming, laborious, and very expensive. More recently, the development of a technique based on the clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated protein 9 (Cas9) adaptive immune system found in *Streptococcus pyogenes* has overcome some of these limitations (Doudna and Charpentier, 2014, Jinek et al., 2012). CRISPR/Cas9 has been developed as a two-component system with a synthetic single-guide RNA (sgRNA) and Cas9 nuclease, which can be used to generate gene knockouts, insertions and point mutations by simply applying it with an sgRNA specific to the gene of interest (Ran et al., 2013b, Doudna and Charpentier, 2014). Since its development, CRISPR/Cas9 has been used to knockout FUT8 in CHO cells, to produce afucosylated mAbs (Sun et al., 2015). More recently, CRISPR/Cas9 was utilsed to engineer CHO cells with simultaneously and independently controllable FUT8 and βGalT expression, under small-molecule inducible promoters (Chang et al., 2019). This enabled the production of mAbs with high galactosylation levels, which are desirable for enhanced CDC and low fucosylation for enhanced ADCC (Chang et al., 2019). These studies demonstrate some of the potential applications of CRISPR/Cas9 for glycoengineering of CHO cells for biologic production. Applications are more extensively reviewed in (Lee et al., 2015b, Donini et al., 2021).

1.4 Mass Spectrometric *N*-glycan profiling

To understand how changes in glycosylation machinery can alter the glycan profile of a cell or of a recombinantly expressed glycoprotein, mass spectrometry is a powerful tool (reviewed in O'Flaherty et al., 2017, Mulloy et al., 2015). Essentially, this technique involves ionising whole molecules and their fragments and analysing their mass:charge (*m/z*) ratios to establish their composition. Before detailed analysis can take place, there are many considerations. Firstly, samples must be prepared, which can mean purifying whole glycoprotein or releasing *N*-glycans from protein preparations or

whole-cell lysates. Secondly, derivatising *N*-glycans is necessary for quantification: typically, samples are either permethylated, or labelled with a fluorescent tag to enable fluorescent quantitation.

When it comes to performing mass spectrometry and selecting the type of mass spectrometer, there are also several considerations. Mass spectrometers basically consist of a sample inlet, an ionisation source, mass analyser(s) and ion detector(s). For *N*-glycan profiling of whole-cell lysates, matrix-assisted laser desorption/ionisation mass-spectrometry (MALDI-MS) and electrospray ionisation (ESI) are the most common ionisation methods, but MALDI-MS is favoured for yielding good quantitation (Wada et al., 2007). However, for industrial analysis of *N*-glycosylated biopharmaceuticals, ESI-MS instruments are more readily available and therefore used for routine analysis. Mass spectrometry is also frequently coupled with high-performance liquid chromatography (HPLC), to separate analytes prior to sample injection. Several types of mass analyser are also commonly used in *N*-glycan analysis. The work presented in this thesis involved the use of two different types of mass analysers: quadrupole and time-of-flight (TOF). The aim of the following sections is to introduce each stage of sample preparation and analysis and the types of MS used to gather the data presented in this thesis, and to show how each method has advanced the field of glycobiology.

1.4.1 Filter- Aided *N*-glycan Separation

Cell cultures are excellent tools for studying glycosylation as they can be easily manipulated under controlled conditions to aid in understanding how, for example, mutation (e.g. Abdul Rahman et al., 2014), differentiation (e.g. Wilson et al., 2016) or drug treatment (e.g. Wilson et al., 2018) affect the whole-cell glycan profile. Until recently, glycan profiling of adherent cell cultures using mass spectrometry required large sample volumes and up to hundreds of millions of cells (North et al., 2010, Bateman et al., 2010). However, a technique developed as part of an earlier collaboration between Prof. Dani Ungar, and Prof. Jane Thomas-Oates enabled *N*-glycan release and profiling of cell culture samples containing less than a million cells (Abdul Rahman et al., 2014). This technique, named filter-aided *N*-glycan separation (FANGS) (figure 1.13), has since been used on multiple cell types (Wilson et al., 2016),

including CHO cells, and further developed by Dr Kirsty Skeene and colleagues to additionally facilitate O-glycan release and protein recovery from the same cell sample, in the same pot (Skeene et al., 2017).
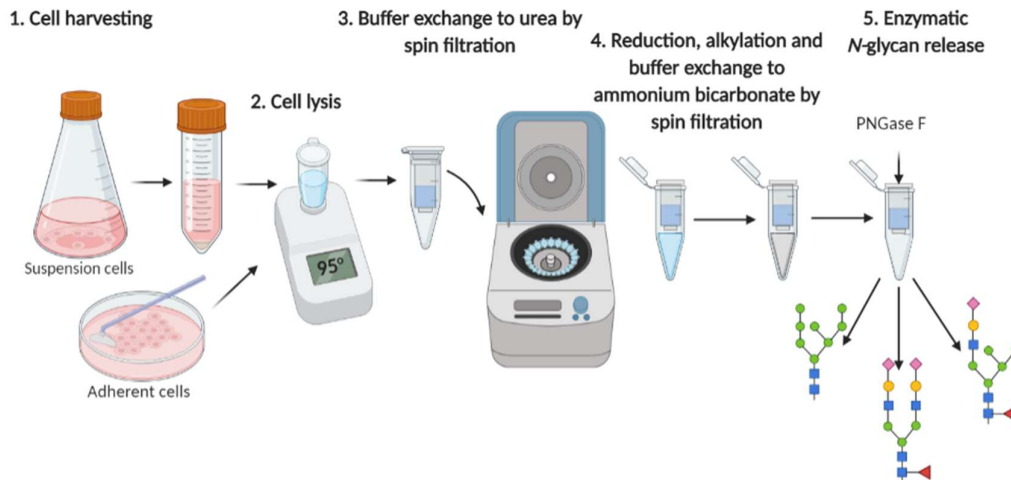
Figure 1.13 | Schematic representation of the FANGS procedure. Cells are harvested either by scraping (adherent) or centrifugation (suspension) and washed before cell lysis by heating in an SDS-containing buffer. SDS is removed by spin filtration and exchanged for a urea buffer prior to reduction and alkylation. The buffer is then exchanged again for ammonium bicarbonate and N-glycans are released by overnight treatment with PNGase F and recovered by washing through the filter membrane in water.

For almost all whole-cell *N*-glycan profiles presented in this thesis, the FANGS method was adopted for cell lysis and *N*-glycan release (figure 1.13). The first stage of FANGS is cell harvesting. For adherent cell cultures, these must first be washed to remove residual protein from foetal bovine serum (FBS) in the medium, before releasing the cells by scraping. For suspension cells, which grow in medium lacking FBS, cell harvesting is quicker as scraping isn't required and fewer washes are necessary. After washing, cells are pelleted and lysed by heating to 95°C in a lysis buffer containing the strong surfactant sodium dodecyl sulfate (SDS), which is highly effective for solubilisation of membrane glycoproteins from whole cells (Abdul Rahman et al., 2014).

The technique is named filter-aided *N*-glycan separation because it makes use of the filter-aided sample preparation (FASP) approach for SDS removal and buffer exchange (Manza et al., 2005, Wiśniewski et al., 2011). FASP involves the use of centrifugal filter

units that consist of an inner, porous membrane-bottomed tube, through which molecules above a certain molecular weight cannot pass, which is suspended inside a collection tube. Buffer exchange is essential because *N*-glycans are released enzymatically using peptide *N*-glycosidase F (PNGase F,) which would be inhibited by SDS. Centrifugal filters are used to exchange the SDS lysis buffer for a urea solution before treatment with iodoacetamide to alkylate thiol groups and prevent reformation of disulfide bonds (Suttapitugsakul et al., 2017). Following alkylation, the buffer is exchanged again for volatile ammonium bicarbonate, which provides the optimum environment for enzymatic *N*-glycan release (Tarentino and Plummer, 1994).

N-glycan release is fortunately straight-forward, due to the natural availability of PNGase F, which cleaves *N*-glycopeptides at the C-N bond between the asparagine residue's side chain N atom, to which the anomeric C atom of the GlcNAc is bound, and the C atom in the Asn side chain amide (figure 1.14). PNGase F is an amidase that was first identified in and purified from *Flavobacterium meningosepticum* (Plummer et al., 1984, Plummer and Tarentino, 1991). Endoglycosidases (Endo H, F_1, F_2 and F_3) are also expressed by this bacterium, though these all cleave in between the two GlcNAc residues of the *N*-glycan chitobiosyl core and each has a preference for a particular type of *N*-glycan (Tarentino and Plummer, 1994). In contrast, PNGase F has very broad specificity (Plummer et al., 1984) and is able to cleave almost all *N*-glycans from peptide, with the exception of some glycan structures found in plants and insects that contain an α1-3 linkage between their innermost GlcNAc and a fucose residue (Tretter et al., 1991). This broad specificity is because PNGase F interacts with the chitobiosyl *N*-glycan core and the attached polypeptide chain, whilst the oligosaccharide branches are largely unrecognised (Tarentino and Plummer, 1994). PNGase F is therefore used in the FANGS procedure for release of mammalian *N*-glycans. Released *N*-glycans can fit through the pores of the membrane in the centrifugal filter unit, so they can be recovered with water and subsequent centrifugation, leaving the large, free protein behind, retained by the membrane.
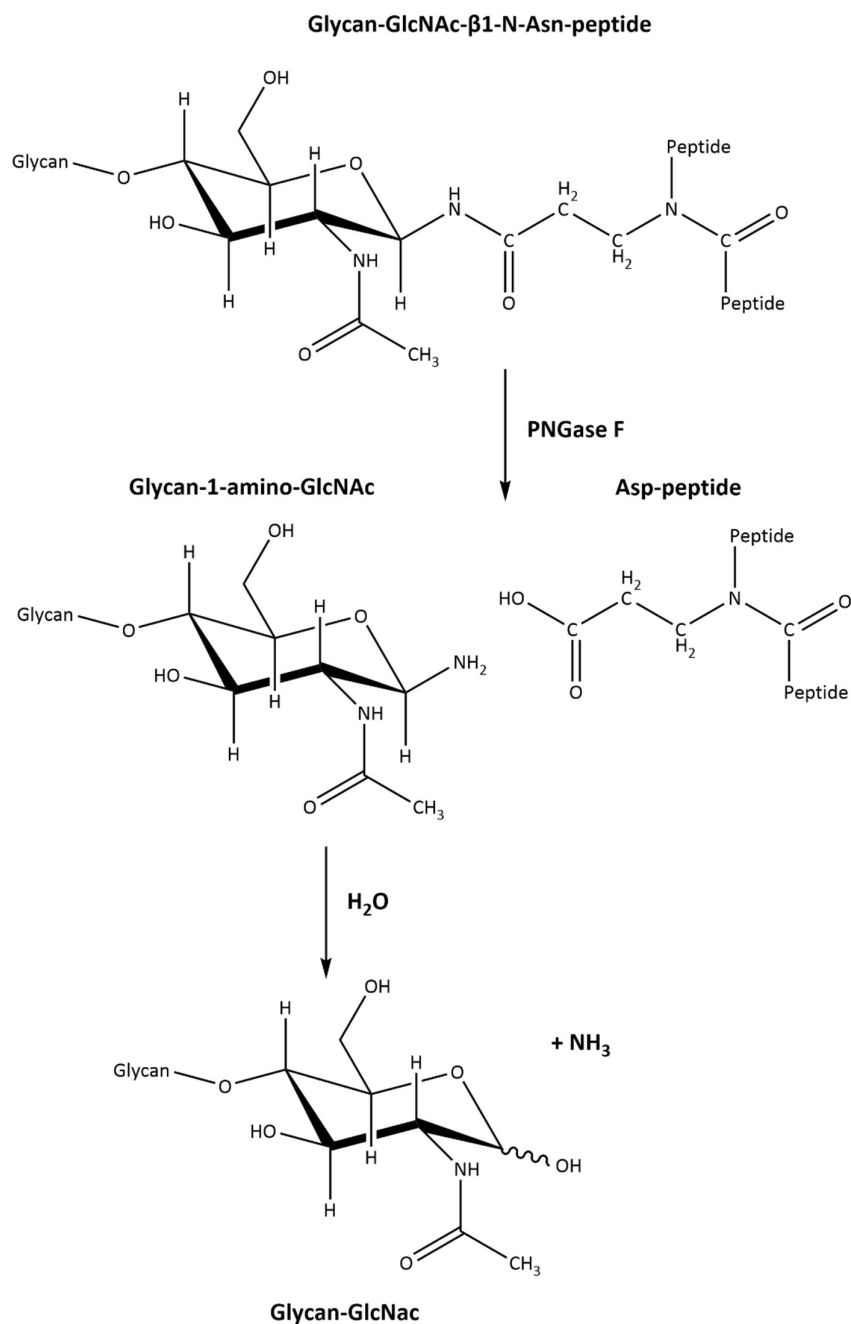
Figure 1.14 | PNGase F- catalysed N-glycan release. PNGase F cleaves N-glycopeptides at the C-N bond between the asparagine residue's side chain N atom to which the anomeric C atom of the GlcNAc is bound and the C atom in the Asn side chain amide. A glycan terminating in 1-amino-GlcNAc forms as a reaction intermediate, but is quickly hydrolysed, releasing ammonia and a glycan with a reducing GlcNAc.

Since the development of the FANGS procedure (figure 1.13), it has been used to release whole-cell glycans from a variety of cell types, revealing glycosylation changes in multiple conditions. It was first demonstrated by Abdul Rahman and colleagues using HeLa and CHO cells of both wild-type (WT) and in cells lacking Cog4 (ΔCog4) (Abdul Rahman et al., 2014). They were able to detect a variety of oligomannose, complex and hybrid glycans in spectra taken for all cells tested and noted that Cog4 knockdown resulted in an increase in hybrid glycan signals in HeLa cell spectra and a reduction in sialylation in both CHO and HeLa cell spectra (Abdul Rahman et al., 2014). The authors also treated HeLa cells with the mannosidase II inhibitor, swainsonine (Elbein et al., 1981), and were unable to detect complex glycan signals in the spectra (Abdul Rahman et al., 2014).

More recently, FANGS has been used in the glycan profiling of mesenchymal stem cells (MSCs) undergoing osteogenesis (Wilson et al., 2016). In an MSC clone that differentiates into bone-forming osteoblasts, the authors showed that differentiation resulted in an increase in complex glycans (Wilson et al., 2016). The authors later showed that glycosylation could modulate differentiation and that swainsonine treatment increased differentiation (Wilson et al., 2018). FANGS was also used in the glycan profiling of HEK293T cells lacking various COG complex subunits (Bailey Blackburn et al., 2016). The authors found several glycosylation defects in ΔCog2, ΔCog4 and ΔCog7 cells, including reduced fucosylation and sialylation compared to WT (Bailey Blackburn et al., 2016).

Together, these studies show that FANGS is sufficient to release whole-cell *N*-glycans, enabling comparisons between different cell lines, different culture conditions, or drug treatments. Furthermore, glycan profiles derived from FANGS-released whole-cell *N*-glycans have provided sufficient information to experimentally validate a new computational model for predicting *N*-glycosylation (Fisher et al., 2019).

1.4.2 Derivatisation of *N*-glycans

Prior to MS analysis, released *N*-glycans are normally derivatised because in their native state, they have poor ionisation efficiency and produce a poor signal response due to having few sites available for protonation (Thomas-Oates and Dell, 1989). A well-adopted method for *N*-glycan derivatisation is permethylation, which involves the

addition of methyl groups to all hydroxyl and *N*-acetyl groups and methyl esterification of sialic acid carboxy functions (figure 1.15). Other derivatisation techniques focus on the reducing position of released *N*-glycans, which can be tagged by fluorescent molecules (e.g. Tomiya et al., 1987). Since *N*-glycans don't contain strong natural fluorophores, tagging with fluorescent molecules offers the advantage of enabling quantitation by fluorescence detectors, which have excellent limits of detection (recently reviewed by Keser et al., 2018).

1.4.2.1 Permethylation

Permethylation involves substituting every acidic proton on a carbohydrate for a methyl group, by treatment with iodomethane in the presence of hydroxide (figure 1.15). The reaction, as most commonly carried out (Ciucanu and Kerek, 1984), involves dissolving dried carbohydrate samples in dimethyl sulfoxide (DMSO), a polar aprotic solvent that solvates both polar and non-polar molecules, before sodium hydroxide (NaOH)-catalysed reaction with iodomethane. In the early stages of method development, some oxidation of alcoxides to carbonyls was reported (York et al., 1990). However, the method has since been improved to eliminate oxidative degradation. Slight modifications to the method proposed by Ciucanu and Kerek include the addition of solid NaOH some time before iodomethane treatment (Needs and Selvendran, 1993), adding iodomethane slowly, introducing a small amount of water in the DMSO and neutralising the base prior to liquid-liquid extraction of the permethylated products (Ciucanu and Costello, 2004).

Permethylation results in the conversion of hydroxide-removable hydrogens (typically, in N-glycans, those on hydrophilic hydroxyl groups, N-acyls, and carboxyl groups) to hydrophobic methyl substituents. This decreases hydrogen bonding between glycans and solvent, increasing the surface activity of the glycan in both MALDI matrix mixtures and in ESI droplets, aiding the formation of gas phase ions and improving signal intensity (Zaia, 2004, Han and Costello, 2013, Dell et al., 1988). This reduced hydrogen bonding also makes permethylated glycans more volatile and amenable to vaporisation. This is helpful in electron ionisation; an energetic ionisation technique developed before MALDI or ESI, which involves sample vaporisation and ionisation by gas phase collision between analyte molecules and high energy electrons. An

advantage of permethylation over other hydroxyl derivatisations, such as acetylation or higher alkylations is that the mass increment it confers is so small (+ 14 Da). This is important because of the sheer number of hydroxyl groups on oligosaccharides, which could cause mass spectrometer *m/z* limits to be exceeded by cumulative mass increments of large derivatising groups.
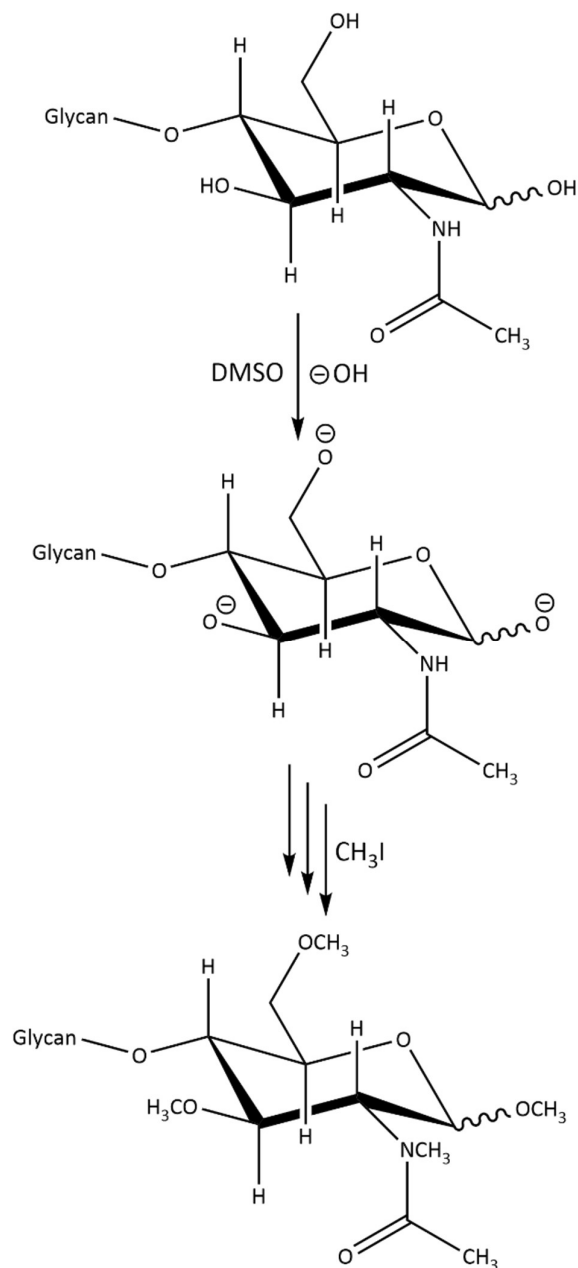
Figure 1.15 | Reaction equation for permethylation of carbohydrates.

Permethylation also esterifies the negatively charged carboxylic acid group of sialic acid residues, producing a neutral derivative with ionisation efficiency equivalent to that of naturally neutral monosaccharides (Wada et al., 2007). This is important because not all glycans are sialylated and the number of un-derivatised sialic residues has previously been shown to affect ionisation efficiency (Sutton et al., 1994). Furthermore, the glycosidic linkage of sialic acid is more labile than linkages between neutral monosaccharides due to the proximity of the sialic acid glycosidic linkage to its carboxyl group. Un-derivatised sialic acids are therefore prone to in-source decay during MALDI, whereas permethylated sialic acid linkages are stabilised in the positive mode (Wada et al., 2007, de Haan et al., 2020).

Most importantly, permethylation of oligosaccharide mixtures and analysis by MALDI-MS has been shown to yield good quantitative results, which correlate well with results from other quantitative methods involving fluorescent labelling (Wada et al., 2007). Therefore, in all the work presented in this thesis, and in all the previous studies of FANGS-released *N*-glycans mentioned in the previous section (Abdul Rahman et al., 2014, Bailey Blackburn et al., 2016, Fisher et al., 2019, Skeene et al., 2017, Wilson et al., 2018, Wilson et al., 2016), permethylation was the derivatisation method of choice, followed by MALDI for mass spectrometric analysis.

1.4.2.2 Reducing-terminal Fluorescent Labelling

A second type of carbohydrate derivatisation makes use of the fact that PNGase F-released *N*-glycans have a free reducing terminus. Explicitly, the innermost GlcNAc of the *N*-glycan core is in equilibrium between its closed ring (usually pyranose) structure and its open straight chain aldehyde form, there is a free carbonyl group (in the aldehyde group). The carbonyl oxygen can undergo reductive amination by reacting with an amine to form a Schiff base reaction intermediate, which is subsequently reduced (Han and Costello, 2013, Harvey, 2011) (figure 1.16a). This reaction is often exploited to introduce aromatic fluorophore tags, such as 2-aminobenzamide (2AB) (figure 1.16b).

An advantage of reducing-terminal tagging is the stoichiometric attachment of one label per released *N*-glycan, which allows for relative quantitation of glycans based on fluorescence intensity. Fluorescence detectors enable accurate quantitation and have

excellent limits of detection so are sensitive enough to detect low abundance glycans (Hase et al., 1981). The sensitivity of fluorescence detectors and accuracy of quantitation is sufficient for industrial glycan profiling of biologics, such as mAbs, the production of which is tightly monitored to limit batch-to-batch variation resulting from *N*-glycan heterogeneity (Kawasaki et al., 2009). The most common fluorescent amine derivative used for *N*-glycan analysis is 2AB (figure 1.16b) (Bigge et al., 1995), though many other tags have been used and these are reviewed in (Harvey, 2011). 2AB was favoured for years because it could be readily detected by fluorescence and samples could be routinely prepared using a labelling kit developed by Ludger.

Figure 1.16 | Reductive amination reaction and the structure of 2-aminobenzamide. A) Reductive amination reaction. B) Structure of 2-aminobenzamide (2AB), an example of a fluorescent amine tag added by reductive amination.

An advantage of reducing-terminal tagging is the stoichiometric attachment of one label per released *N*-glycan, which allows for relative quantitation of glycans based on fluorescence intensity. Fluorescence detectors enable accurate quantitation and have excellent limits of detection so are sensitive enough to detect low abundance glycans (Hase et al., 1981). The sensitivity of fluorescence detectors and accuracy of quantitation is sufficient for industrial glycan profiling of biologics, such as mAbs, the production of which is tightly monitored to limit batch-to-batch variation resulting from *N*-glycan heterogeneity (Kawasaki et al., 2009). The most common fluorescent amine derivative used for *N*-glycan analysis is 2AB (figure 1.16b) (Bigge et al., 1995), though many other tags have been used and these are reviewed in (Harvey, 2011). 2AB was favoured for years because it could be readily detected by fluorescence and samples could be routinely prepared using a labelling kit developed by Ludger.

Despite the ease of fluorescent quantification afforded by 2AB labelling, the process of reductive amination during sample preparation is time consuming. Additionally, ionisation efficiency of 2AB derivatised glycans for MS analysis is relatively poor (reviewed in Mechref et al., 2013, Ruhaak et al., 2010). To address these shortcomings, a new labelling reagent has been developed by Waters and manufactured as a complete *N*-glycan release and labelling kit (Lauber et al., 2015). The kit was developed with rapid sample preparation in mind and utilises a buffer that sufficiently denatures glycoproteins at 95 °C but facilitates rapid PNGase F *N*-glycan release at 50 °C (Lauber et al., 2015). Waters *Rapi*Fluor-MS tag (figure 1.17) does not require reductive amination for reaction and instead exploits a carbamate tagging group, that can facilitate rapid labelling of PNGase F released glycosylamines in less than five minutes through formation of a highly stable urea linkage. A quinoline fluorophore offers excellent fluorescence sensitivity and a tertiary amine tag improves MS signal upon positive ion mode electrospray ionisation (ESI+) due to its hydrophobicity, which reduces hydrogen bonding and increases the surface activity of the derivatised glycan in electrospray droplets (Lauber et al., 2015). This kit is now routinely used by several large pharmaceutical companies for tagging glycans released from purified glycoproteins. It has also been used for *N*-glycan profiling of blood serum from CDG patients (Sturiale et al., 2019, Messina et al., 2021). However, at the time of writing, it has not to the author's knowledge been used for whole-cell glycan profiling.
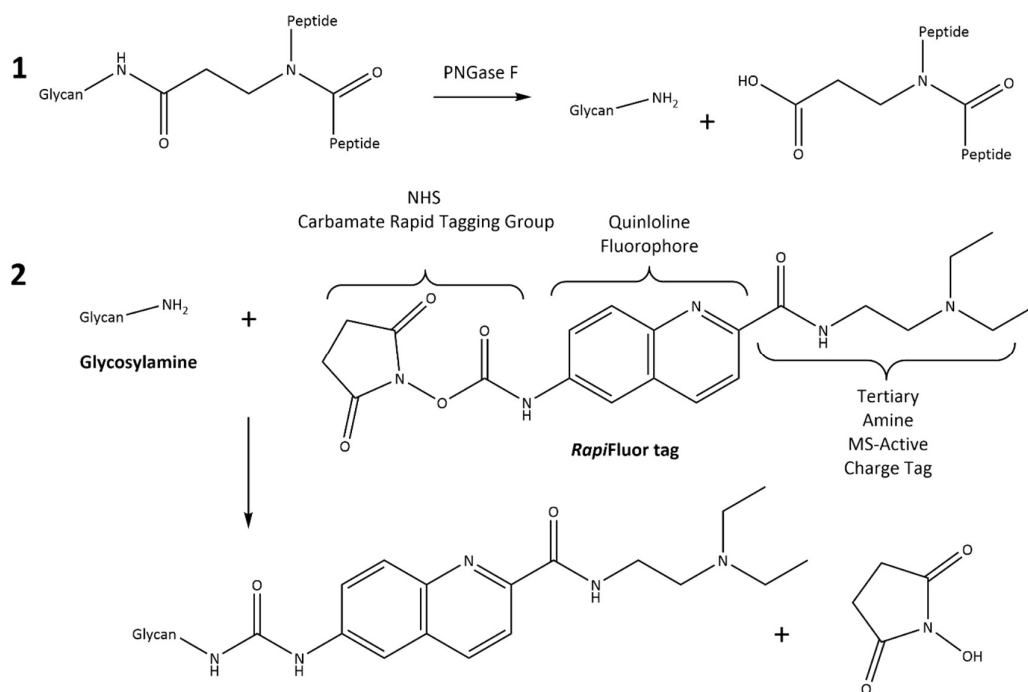
Figure 1.17 | Reaction equations for rapid N-*glycan release and labelling with Water's RapiFluor-MS tag. 1. Rapid* N-*glycan release by PNGase F produces a glycosylamine reaction intermediate. 2. Waters RapiFluor tag reacts rapidly with the glycosylamine by formation of a urea linkage between the glycan and the carbamate portion of the tag, releasing N-hydroxysuccinimide in the process. The structure of Waters RapiFluor tag is labelled, highlighting the N-hydroxysuccinimide (NHS) carbamate reactive group, a quinoline fluorophore, and a basic tertiary amine.* (Lauber et al., 2015).

1.4.3 Coupling Chromatography with Mass Spectrometry

Derivatisation of *N*-glycans by reducing-terminal tagging is often performed prior to separation of the tagged glycans by chromatography. The basic principle of chromatography is that a mixture is carried by a mobile phase, which is a liquid (LC) or gas (GC), through a column containing a stationary phase. The stationary phase is often presented coated onto or made from a porous material with which the analyte molecules interact. Separation is achieved on the basis that different analyte molecules have varying affinities for the mobile and stationary phases and so separate as they partition differently between the two, meaning they elute from the column at different times. Modern liquid chromatographic separation that involves passing mobile phase

79

at high pressure through a stationary phase made up of very small (2.5-5 µm), tightly packed particles is referred to as high performance liquid chromatography (HPLC). Ultra-high performance liquid chromatography (UHPLC) uses even higher pressure and smaller (sub 2 µm) stationary phase particles than HPLC.

Normal phase HPLC, involving a stationary phase that is more polar than the mobile phase, is most commonly used for analysis of fluorescently tagged glycans (e.g. Kalay et al., 2012). When coupled with MS, in online HPLC, it becomes a powerful quantitative and qualitative tool for glycan analysis (Liu et al., 1993, Kawasaki et al., 2000). Mass spectra in this set up provide qualitative information on glycan structure, whilst the retention index scale for HPLC-separated glycans is linear and highly reproducible. This enables rapid glycan structural assignment for repeated runs following initial analysis of the MS data, with quantifiable fluorescence signals from peaks at each retention time. Reversed phase HPLC, where the polarity of each phase is reversed in comparison to normal phase, is commonly used to separate both native and permethylated glycans (Delaney and Vouros, 2001, Han and Costello, 2013, Boulenguer et al., 1988). It frequently involves the use of octadecylsilyl groups (C_{18}) as the stationary phase and typically a mixture of water and methanol as the mobile phase. In both normal and reversed phase HPLC, organic and aqueous mobile phases are applied to the column in a gradient to facilitate the analytes partitioning out of the stationary phase, and so separating.

Hydrophilic interaction chromatography (HILIC) is a variant of normal phase HPLC, first introduced by (Alpert, 1990). It differs from normal phase HPLC because a hydrophobic, mostly organic mobile phase is used with a hydrophilic stationary phase (e.g. $CONH_2$-silica) that sequesters water from the aqueous mobile phase by solvation (Alpert, 1990). Polar carbohydrates interact with this semi-immobilised aqueous layer, partitioning with the organic mobile phase. Elution of analytes, in order of increasing polarity, is achieved by using a combination of organic and aqueous mobile phases, of which the aqueous solution is the stronger eluting solvent (Alpert, 1990). HILIC has been routinely employed for *N*-glycan structural assignment following the development of a method utilising a database of normalised elution times of 2AB-labelled glycans on HILIC amide ($CONH_2$) stationary phases, represented as glucose

units (Royle et al., 2008). This database was developed by normalising the retention times of 2AB labelled glycans against those for a glucose polymer ladder and expressing them as a retention index with respect to the retention of the Glc polymers, in terms of integer and fractional Glc unit retention times (Campbell et al., 2008).

Waters' *Rapi*Fluor-MS kit was optimised for sample preparation prior to UHPLC HILIC using 1.7 μm amide-bonded organosilica particles as the column stationary phase (Lauber et al., 2015). Coupling *Rapi*Fluor-MS *N*-glycan sample preparation with UHPLC HILIC separations has been well adopted in the biotherapeutics industry and has recently been demonstrated for long-term high-throughput IgG *N*-glycan analysis, showing robust and repeatable results (Deriš et al., 2021).

1.4.4 Mass spectrometric ionisation techniques for *N*-glycan profiling

Following release and derivatisation of *N*-glycans, samples can be analysed using a variety of mass spectrometry techniques, with different methods of ionising and analysing these preparations (Mulloy et al., 2015). Two of the most common ionisation techniques for current application to large biomolecules are matrix-assisted laser desorption/ionisation (MALDI) (figure 1.18) (Karas et al., 1985, Karas et al., 1987) and electrospray ionisation (ESI) (Fenn et al., 1989) (figure 1.19). These are both soft ionisation techniques, meaning they result in minimal fragmentation of analyte ions during the ionisation process, as opposed to energetic ionisation processes (e.g., electron ionisation) that cause extensive fragmentation.

1.4.4.1 Matrix-assisted laser desorption/ionisation

Matrix-assisted laser desorption/ionisation, as the name suggests, involves mixing an analyte with a laser-absorbing matrix compound, which facilitates desorption of both the analyte and matrix into gas phase ions when a laser is pulsed at the mixture (figure 1.18). The term MALDI was originally coined by Karas and Hillenkamp in the late 1980s (Karas et al., 1985, Karas et al., 1987). Franz Hillenkamp had previously pioneered a technique called laser-micro-mass analysis (LAMMA), which was a form of laser desorption ionisation (LDI) mass spectrometry capable of analysing inorganic ions embedded in a matrix material (Kaufmann et al., 1979). In a collaboration with Michael Karas, the method was applied to amino acids, leading to the discovery that aromatic amino acids desorbed with less energy than aliphatic ones and that mixtures of

aromatic and aliphatic amino acids were able to desorb simultaneously (Karas et al., 1985). They later used an aromatic nicotinic acid matrix to facilitate LDI of proteins exceeding 10,000 Da (Karas and Hillenkamp, 1988). Around the same time, Koichi Tanaka and colleagues used LDI mass spectrometry with an ultra-fine metal powder matrix for the analysis of proteins and polymers of up to *m/z* 100,000 (Tanaka et al., 1988). However, it was Karas and Hillenkamp who first published the use of MALDI as we know it today, with superior aromatic matrices (Hillenkamp and Karas, 1990). They later tested a new aromatic matrix material, 2,5-dihydrobenzoic acid (DHB), and found it worked well and was insensitive to contaminating salts (Strupat et al., 1991). This matrix material is still commonly used today for analysis of biological samples (reviewed in Harvey, 2018).
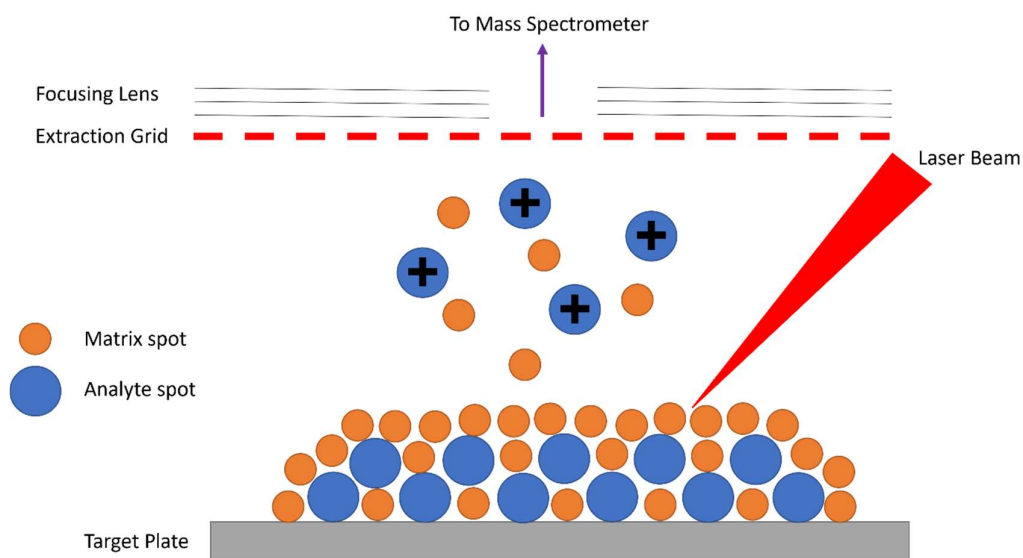
Figure 1.18 | Schematic representation of the ionisation process occurring during MALDI. A short laser pulse is fired at the MALDI target plate on which samples are deposited in spots, which contain a co-crystal of matrix and sample. Energy from the laser is absorbed by matrix molecules, causing ionisation and desorption of both matrix and sample ions. Ions pass through an extraction grid before being focussed through a lens into the mass analyser.

Aromatic matrices, such as DHB, facilitate LDI because their electron-dense aromatic rings absorb laser energy, exciting the matrix and causing an explosive vaporisation process, generating a plume, and pulling mixed analyte molecules into the gas phase

with it. During this desorption, singly charged gas phase analyte ions are formed, although the details of the ionisation process remain a matter of ongoing research. Two main ionisation models have been proposed: the lucky survivor theory (Karas et al., 2000) and the gas phase protonation model (Ehring et al., 1992). The gas phase protonation model postulates that neutral analytes are present in the gas phase, either due to their incorporation into matrix crystals as uncharged species or due to neutralisation in the gas phase plume by counter ions. These gaseous neutral analytes gain their positive charge through proton transfer resulting from gas-phase collision with positively charged matrix ions (Ehring et al., 1992). In contrast, the lucky survivor theory hypothesises that analytes retain the charge states they would normally have in solution when they are incorporated into matrix/analyte crystals. When laser ablation occurs, there are two possible outcomes: either analyte charges are neutralised by counter-ions, or a lack of counter-ions means that positive analyte ions are not neutralised and are detected as "lucky survivors" (Ehring et al., 1992). There is evidence supporting both theories, and now it's accepted that both scenarios are likely to occur in parallel, with various conditions dictating which model occurs to the greater extent in any particular instance (Jaskolla and Karas, 2011).

MALDI is a popular tool for glycan analysis as, for permethylated glycans, it has been shown to be quantitative (Wada et al., 2007). It offers good limits of detection for glycans and increased resistance over other ionisation techniques to conditions such as high salt concentration, which are common in biological samples (Kailemia et al., 2014). However, it generates singly charged ions, meaning the analyte mass detection limit is restricted to the *m/z* limit of the mass analyser and larger glycans may be missed.

1.4.4.2 Electrospray Ionisation

Electrospray ionisation (ESI) mass spectrometry was first introduced by Malcom Dole, who discovered the phenomenon of electrospray by chance, whilst visiting a car manufacturer and observing how very fine, charged droplets of paint were attracted to the metal surface of a car body. By diluting a polystyrene polymer in a volatile solvent and firing it through a charged hypodermic needle, he was able to observe gas phase macroions (Dole et al., 1968a). At the time, he suggested this technique could

be used as an ionisation source for mass spectrometry. However, it wasn't until nearly 20 years later that ESI-MS was developed (Whitehouse et al., 1985). John Fenn and colleagues were able to use ESI as an ion source for a mass spectrometer with a quadrupole mass analyser (section 1.4.4.2), which they later used for the analysis of large biomolecules of up to 130 kDa (Whitehouse et al., 1985, Fenn et al., 1989). For this breakthrough, Fenn was awarded a quarter of the 2002 Nobel Prize for Chemistry.
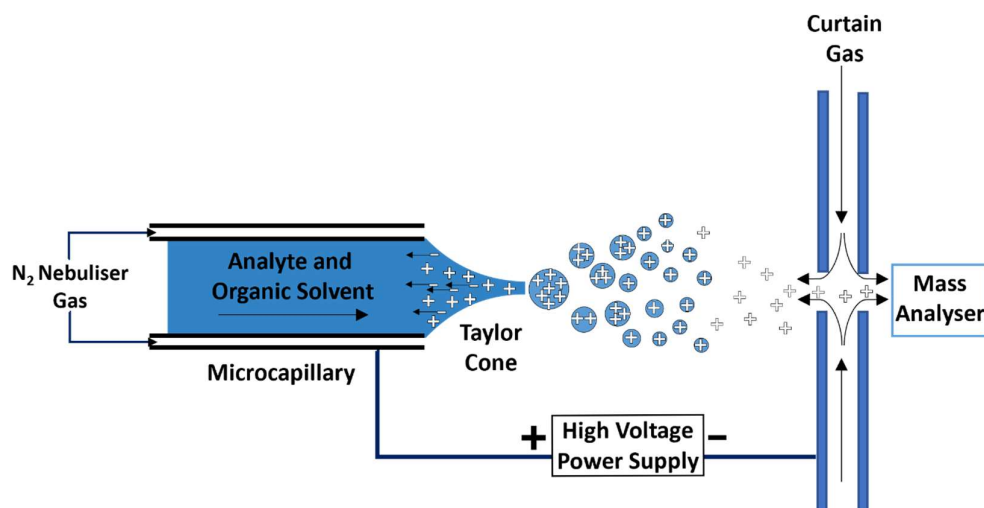
Figure 1.19 | Schematic representation of the processes occurring in an electrospray ionisation source.

The principals of ESI had begun to be established as early as the 1880s. In 1882, Lord Rayleigh calculated the "Rayleigh limit", which is the maximum density of charged particles that a liquid droplet can contain, before expelling fine jets of liquid (Rayleigh, 1882). This phenomenon was studied further by John Zeleny, who utilised glass capillaries to study electrical discharge from liquid droplets (Zeleny, 1914, Zeleny, 1917). Geoffrey Taylor later published his observation of a "Taylor cone" during electrospray, which is a distortion of liquid emerging from a charged capillary into obtuse-angled conical points, from which coloumbic forces expel the charged droplet spray (Taylor, 1964).

ESI sources exploit these phenomena to generate gas phase ions. A positively charged capillary retains negatively charged ions whilst positively charged droplets are sprayed out in solvent/analyte droplets due to repulsion. Nebuliser and curtain gases, usually

N$_2$, aid the process of droplet formation and solvent evaporation to accelerate gas phase ion formation (figure 1.19). However, like MALDI, the mechanism of gas ion formation is still the subject of debate, and two different models have been proposed: the charged residue model and the ion evaporation model.

The charged residue model, proposed by Malcolm Dole (Dole et al., 1968b), postulates that solvent evaporation from charged droplets results in the Rayleigh limit being exceeded, causing finer droplets to be expelled in a process of coloumbic explosion. This process continues in cycles until droplets contain only a single gas phase ion. The ion evaporation model, proposed by Iribarne and Thomson, (Iribarne and Thomson, 1976, Thomson and Iribarne, 1979) differs in that evaporation of solvent does not cause coloumbic explosion. Instead, as the charge density of the droplet increases, ions desorb from the surface of the droplet, temporarily reducing the droplet charge density below the Rayleigh limit. As solvent continues to evaporate, this process repeats until all the solvent has evaporated and only gas phase ions remain. Like the two theories for ionisation in MALDI, there is evidence supporting and contradicting both theories, and it's likely that a combination of the two processes takes place simultaneously (Wilm, 2011).

A major advantage of ESI over MALDI is that it generates multiply charged ions, as there are no forces preventing analytes with multiple protonation sites from picking up more than one charge (Fenn et al., 1989). This effectively increases the size limit of detectable analyte molecules. ESI can also easily be hyphenated with chromatography columns, making it automatable and popular for high-throughput applications, such as in the industrial glycan profiling of biologics. However, absolute quantitation of glycans using ESI, even with internal standards, is unreliable (Grunwald-Gruber et al., 2017), so quantitation often relies on using fluorescent tags and HPLC separation before ESI. Therefore, since MALDI is available on several instruments in York, and offers the best quantitation when combined with glycan permethylation, it is preferred for the analysis of permethylated *N*-glycans from whole cell lysates and is the method most commonly used in our labs (Abdul Rahman et al., 2014, Fisher et al., 2019, Wilson et al., 2018, Wilson et al., 2016).

1.4.5 Mass Analysers

Following ionisation, gas phase ions are transferred into a mass analyser, where they are separated on the basis of their *m/z* values and detected. As with ionisation sources, there are multiple types of mass analyser currently available. The work presented in this thesis made use of time-of-flight and quadrupole mass analysers. The aim of the following sections is to provide a brief description of how each analyser works and the advantages and disadvantages of each.

1.4.5.1 Time-of-Flight

Time-of-flight mass analysis, which was first introduced in the 1940s (Goudsmit, 1948, Hipple and Thomas, 1949), is perhaps the most intuitive type of mass spectrometry. At its simplest, it involves the use of a field-free flight tube, into which gas phase ions are accelerated at a constant velocity by an electric field of known strength (figure 1.20). Essentially, ions with a smaller *m/z* reach higher velocities under the accelerating voltage and reach the detector before ions with larger *m/z* values. The relationship between *m/z* and flight time is given by the equation shown in figure 1.20, which also involves the electronic charge of the ion, the applied acceleration voltage, time of flight and the length of the flight tube.

Acceleration Region | Field Free Drift Zone

$$m/z = 2eV\left(\frac{t}{D}\right)^2$$

Gas Phase Ions from Ionization Source

+ve potential   -ve potential

Detector
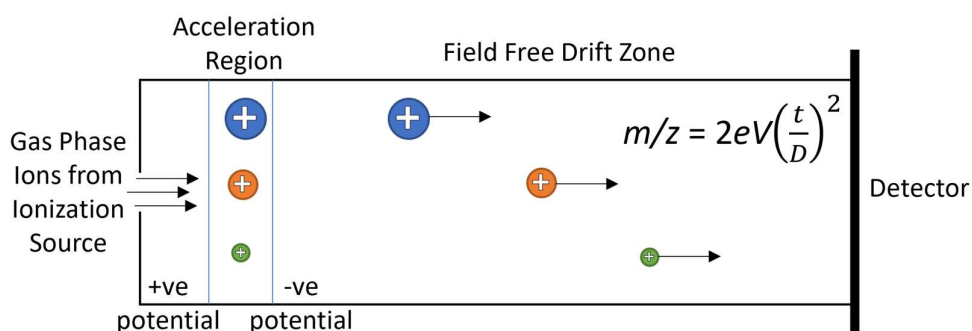
Figure 1.20 | Schematic of a basic time-of-flight instrument. The equation for calculating m/z *is shown, where e = electronic charge, V = the acceleration voltage, t = time of flight and D = distance (length of the flight tube).*

TOF-MS in a linear instrument such as shown in figure 1.20 is well suited to use with MALDI ionisation, as this involves pulsing a laser at the sample, causing a burst of ions to enter the flight tube together. However, despite ions entering the acceleration region together, ions of the same *m/z* each have a distribution of possible flight times,

resulting in poor mass resolution (broader spectral peaks). Reasons for differences in flight times are well reviewed in (Cotter, 1992). Briefly, these reasons include the fact that even with laser pulsing during MALDI, ions are formed at slightly different times in the MALDI plume and so they reach the detector at slightly different times. Additionally, ions can form in different regions of the accelerator or enter the accelerator travelling on slightly different trajectories, also resulting in ions with the same *m/z* reaching the detector at different times. Furthermore, the applied acceleration voltage is always a distribution of voltages, giving ions of equivalent *m/z* a distribution of velocities. Although these factors could be mitigated by increasing the length of the time-of-flight tube to improve mass resolution, this is not a practical solution considering laboratory size constraints. Furthermore, a simple in line linear TOF instrument as is shown in Figure 1.20 would not be suitable for use with EI or ESI as ions are formed continuously rather than in a pulse.

Fortunately, these shortcomings have been overcome with the introduction of time delayed pulse ion extraction (Wiley and McLaren, 1955), the development of an ion reflecting device, called a reflectron (Mamyrin et al., 1973) and the use of orthogonal acceleration (Dawson and Guilhaus, 1989, Guilhaus et al., 2000). These developments are well reviewed in (Radionova et al., 2016). Delayed pulse extraction involves the introduction of a second acceleration region, allowing a time delay between ions being accelerated initially and a second acceleration voltage application that extracts ions into the flight tube. This enabled higher resolution to be achieved with MALDI as the delayed pulse essentially allows ions with the same *m/z* but slower initial velocity to catch up with faster ions during the time delay (Vestal et al., 1995, Brown and Lennon, 1995).

Additionally, most modern TOF instruments are fitted with a reflectron to improve resolution (Mamyrin et al., 1973). Briefly, a reflectron is a stack of ring-shaped electrodes that generates an electric field capable of slowing down and reflecting ions, that can be so set up to correct the distribution of kinetic energies for ions with the same *m/z*. Positioning the reflectron at an angle at the end of one flight tube and adjusting the voltage of the electrode rings causes a change in the flight path of ions as they enter and exit the electric field generated by the device. Ions travelling faster can

penetrate further into the field and so they follow a longer flight path as they are reflected. As a result, ions with the same *m/z* but differing kinetic energies exit the reflectron and re-enter the time-of-flight tube now travelling in the opposite direction leading to arrival at the detector at the same time.

Both delayed pulse extraction and reflectrons were introduced to improve *m/z* resolution in TOF mass analysers. Coupling continuous ionisation sources, such as EI in the early days and later, ESI, with TOF analysers required additional measures to transform the continuous ion beam into a pulsed beam (reviewed in Weickhardt et al., 1996). This could be achieved using a variation on delayed pulse ion extraction in which ion beams are orthogonally accelerated, resulting in a 'slice' of the continuous ion beam being introduced into the TOF tube with a trajectory perpendicular to the original trajectory (Dawson and Guilhaus, 1989, Guilhaus et al., 2000). Orthogonal acceleration TOF (oaTOF) mass analysers work by focussing a beam of ions from their ionisation source (continuous or pulsed) into an accelerator region with an intermittently charged metal plate positioned parallel to the ion beam. The charge applied to the plate matches that of the analyte ions (usually positive), resulting in repulsion of the section of ion beam that aligns with the plate. Pulsed voltages are applied to the metal plate, intermittently altering the ion beam trajectory by around 90° and repelling the slice of the ion beam further into the orthogonal accelerator, where delayed pulse extraction is applied before ion slices are focussed into the flight tube through an aperture (Dawson and Guilhaus, 1989). The result is that ions of the same m/z that enter the accelerator region from the ionisation source at slightly different times are accelerated into the flight tube together.

All spectra presented in this thesis for permethylated glycans from whole-cell lysates were recorded using a MALDI-in-line TOF-MS instrument. Glycan profiles for recombinantly expressed mAbs were generated using ESI-TOF-MS.

1.4.5.2 Quadrupoles

Early magnetic detector mass spectrometers required a constant magnetic field, which made them very expensive, and their operation remained elusive to all but highly specialised mass spectrometrists (Dawson, 2013). Increased industrial demand in the 1950s led to the development of commercial mass spectrometry equipment and the

search for a way to shift from the use of magnetic fields (Dawson, 2013). TOF-MS was one such development during this period and of all the different mass analysers available, it offers the highest theoretical *m/z* limit. However, TOF instruments remained bulky and had low resolution until the introduction of delayed pules ion extraction and the use of reflectrons. In the 1950s, an alternative, much smaller mass analyser called a quadrupole was developed (Paul and Steinwedel, 1953) (figure 1.21). Quadrupole mass spectrometers are less intuitive than TOF mass spectrometers, but are robust and cheap, and can be coupled with ESI and therefore hyphenated with HPLC equipment easily.

Essentially, quadrupoles consist of four parallel, usually cylindrical, metal rods, arranged equidistantly from a central axis. Opposite pairs of rods are charged positive or negative, by the application of direct current (DC) and radio frequency (RF) voltages, with at any one time, opposite rods carrying like charges and adjacent rods carrying the opposite charge. Gas phase ions are focussed into the centre of the quadrupole. Due to the superposition of the RF voltage on the rods already bearing their DC voltages, the charges on the paired electrodes oscillate between positive and negative – the rods referred to as the positive rods spend the majority of their time carrying an overall positive charge with brief negative excursions, while those referred to as the negative rods are predominantly negatively charged. Consequently, ions travel in a roughly corkscrew motion through the oscillating electric field resulting from the superposition of the RF voltage onto the DC voltages, as they are alternately pulled towards negatively charged adjacent rods, and then repelled when the rods become positively charged. At a particular set of DC and RF voltages, some ions take stable trajectories whilst others don't (figure 1.21). These voltages can be changed to scan through *m/z* values and enable ions of different *m/z* values to become stable.
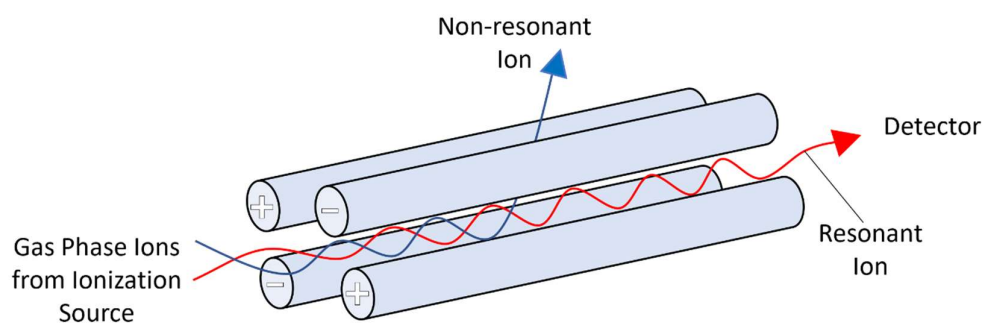
Figure 1.21 | Schematic representation of a basic quadrupole mass analyser, featuring a resonant ion that reaches the detector and a non-resonant ion that filters out of the quadrupole.

Quadrupoles are often used as mass filters, as the trajectory of an ion through the rods is dependent on its *m/z.* DC and RF voltages can be applied such that only ions in a particular *m/z* range can take a stable trajectory through the quadrupole and reach the detector – these are so-called resonant ions. Ions outside of the selected *m/z* range take an unstable trajectory, collide with the metal rods and do not reach the detector – they are non-resonant ions. This makes quadrupoles ideal for applications where the range of *m/z* values is known, for example in mAb glycan analysis. Users can also scan through a range of *m/z* values to detect a broad range of ion signals. Quadrupoles can also be used as ion focussing devices without mass filtering if only RF voltage is supplied. This makes them suitable mass analysers for hyphenation with chromatography, as components of the mixture are separated before they enter the quadrupole.

Glycan profiles presented in this thesis for glycans derivatised with the *Rapi*Fluor tag were generated using UHPLC HILIC-ESI hyphenated with a quadrupole mass analyser.

1.5 Thesis Summary and Research Aims

The overall aim of the previous introductory sections was to describe the relevant background information necessary for understanding the work presented in this thesis. To summarise, glycosylation (introduced in section 1.1) is an important modulator of protein structure and function (section 1.1.7), which is relevant both for controlling cellular characteristics and for regulating the properties of recombinantly expressed glycoprotein biopharmaceuticals (section 1.3), such as monoclonal antibodies (mAbs) (section 1.3.2). Glycosylation of biopharmaceuticals is therefore tightly monitored and controlled either through manipulating host cell culture conditions or through the use of genetically engineered host cells lacking a specific glycosyltransferase (section 1.3.4).

In the absence of a molecular template, glycosylation is controlled by localisation of enzymes within the Golgi apparatus, which requires tightly regulated membrane trafficking (section 1.2), involving highly specific vesicle tethering interactions (section 1.2.3). These tethering interactions are coordinated by the herooctomeric, multi-subunit tethering, COG complex (section 1.2.5). Mutations in COG subunits have been shown to cause glycosylation changes, exemplified in congenital disorders of glycosylation (section 1.2.6) and further demonstrated in genetically engineered COG mutant cell lines (section 1.2.4.2). Considering the well-established role played by the COG complex in the process of glycosylation, the lab groups of Prof. Dani Ungar and Prof. Jane Thomas-Oates have extensively studied glycosylation in COG subunit knock-out cell lines, using mass spectrometric glycan profiling (section 1.4).

Knocking out COG subunits, as previously carried out in these labs and others (section 1.2.5.2), has previously been shown to alter glycosylation. However, the effects of knocking out individual subunits are similar regardless of which subunit is targeted, as disrupting any one subunit destabilises the whole complex (Bailey Blackburn et al., 2016). The work in this thesis builds instead on unpublished work to generate functional point mutations in COG subunits that target specific interactions between COG subunits and Rab proteins. These COG mutations were originally generated with the goal of dissecting the functions of specific tethering interactions in controlling glycosylation and their influence on cell properties. In light of an observation

(unpublished data) that one of these mutations in Cog4 conferred a glycan profile distinct from either wild-type or Cog4 knockout (ΔCog4) when expressed in a CHO-K1 cell line, the potential for engineering glycosylation by targeting COG subunits was highlighted. The work presented in this thesis therefore builds on this previous research with the overall aim of testing a novel approach to engineering production and glycosylation of glycoprotein biopharmaceuticals (see section 1.3.4 for existing approaches), by generating a COG subunit mutant CHO-K1 cell line. This work was devised in collaboration with GlaxoSmithKline (GSK), under additional supervision by Dr Kirsty Skeene. The following paragraphs summarise the specific aims and objectives of each thesis chapter and how they combine to address the overall aims of the research presented in this thesis.

Chapter two introduces and describes how the research builds on previous work in the Ungar group to identify and characterise an L36P mutation in Cog4 that specifically impairs its interaction with Rab30. Some of this work was carried out using WT and Cog4L36P knock-in adherent CHO-K1 cell lines previously developed in the Ungar lab. However, these knock-in cell lines express Cog4 to different extents so a key objective was to engineer the point mutation under its endogenous promoter using CRISPR-cas9 site-directed mutagenesis. After successfully engineering a Cog4L36P mutant adherent CHO-K1 cell line and adapting it to suspension culture, the first objective was to validate the impairment to Rab30 binding, originally identified in a yeast-two-hybrid system, in Cog4L36P containing COG complex expressed by the mutant cells using a GST pull-down approach. Characterisation experiments with both the knock-in and CRISPR-engineered adherent and suspension cells aimed to establish the effects of Cog4L36P mutation or Cog4 knockout on cell proliferation, viability, and metabolic activity. Additionally, in light of the influence of glycosylation on cell adhesion (section 1.1.7.4), another aim was to characterise adhesion in the WT and mutant cells by testing how well the cells adhered to a range of substrates, including ECM secreted by either WT or Cog4L36P cells.

The aim of the work presented in chapter three was to characterise the effects of Cog4L36P mutation on glycosylation on a whole-cell level by mass spectrometric glycan profiling (section 1.4). One objective was to generate whole-cell glycan profiles for both

adherent- and suspension-cultured WT, ΔCog4 and Cog4L36P cells using mass spectrometry in order to understand how Cog mutation impacts glycosylation. Another objective was to compare glycan profiles for adherent and suspension-cultured varieties of each cell type to understand how glycosylation in the different cell-types changes in response to different culture conditions. This work utilised an approach previously developed in the groups of Prof. Jane Thomas-Oates and Prof. Dani Ungar, involving filter-aided *N*-glycan separation (section 1.4.1) and permethylation (section 1.4.2.1) prior to analysis by MALDI-MS (sections 1.4.3.1 and 1.4.4.1). To aid collaborative work with GSK, where their industrial labs lack MALDI-MS instrumentation and a semi-automated, high-throughput HPLC-based approach to mAb glycan analysis is routinely used, another aim was to adapt GSK's industrial mAb glycan profiling method for whole-cell glycan profiling. This approach involves using fluorescent labelling (section 1.4.2.2) prior to UHPLC-ESI-MS (sections 1.4.3, 1.4.4.2 and 1.4.5.2).

Chapter four describes work to put the overall aim of this research to the test by assessing the production and glycosylation of biopharmaceuticals in Cog4L36P mutant suspension CHO-K1 cells compared to their WT parent cells. This involved monitoring the production of a readily available mAb, Herceptin (section 1.3.2.1), in Cog4L36P and WT cells by performing a transient transfection and analysing the recombinantly produced protein. A highlight of the results presented in this chapter is the observation of enhanced Herceptin production levels in Cog4L36P cells with unaltered glycosylation, demonstrating the potential of this cell line for aiding industrial mAb production.

The CRISPR-engineered suspension-adapted cells developed as described in chapter two however, are not optimal host cells for biologic production. Production CHO-K1 cells have been adapted and optimised for suspension culture over many years. The process of adaption from adherent to suspension culture used for this thesis was on a much shorter timescale and the cells retain some adherent properties, causing them to clump together under certain conditions. Therefore, another objective of the work presented in chapter four was to use CRISPR-cas9n to engineer the Cog4L36P mutation in a legacy CHO-K1 suspension cell line, with properties similar to mAb production host-

cells, provided by GSK. Herceptin production levels were also monitored in this cell line and compared to those of its WT parent cell line to test the potential of the Cog4L36P mutation for enhancing mAb production in conditions more closely resembling those used during industrial production.

Overall, the aim of this thesis is to explore the potential of targeting the COG complex for aiding the production of biologics. Research chapters presented in this thesis together explore how whole-cell characterization observations in WT and Cog4L36P cells (chapter 2) relate to their observed glycan profiles (chapter 3) and biologic production levels (chapter 4). Together, they address the overall aim of this thesis by beginning to explain how altering Cog4 can cause specific glycosylation changes and how glycosylation relates to changes in cell properties and protein production.

Chapter 2: Generation and Characterisation of a CHO-K1 Cell Line with an L36P Mutation in Cog4

2.1 Introduction

2.1.1 Rationale for targeting the conserved oligomeric Golgi complex for engineering glycosylation

The COG complex is a key coordinator of enzyme sorting in the Golgi and disrupting the complex or its subunits can result in serious consequences on glycan processing (Shestakova et al., 2006, Smith and Lupashin, 2008). This is exemplified in patients with congenital disorders of glycosylation (CDGs) caused by COG mutations, which are discussed in more detail in section 1.2.5 (Foulquier et al., 2006, Wu et al., 2004). CDG patients exhibit a range of pleiotropic symptoms and altered glycosylation (Jaeken, 2011, Leroy, 2006). For example, an R729W mutation in Cog4 disrupts Golgi structure, causing reduced levels of galactosylation and sialylation (Reynders et al., 2009). Similar galactosylation and sialylation defects have also been seen in a patient carrying an L773R mutation in Cog4, whose fibroblasts also displayed a deficiency in Brefeldin A induced retrograde trafficking (Ng et al., 2011). CDGs affecting different COG subunits can affect glycosylation in different ways, altering the severity of the phenotype. For example, CDGs affecting Cog7 (Wu et al., 2004) cause a more severe glycosylation defect and phenotype than any other CDG, resulting in lethality (recently reviewed in D'Souza et al., 2020). Different mutations in the same COG subunit can also affect glycosylation in different ways: L773R and R729W Cog4 CDG patients have reduced sialylation and galactosylation (Ng et al., 2011, Reynders et al., 2009), whereas other Cog4 CDG patients with a G516R mutation showed no glycosylation defects to serum transferrin but altered glycosylation of a secreted glycoprotein, decorin (Ferreira et al., 2018).

These observations highlight the potential for changing glycosylation in different ways by targeting specific subunits of the COG complex. So far, COG subunit deficiencies have been studied in knock-down experiments using multiple different cell types,

including Hela (Zolov and Lupashin, 2005, Oka et al., 2005, Pokrovskaya et al., 2011, Willett et al., 2013a, Fisher et al., 2019), MSCs (Wilson et al., 2018) and even Vero (monkey kidney epithelial) cells (Smith et al., 2009). All COG subunit knock-downs resulted in glycosylation and trafficking defects (reviewed in Blackburn et al., 2019). CRISPR/cas9 has also previously been used to completely knock out COG subunits in HEK293T cells (Bailey Blackburn et al., 2016, Blackburn and Lupashin, 2016). These cell lines revealed deficiencies in cis/medial Golgi glycosylation, altered sialylation and fucosylation, and defects in retrograde trafficking and sorting (Bailey Blackburn et al., 2016). Importantly, knockdown of different COG subunits altered glycosylation in different ways: for example, Cog2, Cog4 and Cog7 deficient cells displayed elevated levels of oligomannose glycans, notably increased $GlcNAc_2Man_5$ compared to WT, but relative abundance varied depending on which subunit was effected, with less $GlcNAc_2Man_5$ detected in ΔCog4 cells than either ΔCog2 or ΔCog7 cells (Blackburn and Lupashin, 2016).

As well as changes to glycosylation and trafficking, COG deficiency commonly causes morphological changes to the Golgi structure (Bailey Blackburn et al., 2016, Laufman et al., 2013a, Shestakova et al., 2006) and redistribution of glycosylation enzymes, due to trafficking defects (Oka et al., 2004, Pokrovskaya et al., 2011, Shestakova et al., 2006). In yeast, depletion of any COG lobe A subunits also causes growth defects (VanRheenen et al., 1998, Whyte and Munro, 2001). These findings highlight the importance of cellular characterisation in COG deficient cells.

On the basis that all evidence so far suggests disrupting COG alters glycosylation, and that targeting different COG subunits causes different glycosylation defects, this thesis explores targeting the COG complex, specifically Cog4, for engineering glycosylation of recombinantly expressed biologics. This chapter focuses on generating a point mutation in Cog4 and characterising the effects of the mutation on a CHO-K1 cell line. So far, the only published examples of COG deficient CHO cells are those lacking Cog1 or Cog2, termed ldlB and ldlC cells, in which the COG complex was first identified (Krieger et al., 1981, Podos et al., 1994, Chatterton et al., 1999, Ungar et al., 2002). As with all other cell types lacking COG subunits, these cell lines both display glycosylation defects (Kingsley et al., 1986, Abdul Rahman et al., 2014). However, at the time of

writing this thesis, there is no published data for CHO-K1 cells lacking Cog4. Furthermore, to the author's knowledge, CRISPR/cas9 has never been used to generate a COG point mutation in CHO-K1 cells. Mutating Cog4 using CRISPR/cas9 is therefore a novel approach to engineering glycosylation.

2.1.2 Identification of the Cog4L36P mutation, which does not bind Rab30

Of all the COG subunits, Cog4 has the most extensive structural characterisation (Richardson et al., 2009) and many of its binding partners have been identified (Loh and Hong, 2004, Laufman et al., 2009). Cog4 normally interacts with several Rabs, including a strong interaction with Rab30 (Miller et al., 2013), which does not interact with any other COG subunit. Therefore, prior to the start of this PhD, the Cog4-Rab30 interaction was targeted in a random mutagenesis-based selection to investigate its role in enzyme sorting at the Golgi (personal communication with Prof. Daniel Ungar). A Cog4 mutant library was generated by Vicky Miller and cloned into Gal4 activation domain (GAD) vectors to generate yeast strains for a reverse-2-hybrid screen, which was carried out by previous project student, Jenny Burgum. The screen assessed the interaction of mutant Cog4-GAD with Rab30-Gal4 binding domain (GBD) fusion protein, utilising expression of the URA3 reporter gene. In the presence of 5-FOA, only yeast transformed with mutants that did not interact with Rab30 were able to survive. Several Rab30 non-binding Cog4 mutants were identified, and the result was further validated using a traditional yeast-2-hybrid (Y2H) approach, with growth on media lacking uracil indicating an interaction (data not shown).

This approach led to the identification of a Cog4 mutant, containing several point mutations, which did not bind Rab30. A previous PhD student in the Ungar group, Peter Fisher, worked to generate plasmids with each individual point mutation, to identify if any mutation alone could account for the observed phenotype. Of all these mutations, only a leucine to proline mutation at position 36 (L36P) could account for reduced Rab30 binding (data not shown).

2.1.3 Rationale for Engineering the Cog4L36P Mutation

Prior to the start of this PhD, an adherent CHO-K1 cell line bearing the Cog4L36P point mutation was generated using CRISPR-Cas9 to knock out the Cog4 gene (ΔCog4), before stably transfecting in the Cog4L36P gene using a plasmid. The ΔCog4 cells were also

used to generate a WT Cog4 knock-in control cell-line. Glycan profiling to compare Cog4L36P cells to both WT and ΔCog4 cells revealed the presence of glycans that were unique to the mutant cell line (personal communication from Dr. Dimitrios Kioumourtzoglou- data not shown). I worked with these cell lines initially at the start of this PhD and supervised a final year project and a master's student with their cellular characterisation. We observed distinct phenotypic differences between Cog4L36P CHO cells and the WT control, which are discussed later in this chapter. These included reduced adhesion, which is favourable for biologic production in suspension culture, and differences in growth rate and metabolic activity.

These findings together highlighted the potential for engineering glycosylation by generating COG mutants and suggested that the Cog4L36P cell-line could be a promising host-cell line for biologic production. However, the Cog4L36P knock-in cell line overexpresses Cog4 compared to WT (figure 2.1). Although we have a WT Cog4 knock-in cell line for comparison, it also overexpresses Cog4 to a different extent (figure 2.1). Attributing phenotypic differences to the mutation alone may be an inaccurate assumption as they could be due to expression levels. It is also unclear where the Cog4 transgene was integrated into the genome in either of the cell lines and the consequences of differential integration sites for the Cog4 gene are unknown. Furthermore, transgenes are prone to silencing, and expression of Cog4L36P in these cells could easily be lost without maintaining antibiotic selection. Therefore, accurately characterising the effects of this mutation necessitates generating a point mutation expressed under its own promoter in the genome.
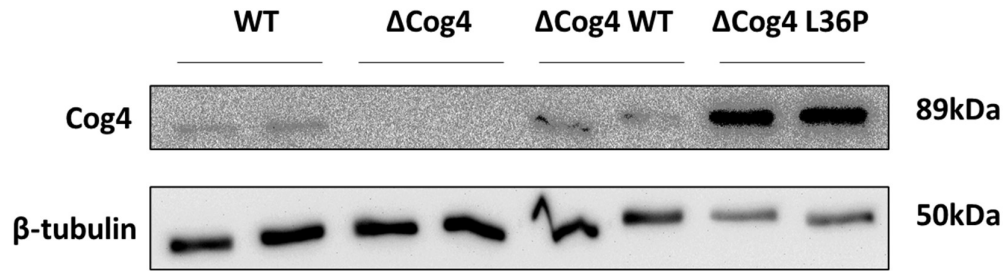
Figure 2.1| Western blots showing Cog4 levels in WT, ΔCog4, ΔCog4 WT and ΔCog4 L36P CHO-K1 cells. Cell lysates were prepared from 100% confluent wells of a 6 well plate, with duplicate wells for each CHO-K1 cell type. Lanes were loaded with 5% of the total lysate from each well. The membrane was cut and probed with affinity purified α-Cog4, diluted 1:200, and α-β-tubulin diluted 1:500 in 5% milk/PBST.

2.2 Aims

This thesis builds on extensive previous research in our lab to identify and begin to characterise the effects of the L36P mutation in Cog4. A key aim of the work presented in this chapter was to engineer the Cog4L36P point mutation under its endogenous promoter in a CHO-K1 cell line, using CRISPR/cas9, and adapt the cells from adherent to suspension culture. This cell line can later be tested as a host for biologic production, to assess the effects of the mutation on glycosylation of recombinantly expressed protein.

This aim of the work presented in this chapter was to generate and characterize a Cog4L36P cell line, to assess properties affecting the cells' suitability as biologic host cells. Work presented in this chapter includes characterisation of CRISPR-engineered Cog4L36P cells and the original knock-in ΔCog4L36P cell line. *In vitro* characterisation aims included testing the binding of COG complex containing Cog4L36P to Rab30 by GST pull-down. This experiment was necessary because the mutation was originally identified based on impaired Rab30 binding in a yeast-two-hybrid system, so it was important to validate this finding in the context of the whole COG complex expressed by a mammalian host. *In vivo* characterisation aims included analysing proliferation, viability, and metabolic activity of Cog4L36P compared to WT cells, to test whether the mutation impacts cell growth or metabolism, both of which are important properties for consideration of this cell line as a potential host for recombinant protein expression.

Another aim of *in vivo* characterisation experiments was to assess adherent properties of Cog4L36P cells, following the observation that the cells differ in their morphology compared to WT, appearing rounder and less elongated, suggestive of defective adhesion.

2.3 Materials

Table 2.1 | Summary of buffers used

Buffer Name	Composition
Phosphate Buffered Saline (PBS)	137 mM NaCl, 13 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 2.7 mM KCl (pH 7.4 – HCl)
Tris-acetate-EDTA (TAE) buffer	40 mM tris, 20 mM glacial acetic acid, 1mM EDTA
Fairbanks A	0.05% w/v coomassie brilliant blue (CBB), 10% v/v acetic acid, 25% v/v isopropanol in dH_2O
Fairbanks B	0.005% w/v coomassie brilliant blue (CBB), 10% v/v acetic acid, 10% v/v isopropanol in dH_2O
Fairbanks C	0.002% w/v coomassie brilliant blue (CBB), 10% v/v acetic acid in dH_2O
Fairbanks D	10% acetic acid
PBST	PBS with 0.1% Tween-20
Blocking	5% w/v dried milk in PBST
Cell Lysis	50mM Hepes-HCl (pH7.4), 200mM KCl, 1.5% Triton X-100
SDS Sample	5% glycerol, 50 mM Tris-HCl pH 6.8, 50 mM DTT, 1% SDS, 0.7 mM Bromophenol Blue
6 x SB	20mM Tris-HCl pH 8, 60mM EDTA ph 8, 0.1% SDS, 60 % glycerol, 0.2% orange G
GST Lysis	20mM Hepes-HCl (pH 7.4), 200mM KCl, 1mM EDTA, 2mM $MgCl_2$
Protease Inhibitor Cocktail	1mM DTT, 2µM Pepstatin A, 0.3µM Aprotinin, 1.2µM Leupeptin, 1mM PMSF supplemented with DNase I

Nucleotide Exchange (NE)	GST-lysis buffer + 3mM MgCl$_2$ (5mM total MgCl$_2$), 10mM EDTA
Nucleotide Binding (NB)	NE buffer lacking EDTA
GTP Binding	NB Buffer with 1mM GTP
GDP Binding	NB Buffer with 1mM GDP

Table 2.2 | Summary of Chinese hamster ovary (CHO-K1) cell lines used. All cells were kindly provided by the lab, and all non-wild-type cells were previously generated using the CRISPR-Cas9 technique (Cong et al., 2013). ΔCog4 cells were engineered by Estere Seinkmane and both knock-in cell lines were made by Sarah Tindall.

Cell Line	Gene Knockouts	Gene Knock-ins	Knockout Method	Knock-in Method
WT	-	-	-	-
ΔCog4	Cog4	-	CRISPR/Cas9	Stable Transfection
ΔCog4WT	Cog4	WT Cog4	CRISPR/Cas9	Stable Transfection
ΔCog4L36P	Cog4	Cog4L36P	CRISPR/Cas9	Stable Transfection

Table 2.3| Summary of all oligonucleotides used for CRISPR. ssgDNAs, a U6 fwd primer for PCR and an ssODN template, were purchased from Integrated DNA Technologies. BbsI overhangs for ligation of ssgDNA into cas9n PX462 are highlighted in blue. In the ssODN repair template, silent mutations that abolish protospacer adjacent motif (PAM) sites are highlighted in green and an additional silent mutation for abolishing a SacI restriction site is highlighted in yellow. Phosphorothioate bonds are indicated by *s.

Name	Sequence (5'-3')
gDNA 1.1	CACCGCGGTGGAGATTTCGGAGCAG
gDNA 1.2	AAACCTGCTCCGAAATCTCCACCG
gDNA 2.1	CACCGTCTGACAGAGCTGCAGGAGC
gDNA 2.2	AAACGCTCCTGCAGCTCTGTCAGA

U6 Fwd	GGACTATCATATGCTTACC
CRISPR Fwd	AACAACGAGGCCATCTCAAG
CRISPR Rev	TCCGCCTGATAGGAAGATGT
L36P ssODN	C*A*G*GAGCGCCTCCGCCACCGGAGGGTTTGCGA GGAGGGCACTGCTCGGAAATCTCCACCGAGCTTA TTCGCTCTCCGACAGAGCTGCAGGAGCTTGAGTC TGTGTACGAACGCCTCTGCGGCGAGGAGGTGGG GAGCTGATCTTCCTT*T*T*T

2.4 Methods

2.4.1 Adherent Cell Culture

Adherent CHO-K1 cells were cultured as previously described (Chatterton et al., 1999). CHO-K1 cells (table 2.2) were grown in adherent cell medium (Ham's F-12 medium, 5% v/v foetal bovine serum (FBS), 1% v/v penicillin/streptomycin (pen/strep- Gibco)) and kept at 37°C and 5% CO_2 in a 90% humidified incubator. Cells were passaged regularly by gentle washing with Dulbecco's phosphate-buffered saline (Gibco), followed by trypsinisation with trypsin/EDTA (Gibco) for 5 minutes at 37°C and re-suspension in adherent cell medium. All adherent cells were cultured in T75 flasks unless otherwise stated.

2.4.2 Suspension Cell Culture

Suspension cells were either cultured in OptiCHO medium (Gibco), supplemented with 8mM glutamax, 1% v/v pen/strep and 1:1000 anti-clumping agent (Gibco) or FortiCHO medium (Gibco) supplemented with 8 mM glutamax and 1% v/v pen/strep. Cells were cultured in either suspension T25 flasks for routine culturing or Erlenmeyer flasks (125 mL or 250 mL) for experiments requiring more cells. Flasks were incubated on an orbital shaker at 125 rpm in a humidified incubator with 5% CO_2. Cells were allowed to grow to a density of ~3 x10^6 cells/mL before passaging to 0.25 x10^6 cells/mL by transferring the required volume to a new flask and topping up to 30 mL with fresh medium. Cell density and viability were monitored regularly by incubating 300 μL cells with 300 μL TrypLE (Gibco) for 5 mins at 37°C before counting using a Vi-Cell XR automatic cell counter (Beckman Coulter).

2.4.3 Cell freezing and revival

Frozen stocks of cell lines were made regularly. For adherent cells, cultured in medium containing 5% FBS, cells were grown to 100% confluency in 10cm dishes and harvested by trypsinisation before pelleting at 1,000 $x g$ for 5 mins. Cells growing in suspension were counted using a Vi-Cell XR automatic cell counter to calculate the required volume for freezing 15 million cells per vial. Suspension cells were then harvested by centrifuging at 1,000 $x g$ for 5 minutes. Each cell pellet was re-suspended in 1.5mL freezing medium consisting of either FBS with 10% v/v DMSO or culture medium with 7.5 % v/v DMSO for adherent and suspension cells respectively. Cells were then transferred to cryovials in a room temperature Mr Frosty™ freezing container (Gibco), which was stored at -80 °C for at least 24 hours before transferring the vials to a sample box within a liquid nitrogen dewar.

For adherent cell revival, cryovials were thawed at room temperature before pipetting the vial contents into a 10cm dish or T75 flask containing 10 mL adherent cell medium. The following day, medium was aspirated from the cells and replaced with fresh adherent cell medium. For suspension cell revival, cryovials were thawed by pipetting warm suspension cell medium directly onto the frozen pellet. Vial contents were then aspirated gently into 5 mL medium using a serological pipette, before transfer to a 15 mL falcon tube and centrifugation at 1,000 $x g$ for 5 mins. The supernatant containing DMSO was then discarded, and cell pellets re-suspended in 30 mL suspension cell medium and transferred into 125 mL shake flasks.

2.4.4 Adaptation from adherent to suspension culture

For adaption to suspension culture, cells were cultured in T75 flasks in adherent media (Hams F-12, 1% pen/strep) containing decreasing concentrations of FBS, each for a period of 1 week with two 1:10 passages at each concentration. At the minimum FBS concentration required for adhesion (1% v/v), cells were grown to 100 % confluency in 4 x T175 flasks, harvested by trypsinisation and counted using a ViCell. The required volume for 30 million cells was centrifuged at 1,000 $x g$ for 5 mins, re-suspended in 30 mL medium containing 50% suspension cell medium (OptiCHO- Gibco, supplemented with 8 mM glutamax and 1% pen/strep) and 50% adherent medium (Hams F-12 with 1% FBS) for a density of 1 million cells per mL. The amount of adherent medium in the

mixture was gradually lowered, starting with 50% v/v and lowering to 25%, 10% and finally 0% (100% suspension medium), and cells were cultured in each medium for a minimum of 1 week (as described in section 2.4.2). At each stage of the adaptation process, frozen cell stocks were made as back-ups in case of contamination. The freezing procedure detailed in section 2.4.3 was followed with the exception that freezing medium consisted of fresh culture medium matching the composition of the medium cells were growing in at that point of the adaptation process, with 10% DMSO.

Cells were considered fully adapted once their viability stabilised above 95% in 100% suspension cell medium. This is summarised in the schematic shown in figure 2.2.
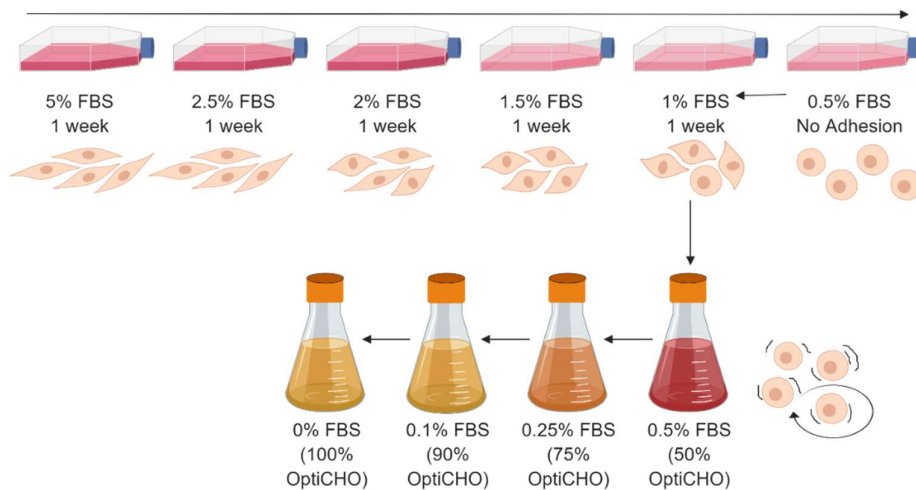
Figure 2.2 | Schematic representation of the adherent to suspension cell culture adaptation process.

Cells were initially adapted into OptiCHO medium with 8mM glutamax and 1% pen/strep. Cells were later adapted to grow in FortiCHO medium (Gibco), also supplemented with 8mM glutamax and 1% pen/strep, by thawing a cell pellet made from cells growing in 75% OptiCHO, 25% Hams F-12, 0.25% FBS into 90% FortiCHO, 10% Hams-F12, 0.1% FBS before continuing with the adaptation process using FortiCHO medium.

2.4.5 CRISPR-cas9n Genome Editing

2.4.5.1 CRISPR gDNA and Repair Template Design

A CRISPR point-mutation strategy was designed based on guidelines in (Ran et al., 2013b) utilising Cas9n PX462 (Addgene) plasmids. Single-stranded (ss) 20 bp gDNAs and a 151 bp repair template containing the L36P point mutation in the centre were designed using CHO-K1 genomic DNA as a template. DNA sequences were designed and visualised using SnapGene software and guides were tested for theoretical off-target effects using CRISPR guide validation software by Synthego. All oligonucleotides are listed in table 3.

2.4.5.2 Ligation of gDNA into PX462 plasmids

For ligating complementary ssgDNAs and cloning into cas9n PX462 plasmids, protocols detailed in (Ran et al., 2013) were followed. Firstly, 10 µL reaction mixtures consisting of 10 µM of each complementary single stranded guide oligo, T4 polynucleotide kinase (PNK) and T4 PNK buffer in ddH$_2$O were prepared in PCR tubes. Reactions were then carried out in a thermocycler using the following parameters: 37 °C for 30 min; 95 °C for 5 min; ramp down to 25 °C at 5 °C min^{-1}. The resulting annealed, phosphorylated oligos were diluted 1:200 with ddH$_2$O before cloning into two cas9n PX462 plasmids. Digestion of the cas9n plasmids and ligation of the guide DNAs into the digested plasmid were carried out in parallel, using a 20 µL reaction mixture consisting of 100 ng plasmid DNA, 2 µL diluted oligo duplex, Tango buffer (Thermo Scientific™), 500 µM DTT, 500 µM ATP, 1 µL fast digest BbsI and 0.5 µL T7 ligase in ddH$_2$O. The reaction mixture was incubated for 1 hour in a thermocycler, using 6 cycles of 37 °C for 5 mins followed by 21 °C for 5 mins.

2.4.5.3 Transfection

For co-transfection of cas9n with a single stranded oligodeoxynucleotide (ssODN) repair template, a transfection mixture of both cas9n plasmids and the repair template was mixed in the ratio advised in (Ran et al., 2013), which equated to 0.9:1 ssODN to plasmid DNA up to a total of 2.5 µg DNA per transfection. Cells were seeded at 60% confluency in a 6-well plate and transfected using Xfect reagent, following the manufacturer's protocol (Clontech). A control transfection was also done using the same amount of DNA but with a mixture of eGFP plasmid and ssODN repair template.

24 hours after transfection, CRISPR-transfected cells were incubated in medium containing 15 µg/mL puromycin for 72 hours.

2.4.5.4 Clonal Isolation and Expansion

48 hours post-transfection, eGFP transfected control cells were visualised using a fluorescent microscope to estimate transfection efficiency. After puromycin selection, CRISPR treated cells were washed 3 times in DPBS (Gibco), trypsinised and concentrated from 3 mL into 500 µL media and counted using a hemacytometer. Clonal isolation was performed by serial dilution in medium containing 60% filtered spent media collected previously from WT CHO-K1 cells. Cells were seeded in five 96-well plates at a density of 0.6 cells per well (100 uL medium volume). 10 days after seeding, colonies originating from a single cell were identified and maintained with bi-weekly medium changes until they reached confluency. Confluent cells were expanded into 24-well plates for genotyping before further expansion of L36P positive cells.

2.4.6 Molecular Biology Methods

2.4.6.1 *E. coli* Transformation

For DNA purification and protein expression, competent DH5α cells (New England Biolabs) and BL21 cells respectively were transformed according to manufacturer's protocol. ~100ng of plasmid DNA was added to 50µL of *E. coli* in a transformation tube and incubated on ice for 30 minutes. Bacteria were then heat shocked for 30 seconds at 42°C and returned to ice for 5 minutes. 950µL of room temperature SOC media was added before a 1 hour shaking incubation at 37°C. *E. coli* were plated on selective LB plates [2% Agar, 2% Lysogeny Broth (LB)], at a range of dilutions from 1-90% and spread with glass beads. Plates were incubated at 37°C overnight.

2.4.6.2 DNA Purification

For extracting CHO cell genomic DNA, QIAGEN's DNeasy blood and tissue kit was used following the manufacturers protocol. DNA was extracted from confluent wells of a 24-well plate by trypsinising cells and harvesting a cell pellet by centrifugation at 1000 *x g* for 5 minutes. Cell pellets were re-suspended in 200 µL PBS before proceeding with the kit protocol.

For extracting plasmid DNA from *E. coli*, 5mL LB [2% w/v LB in ddH$_2$O] containing the required selection antibiotic was inoculated with a single colony from a plate of *E. coli* cells transformed with the DNA of interest and incubated at 37 °C overnight. For sequencing, DNA was extracted using a Promega mini-prep kit according to the manufacturer's protocol. For mammalian cell transfection, DNA was prepared using a Qiagen midi-prep plasmid purification kit. This required a second overnight *E. coli* culture of 100 mL volume, which was inoculated with 500 μL *E. coli* from the initial 5 mL overnight culture, before preparing DNA following the manufacturer's protocol. Concentration and purity were measured spectrophotometrically using a nanodrop.

2.4.6.3 Restriction Digestion

Restriction digests were carried out according to general molecular biology practices, using the recommended buffer in a 30 μL reaction volume with 1 μL enzyme to minimise star activity. For CRISPR genotyping, analytical digests were performed using between 0.5-1 μg DNA digested with SacI (NEB) enzyme in CutSmart buffer (NEB) overnight at 37°C.

2.4.6.4 DNA Gels and Gel Purification

DNA samples were prepared by diluting with 6 x concentrated sample loading buffer (6xSB). Samples were loaded in agarose gels containing between 1-3% agarose powder dissolved in TAE buffer with SYBR Safe dye. DNA gels were run at 100 V for 20-40 minutes (until required separation was achieved).

2.4.6.5 PCR and Sequencing

For CRISPR genotyping, PCR using CHO-K1 genomic DNA as a template was done for an 832 bp region of DNA surrounding the CRISPR cut site, with primers (table 2.1) designed to bind to genomic DNA. For each potential mutant cell line, a 50 μL reaction was set up and run alongside controls containing WT genomic DNA and ddH2O as templates. Each 50 μL reaction contained the following: 400 nM of each CRISPR forward and reverse primer, ~100 ng of genomic DNA and 25 μL 2X PCR Bio Ultramix (PCR Biosystems).

For verifying ssgDNA incorporation into cas9n vectors, colony PCR was done on DH5α cells transformed with either px462 plasmids ligated with gDNAs 1/2 or px462 with no

insertion. Several colonies from each plate were picked and added to 30 µL PCR reaction mixtures, containing 1 µL taq polymerase (purified by Dani Ungar), taq buffer, 200 µM dNTPs and 500 nM of each primer including U6 fwd and either gDNA1.2 or gDNA2.2 oligos as the reverse primer. Reactions were run with the settings summarised in table 2.4.

Table 2.4 | Stepwise summary of thermocycler settings for CRISPR genotyping PCR and Cas9n colony PCR reactions.

	CRISPR Genotyping PCR	Cas9n Colony PCR
1	95°C, 1:00 min	95°C, 5:00 min
2	95°C, 0:15	95°C, 0:30
3	53°C, 0:15	55°C, 0:30
4	72°C, 1:00	72°C, 1:00
5	Repeat 2-4, 35x	Repeat 2-4, 30x
6	72°C, 10:00	72°C, 10:00
7	4°C, ∞	4°C, ∞

Reaction products were purified using a QIAquick PCR purification kit, following the manufacturers protocol (QIAGEN).

All sequencing was done by sending samples to Eurofins sequencing facility using their TubeSeq service.

2.4.7 GST Pull-down

2.4.7.1 Cell Lysis for GST Pull-down

Cells were cultured in T-175 flasks, harvested by trypsinisation at 90-100% confluency and pelleted (1000 x g for 5 mins). Lysis was carried out at 4 °C. Pellets were washed 6x with 30 mL cold phosphate-buffered saline (PBS). Pellets were re-suspended in cell lysis buffer (1 mL per 4.5 x 10^7 cells) supplemented with 1 mM PMSF and rotated for 30 mins. Lysates were cleared by centrifugation (16,000 x g for 45 mins).

2.4.7.2 Protein Expression in E. coli

GST-Rab30 constructs (Miller et al., 2013), were expressed by induction with 0.5 mM IPTG at A_{600} = 0.8, followed by overnight incubation at 18°C for GST-Rab30 or 3.5 hour 37 °C incubation for GST. After incubation, cells were harvested by centrifugation (5000 x *g* for 10 mins at 4 °C), re-suspended in *E. coli* GST-lysis buffer, split into aliquots and re-pelleted by centrifugation (5000 x *g* for 10 mins at 4 °C). Pellets were stored at -80°C.

2.4.7.3 E. coli Lysis for GST Pull-down

Pellets of GST or GST-Rab30 expressing *E. coli* were re-suspended in a volume of *E. coli* GST-lysis buffer equivalent to a tenth of the original culture volume, with additional protease inhibitor cocktail. Cells were lysed by passage through a 4 °C cell disruptor at 25 psi. Lysates were cleared by centrifugation (15,000 x *g* for 15 mins at 4 °C).

2.4.7.4 GST Pull-down Experiments

All steps were performed at 4 °C. Glutathione-agarose (Clontech), equilibrated with GST-lysis buffer was added to cleared *E. coli* lysates and rotated for 1 hour. Beads were pelleted by centrifugation (8,000 x *g* for 5 mins), and washed 4 times with GST-lysis buffer, followed by a wash with nucleotide exchange (NE) buffer. Washed GST-only control beads were re-suspended in NE buffer containing 1 mM GTP whilst GST-Rab30 bound beads were split into two fresh Eppendorf tubes, centrifuged as before, and re-suspended in NE buffer containing either 1 mM GDP or GTP and incubated overnight.

The next day, beads were split again, washed with nucleotide binding (NB) buffer, and incubated in either GDP or GTP binding buffer for 2 hours. Beads were split into tubes for each CHO cell line and pelleted by centrifugation (8000 x *g* for 5 mins) before resuspension with 600 μL CHO cell lysate, 1 mM $MgCl_2$ and 1 mM GDP or GTP. Beads were rotated in lysate for 1 hour before washing in NB buffer with 0.1% v/v Triton x-100 added). Bound proteins were eluted and analysed.

2.4.8 SDS-PAGE and Coomassie® Staining

All SDS samples were prepared by mixing samples with 2% SDS sample buffer in microfuge tubes before heating and cooling as described above. Prior to loading onto a gel for SDS-PAGE, all samples were sheared by syringing up and down with a needle.

For SDS-PAGE, 10–15 µL SDS samples (diluted with additional 2% SDS sample buffer where necessary) were added to wells of 8%-12% acrylamide gels. The gel was then run at 120 V until the bands reached the stacking gel and 180 V until the dye reached the bottom.

For complete protein staining on SDS-PAGE gels, a heat accelerated Coomassie® blue staining and de-staining protocol was followed (Fairbanks et al., 1971). Gels were soaked in Fairbanks A solution and boiled in a microwave for approx. 1 minute before cooling on a shaker at room temperature, washing with water and replacing the solution with Fairbanks B. This process was then repeated through Fairbanks C and D solutions, before leaving the gel to de-stain in Fairbanks D.

2.4.9 Western Blotting

Proteins, from an SDS-PAGE, gel were transferred to a polyvinylidene difluoride (PVDF) membrane. The membrane was blocked in blocking buffer before incubation with primary antibodies in blocking buffer. Sources and dilutions of primary antibodies used are indicated in relevant figure legends. Both blocking and primary incubation were carried out either for 1 hour at room temperature or overnight at 4 °C. The membrane was then washed three times with blocking buffer and incubated with secondary antibody in blocking buffer for 1 hour at room temperature. Finally, the membrane was washed three times with blocking buffer and three times with PBST before imaging using Immobilon HRP substrate (Millipore) and the Syngene GeneGenius ECL imaging system. Quantification of Western blot images was performed using ImageJ software.

2.4.10 Adhesion Assays

Adherent CHO-K1 cells were expanded into T175 flasks prior to assaying their adhesion. Cells were allowed to grow to 80-100% confluency before harvesting them either by trypsinisation (see section 2.4.1) or by EDTA release. EDTA release involved washing the cells with PBS before adding filtered 10 mM EDTA and incubating at 37 °C for 10 mins. Trypsin released cells were re-suspended in medium and EDTA released cells were pelleted at 1,000 x g for 5 mins before removing the EDTA and replacing with medium. Cells were counted on a Vi-Cell XR automatic cell counter to calculate the volume required for adding either 600,000 cells per well of a 6-well plate or 120,000 cells per well of a 24 well plate. Wells were then topped up to 1 mL or 3 mL for 24 and

6 well plates respectively with adherent cell medium. For testing their adhesion to plastic, cells were incubated for 1.5 hours at 37 °C to allow some to adhere. Non-adhered cells were then removed by washing wells three times with PBS. Remaining cells were trypsinised and counted using either a Vi-Cell for 6-well plate assays or a hemacytometer for 24 well plate assays. Assays in 24-well plates were performed by Maya Trivedy under the author's supervision.

For coating cells in ECM, 120,000 or 600,000 trypsin-released cells were added to wells of a 24-well or 6-well plates respectively and incubated overnight at 37 °C. The following day, cells were released with EDTA, following the procedure outlined above. Wells were then thoroughly washed three times with warm PBS to remove any remaining cells whilst leaving the ECM in-tact. These ECM-coated wells were then used for adhesion assays with trypsin-released cells, as outlined above.

2.4.11 Proliferation and Viability Assays

For proliferation and viability assays, cells were counted daily using a Vi-Cell cell counter. For adherent cells, this involved trypsinising (see section 2.4.1) and counting cells from T75 flasks before seeding multiple wells of a 6-well plates with 30,000 cells per well. Two 6-well plates were filled per cell line and three wells were harvested for counting daily at 24, 48, 72 and 96 hours. For suspension cells, these were passaged to 0.25×10^6 cells/mL in T25 flasks. 300 µL samples were taken daily for counting (see section 2.4.2).

2.4.12 Metabolic Activity Assays

2.4.12.1 MTT Assay

Cells growing in 10 cm dishes were trypsinised and counted using a hemacytometer before seeding into 96-well plates at a density of 5,000 cells/well and incubated at 37 ˚C and 5% CO2 atmosphere. Media was removed from all 96 wells after 24 hours and 100 µl media and 25 µl MTT (SIGMA M5655, 5mg/ml in PBS) were added to the wells to be tested that day. Only medium (no MTT) was added to negative control wells. Prior to each time point, cells were incubated with MTT or medium only for 3 hours in 37 ˚C and 5% CO2 atmosphere, followed by the removal of the solution and the addition of 100 µl acidic isopropanol (0.4M HCl in absolute isopropanol). The absorbance was then

read at 570 nm. This was repeated at 48, 72 and 96 hours. This assay was performed by Emily Priest under the author's supervision

2.4.12.2 AlamarBlue Assay

Suspension cells were counted and seeded in 40 mL suspension cell medium at 0.25 x 10^6 viable cells/mL (see section 2.4.2). Every day, 1 mL of cells were removed, centrifuged at 1000 x g for five minutes and re-suspended in 1 mL 10% AlamarBlue in suspension culture medium. Cells were incubated at 37 °C, shaking at 125 rpm, for 90 minutes. After incubation, cells were transferred to microfuge tubes and centrifuged at 1000 x g for five minutes to separate cells from the medium. Medium was transferred to wells of a black clear bottomed 96 well plate and fluorescence intensity was measured using an excitation wavelength of 545 nm and emission wavelength of 600 nm using a CLARIOstar plate reader.

2.5 Results

2.5.1 Engineering a Cog4L36P CHO-K1 Cell Line Using CRISPR-Cas9n

2.5.1.1 Experimental design

A CRISPR/cas9 editing strategy to introduce the L36P point mutation in CHO-K1 COG4 genomic DNA was designed following guidelines set out in (Ran et al., 2013b) (figures 2.3-2.4). This method utilises a cas9 nickase (cas9n) mutant, which has a D10A point mutation in the active site that partially inhibits the nuclease activity of the enzyme, resulting in a single stranded nick as opposed to a double stranded break (Ran et al., 2013a). Breaking both strands of DNA therefore requires the action of two cas9n enzymes on opposing strands, in close proximity. Each cas9n enzyme uses a different guide RNA to bind to the target region of DNA, meaning off-target effects are reduced as the chances of the same two target sequences being present near to each other elsewhere in the genome are very small. Introducing a point mutation is possible by using cas9 nickases and a single stranded repair template, containing the desired point mutation and DNA either side, which is complementary to the region of interest. This enables the cell to repair the damaged DNA using homology directed repair (HDR), incorporating the repair template into the genome in the process (figure 2.3).
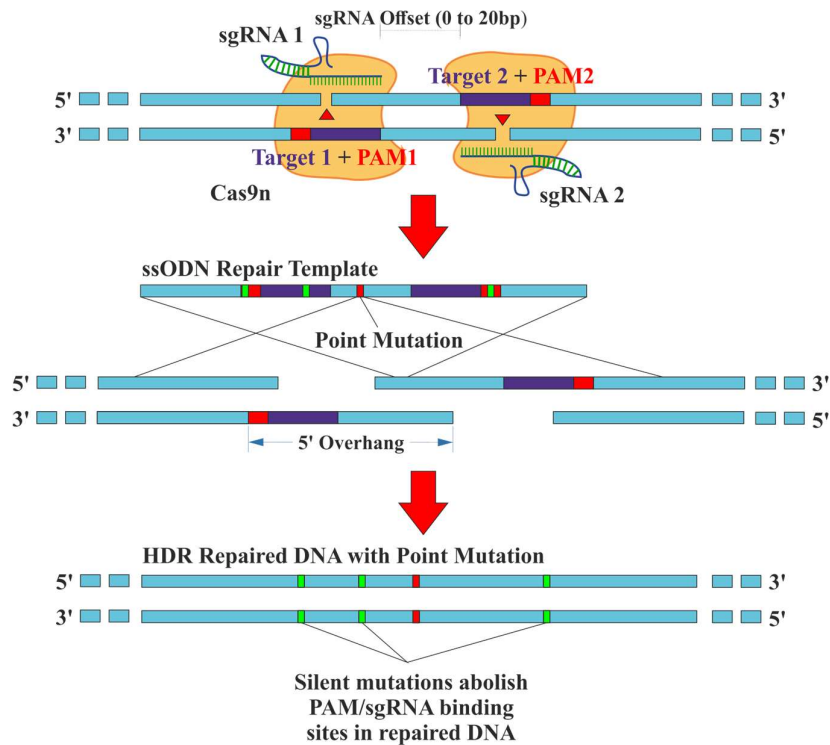
Figure 2.3 | Schematic representation of the strategy used to create a point mutation using CRISPR/Cas9n homology directed repair (HDR). Two Cas9n enzymes are directed to either strand of DNA using a pair of 20nt sgRNAs, which are complementary to target sections of DNA (purple) adjacent to PAM sites (red) (5'-NGG-3'). sgRNAs are positioned 0-20bp apart in close proximity to the nucleobase selected for point mutation. Cas9n enzymes each cut a single DNA strand, resulting in cleaved DNA with 5' overhangs. HDR is enabled by the use of an ssODN repair template, which undergoes recombination with the damaged section of DNA. The ssODN contains the point mutation of interest, along with silent mutations to abolish PAM sites and prevent re-cleavage of repaired DNA by Cas9n.

It is, however, possible that persistent cas9n could cut DNA again, even after repair by HDR. Therefore, silent mutations at each of the PAM recognition sites were incorporated into the repair template to prevent further binding of cas9n to the region of interest (figures 2.3-2.4). Base changes for silent mutations were carefully chosen to minimise changes to mRNA stability. Another key element in the design of this CRISPR-editing strategy is the incorporation of an additional silent mutation in the repair template, which abolishes a SacI restriction site in the WT CHO-K1 genomic DNA. In theory, by amplifying a region of genomic DNA surrounding this site and digesting it with SacI, CRISPR edited cells could be identified based on a distinct DNA gel band, with no digestion product bands after SacI treatment.
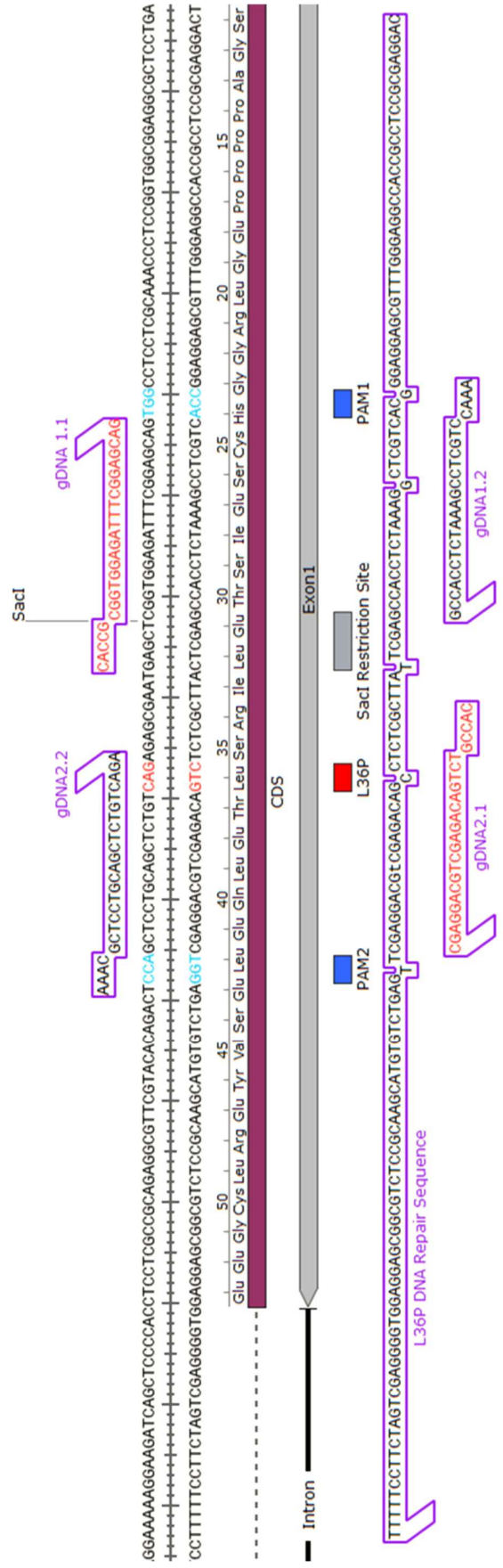
Figure 2.4 | SnapGene diagram showing the position of sgRNAs and ssODN repair template within genomic CHO-K1 COG4 DNA. A 151base ssODN spans exon 1, containing leucine 36, and adjacent intronic DNA. The T-C substitution responsible for L36P is centred within the ssODN with 75bp homology arms either side. All mutations within the template, including L36P and silent mutations, are highlighted by their offset position in relation to the rest of the strand. A SacI restriction site, which will be used for genotyping, is labelled, and highlighted in grey. PAM sites are highlighted in blue and the L36P mutation shown in red. Each single stranded gDNA has a complementary strand, each with overhangs for cloning into a commercial Cas9n vector (px462) digested by BbsI.

114

A control experiment (figure 2.5) with WT CHO-K1 cells was carried out to test the process of genotyping cells using the planned restriction fragment analysis. The aim was to test whether sufficient DNA could be extracted from cells grown in a 24-well plate and to test the PCR primers (table 2.3) designed for amplifying genomic DNA. Results validate the efficacy of this method. The analytical gel image (figure 2.5) clearly shows a distinct 834 bp uncut product, and 2 digested products at 522 bp and 312 bp, which are easy to distinguish from uncut DNA. This suggested it would be possible to screen for unedited WT clones and HDR edited CRISPR clones containing the L36P mutation and lacking the restriction site.
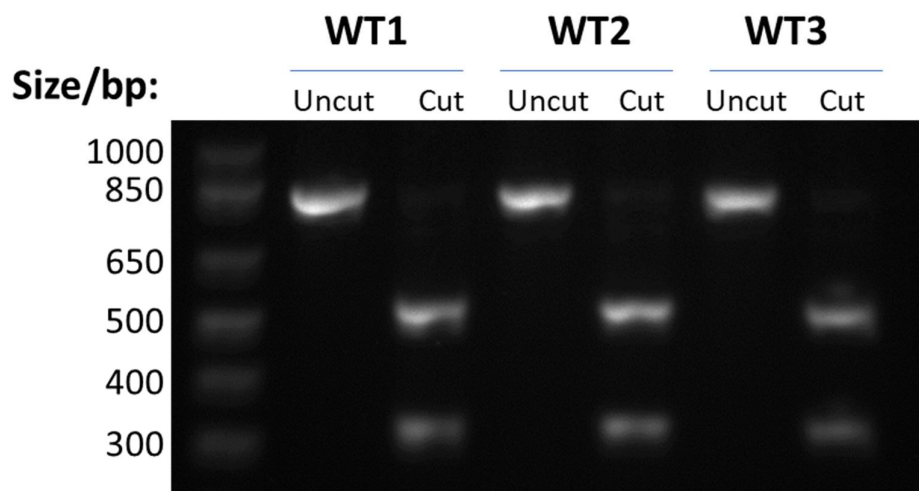
Figure 2.5 | SacI restriction digest of PCR amplified WT CHO-K1 Genomic DNA. WT CHO-K1 genomic DNA from 3 wells of a confluent 6-well plate was amplified using QIAGEN DNeasy kit. PCR primers, CRISPR Fwd and CRISPR Rev, (table 2.3) were designed to bind to intronic DNA either side of COG4 exon 1; amplifying an 834 bp strand of genomic DNA centred asymmetrically on a single SacI restriction site. PCR purified DNA was digested overnight with SacI and analysed on a 1.5% agarose gel.

2.5.1.3 Identification of L36P positive CRISPR clones

Following successful validation of the genotyping approach using non-transfected cells, WT CHO-K1 cells were co-transfected with two cas9n plasmids, each containing a different guide RNA sequence, and the ssODN repair template. Clones were then isolated using serial dilution. Twenty-two clones survived and all were genotyped as planned: using restriction fragment analysis (figure 2.6).
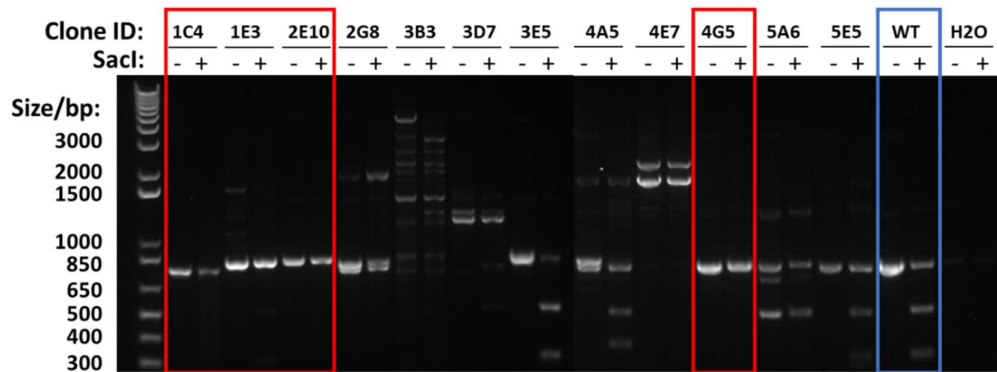
Figure 2.6 | Restriction fragment analysis of twelve CRISPR colonies. The first twelve confluent colonies in a 24-well plate were screened for CRISPR HDR editing using restriction fragment analysis. An 834bp strand of genomic DNA, centred asymmetrically on a single SacI restriction site was amplified and PCR purified. DNA was digested overnight with SacI and analysed on a 1.5% agarose gel. Suspected positive clones are boxed in red and the WT control is highlighted in blue. As an additional control for the PCR reaction, ddH2O was also used as a template to rule out contamination.

Clones 3E5, 4A5 and 5E5 generated digestion fragments like WT, suggesting HDR editing had not occurred in these cells as they retained their SacI restriction site. Other clones, including 3D7, 4E7, and 5A6 gave completely different restriction patterns, potentially indicating that their DNA had been edited in a different way by the CRISPR process. Multiple DNA fragments present in clone 3B3 suggested the PCR reaction didn't work properly for this DNA sample. DNA from four clones resisted SacI digestion and generated a digestion pattern similar to uncut WT (figure 2.6); indicating that at least the silent mutation in the SacI site had been incorporated by HDR with the ssODN

116

template. These four clones were sent for sequencing and both 2E10 and 1E3 came back with three mismatches against the WT COG4 sequence. The repair template contained five mutations in total, so only three mismatches indicated that the ssODN repair template was only partially incorporated by HDR. Nevertheless, both clones were positive for the L36P point mutation (figure 2.7). DNA from clone 1C4 ran at a lower molecular weight than the other clones (figure 2.6) and sequencing confirmed the region of DNA containing the leucine 36 site and the SacI restriction site was missing (data not shown), suggesting that the repair template had not been incorporated; the DNA was likely repaired by non-homologous end joining instead, causing a deletion.
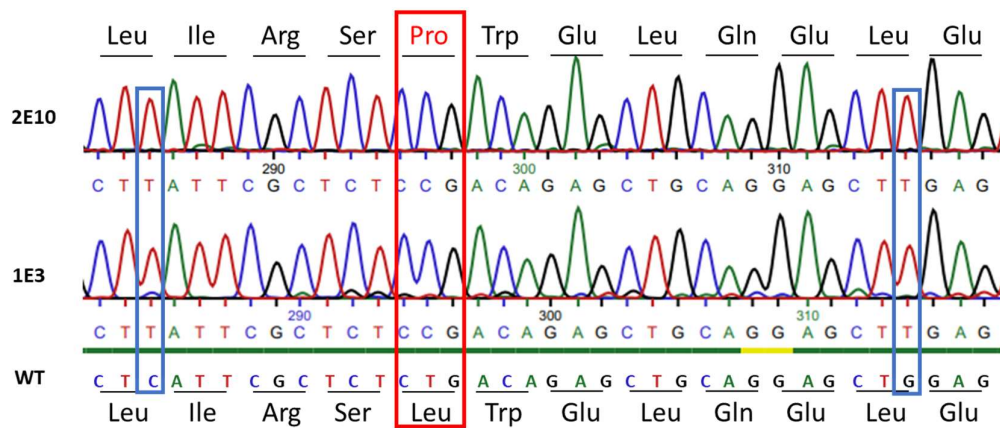
Figure 2.7 | Sequence traces of L36P positive clones. Purified PCR products for all suspected L36P cells were sent for sequencing (Eurofins TubeSeq) with L36Pfwd and L36Prev primers (table). Sequences were run through NCBI Nucleotide BLAST against the Chinese Hamster genome and both 2E10 (top) and 1E3 (bottom) aligned with COG4, with 3 mismatches. The L36P mutation is highlighted in red and silent mutations are highlighted in blue. For comparison, the WT COG4 sequence is written below the sequence traces.

2.5.1.4 Adaptation to suspension culture

Clone 2E10 was taken forward for further characterisation, due to uncertainty over whether clone 1E3 was derived from a single cell. Since the purpose of generating this cell line is to test whether it is a suitable host for generating glycan engineered mAbs, the cells were adapted from growth in adherent culture to suspension culture. One

reason for this is that adherent medium includes FBS, which contains bovine IgG that cannot easily be separated from recombinantly expressed mAbs, whereas suspension medium has no FBS. Furthermore, cells can grow at a much higher density in suspension culture, which is favorable for maximising biologic production. Initially, cells were adapted to grow in OptiCHO medium, which was readily available in our lab. However, subsequent cell culture was done in FortiCHO medium, which is optimised for CHO-K1 suspension cell growth and recombinant protein expression.

Since the reason for engineering the Cog4L36P mutation using CRISPR-cas9 was because the existing knock-in ΔCog4L36P cells overexpress Cog4 compared to WT and ΔCog4WT cells, levels of Cog4 in the 2E10 Cog4L36P suspension-adapted cells were analysed alongside WT cells by Western blot (figure 2.8). Precise quantification is hampered by the overexposed loading control, however quantification of the Cog4 blot alone shows that the Cog4 bands at 89 kDa are of similar intensity in both WT and Cog4L36P derived lysates. Although there is less Cog4 in the Cog4L36P lysates than in the WT lysates, the difference between WT and Cog4L36P not statistically significant (t-test; $p > 0.05$) and is much smaller than the difference observed in the knock-in cells (figure 2.1). Higher molecular weight bands are also visible on the Cog4 blot; however, these are likely non-specific bands due to the promiscuity of the affinity-purified antibody. Comparable Cog4 western blot signals suggest Cog4 protein levels are similar in both WT and Cog4L36P cells. This validates the use of CRISPR-cas9 for engineering the Cog4L36P mutation under its endogenous promoter.
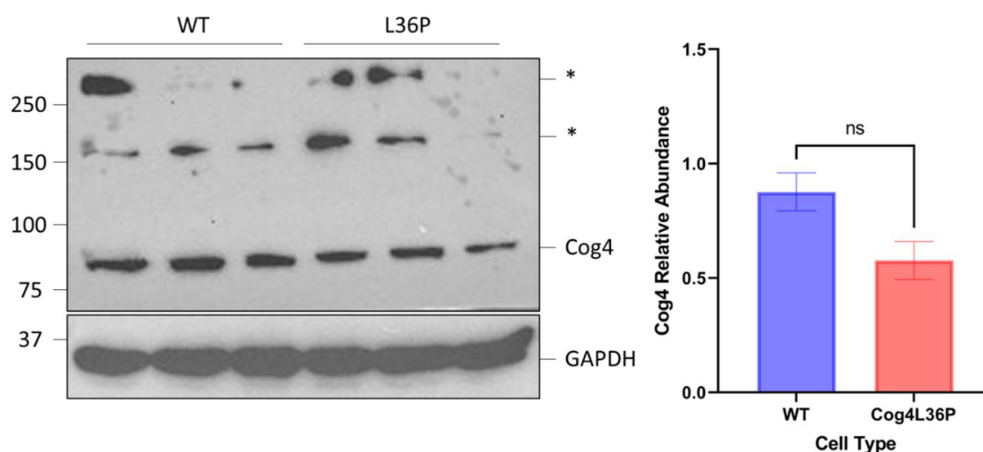
Figure 2.8 | Western blot showing Cog4 levels in WT and 2E10 L36P suspension-adapted cells. 200 μL SDS-PAGE samples were prepared containing a total of 1 million cells from each cell-line. Samples were prepared in triplicate and a 10% acrylamide gel loaded with 10 μL of each, totaling three lanes per cell-line. After transfer, the membrane was cut and the top half probed with 1:500 affinity purified rabbit α-Cog4 (purified by the author on a Cog4 affinity column made by Prof. Dani Ungar), and the bottom half probed with mouse 1:200,000 α-GAPDH (Applied Biosystems AM4300). Asterisks indicate non-specific protein bands. Cog4 bands were quantified using image J and averaged for each cell type. Error bars represent +/- 1 standard error of the mean. An unpaired t-test was performed to assess statistical significance (p > 0.05, not significant).

2.5.2 *In Vitro* Characterisation of the Cog4L36P Mutation

2.5.2.1 L36P mutation in Cog4 impairs binding to Rab30

One of the first questions to address with this cell line was whether the Cog4L36P binds Rab30, as the mutation was originally identified based on its lack of interaction with Rab30 in a yeast2hybrid experiment. It was imperative to establish this using endogenously expressed Cog4, in the context of the whole COG complex, as recombinantly expressed Cog4 in a yeast-2-hybrid experiment is not representative of biological conditions. Prior to development of the Cog4L36P CRISPR cell line, Cog4-Rab30 binding was tested using knock-in adherent cells (ΔCog4 L36P and ΔCog4 WT). A GST-pull-down assay was used to analyse binding of WT Cog4 and Cog4L36P with either GST-Rab30 or a GST-only control (figure 2.9).

119
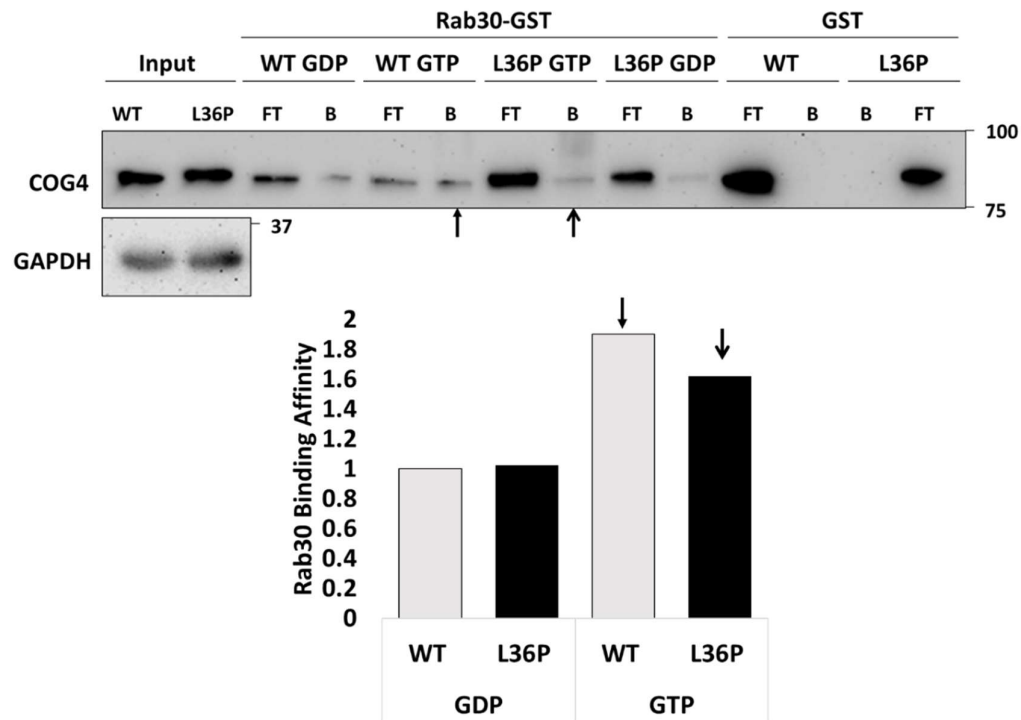
Figure 2.9 | Cog4 interactions with Rab30: comparison of L36P mutant and wild-type Cog4 knock in cells. Top panel: image of a single Western blot loaded with flow-through (FT) and bound (B) samples from each experimental condition of the GST pull-down assay. GST tagged Rab30 was purified by binding to glutathione beads, which were incubated first with GDP or GTP to lock Rab30 into an inactive or active state respectively. Rab30-bound beads were then incubated with whole cell lysates from ΔCog4WT and ΔCog4Cog4L36P CHO-k1 cells. Unbound protein was removed (FT), beads were washed and an SDS sample was prepared from the remaining bead-bound protein. FT samples were diluted 50-fold before loading alongside bound samples as Cog4 binding is very weak, so only a small fraction of the input Cog4 is collected in the bound sample, and most is retained in the flow-through. Membranes were blotted with α-Cog4 (affinity purified by the author) and α-GAPDH (Ambion AM4300) antibodies. Bottom panel: Quantitation of Cog4 in the bead samples on the blot was carried out using ImageJ and plotted for each Rab30 binding experiment. Arrows correspond to arrows on the blot images for clarity.

Quantitation of this experiment suggests that Cog4L36P interacts more weakly with Rab30 than does WT, as indicated by less Cog4L36P in the bound sample (figure 2.9). This effect appears to be GTP dependent as there was no noticeable difference in binding to the GDP bound form of Rab30 between WT and Cog4L36P. Results also appear to show a noticeable increase in concentration of Cog4L36P in the flow-through (FT) sample compared to WT, despite an equal input Cog4 concentration (figure 2.9). This also suggests there is a weakened interaction with Rab30, as Cog4L36P is more easily washed away from the Rab30- GST beads.

However, it was important to repeat this experiment using the CRISPR engineered Cog4L36P cell line, which expresses Cog4 to the same level as WT cells, in contrast to the knock-in cells that overexpress Cog4 (figure 2.1). Despite differences between the cells, it was hypothesised that the interaction between Cog4L36P and Rab30 would still be weaker in these cells than WT, as the Cog4L36P protein structure in both the CRISPR and knock-in cells should be the same. The experiment was repeated with the CRISPR-engineered cell-line following adaption from adherent to suspension culture, as cells could be harvested more efficiently in higher numbers.

As observed in the knock-in cells, the interaction between Cog4 and Rab30 was found to be weaker in the CRISPR engineered Cog4L36P cell line compared to WT (figure 2.10). In fact, no Cog4 was detectable in the Cog4L36P bound bead sample (figure 2.10). A GST control blot confirmed that levels of GST- tagged Rab30 were comparable between WT and Cog4L36P cells, validating that the lack of detectable Cog4 was due to an inhibited interaction, as opposed to less Rab30 in the bead sample.
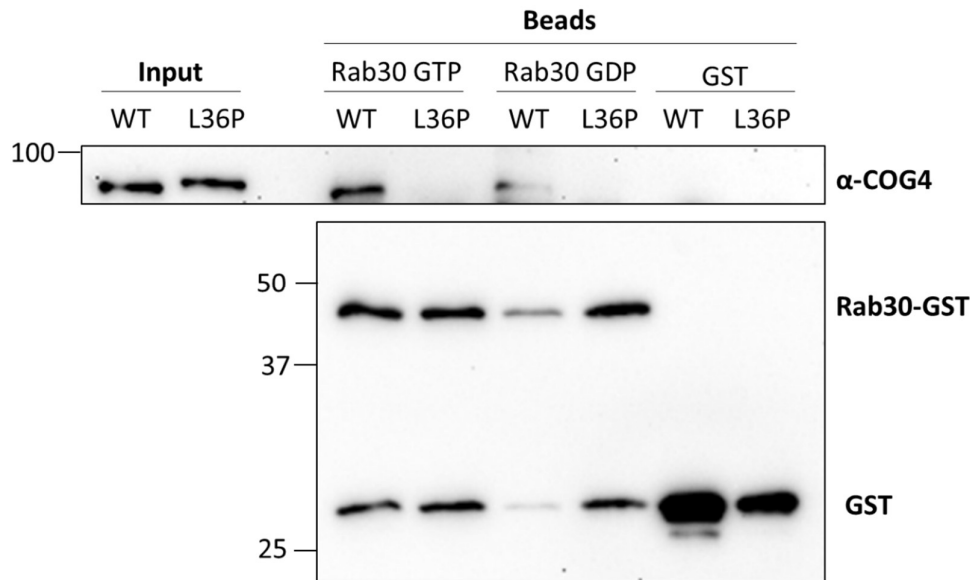
Figure 2.10 | Cog4 interactions with Rab30: comparison of CRISPR engineered Cog4L36P and wild-type Cog4 cells. A representative immunoblot, probed with α-Cog4 antibody, loaded with samples from each experimental condition of the GST pull-down assay. GST and GST tagged Rab30 were bound to glutathione beads, loaded with GDP or GTP and incubated with or without whole cell lysates from WT and Cog4L36P CHO-K1 cells as indicated. 3 µL of each input sample was loaded alongside 15 µL of each bead sample. Bead samples were also diluted 100-fold and 1 µL of each was loaded onto a gel for a second blot and probed with an anti-GST antibody.

2.5.3 *In Vivo* Characterisation of Cog4L36P Cells

2.5.3.1 Proliferation is impaired in Cog4L36P cells

When culturing Cog4L36P cells alongside WT, it became apparent that their growth rates differed. In monolayer culture, Cog4L36P cells took longer to reach confluency than WT. This appears to be the case for both the knock-in cell line and the CRISPR engineered cell line. Before generating the Cog4L36P CRISPR cell line, an undergraduate project student, Emily Priest, under supervision of the author, characterised proliferation of ΔCog4WT and ΔCog4L36P CHO-K1 cells in monolayer. She found that the ΔCog4L36P cells had a significantly reduced growth rate compared to ΔCog4WT (data not shown). ΔCog4WT cells doubled in 19.8 hours, compared to 27.4 hours for ΔCog4L36P.

Growth curves for WT, ΔCog4 and CRISPR-engineered Cog4L36P adherent cells were also measured by the author and fitted to exponential growth curves (figure 2.11). Average doubling times were calculated as 18.5, 22.9 and 18.9 hours for WT, Cog4L36P and ΔCog4 cells respectively. Interestingly, the WT and ΔCog4 cells grew at a similar rate, whilst the Cog4L36P cells proliferated much slower than either of the other cell lines. This suggests that the Cog4L36P mutation causes a more severe adherent growth defect than a complete knock-out of Cog4.
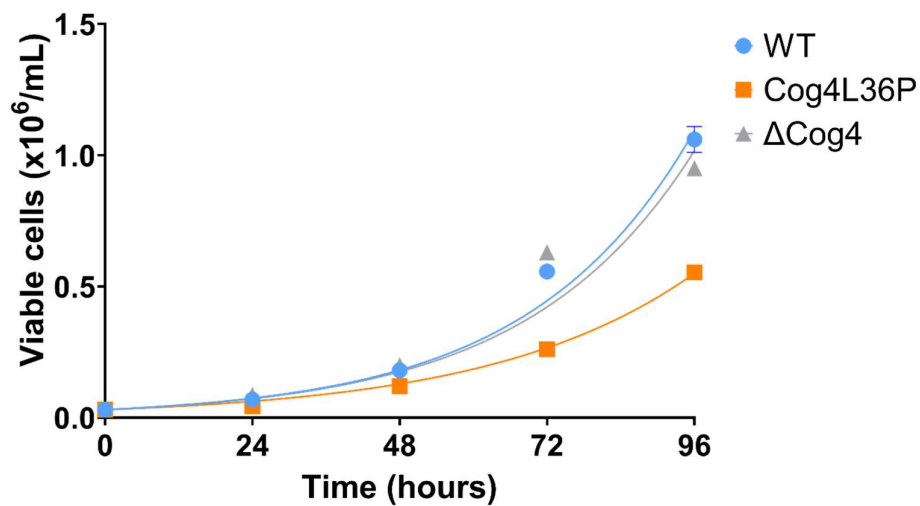
Figure 2.11 | Rate of Proliferation of WT, Cog4L36P and ΔCog4 adherent cells. All cell lines were seeded in triplicate wells at 30,000 cells/well onto culture dishes, incubated for 24, 48 or 72 hours, trypsinised, re-suspended, and counted using by a Vi-cell cell counter. Data were plotted and fit using GraphPad Prism software. Datasets were fit to an exponential, Malthusian growth curve of the form $Y = Y0*e^{kx}$, where Y0 is the starting population, which was assumed to be 30,000 cells, k is the rate constant and x is the time in hours. R^2 values for each curve were 0.9741, 0.9910 and 0.9192 for WT, Cog4L36P and ΔCog4 cells respectively. Error bars, where visible, represent +/- 1 standard error of the mean.

In suspension culture, CRISPR generated Cog4L36P and ΔCog4 cells grow slower than WT cells (figure 2.12a). Average doubling times were calculated as 32.2, 36.8 and 38.3 hours for WT, Cog4L36P and ΔCog4 cells respectively. All doubling times for the cells were slower in suspension culture than in adherent culture, suggesting that the cells grow better in adherent culture than suspension. In suspension culture, both Cog4L36P and ΔCog4 cells have slower growth rates than WT, indicating that fully functioning Cog4 is needed to maintain normal cell growth. Viability of all cell lines was similarly high (>95%) when the growth curve was set up (figure 2.12b), and further viability measurements were taken every day when the cells were counted. For all cell lines, viability remained relatively stable throughout the growth curve time course, with a gradual declining trend as the cell numbers become higher. For WT cells, the viability dropped more rapidly by the end of the week, likely because the cells had grown beyond a sustainable concentration and nutrients in the medium may have been exhausted.

A
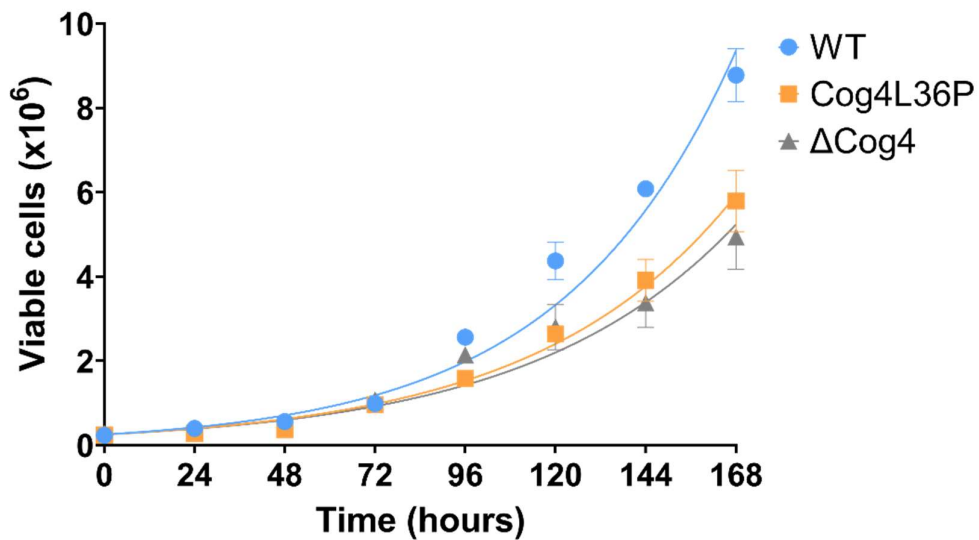
B
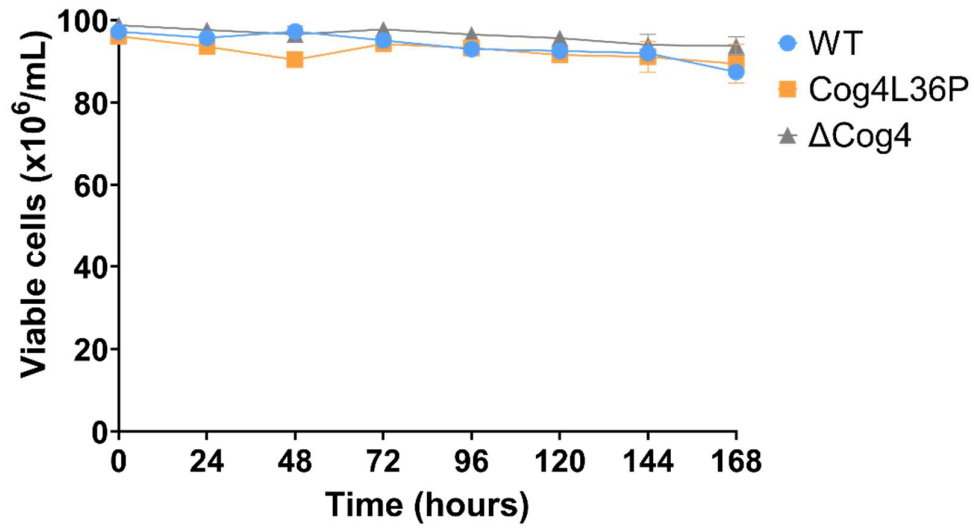
Figure 2.12| Rate of proliferation and viability of Cog4L36P and WT suspension cells. On day 0, cells were passaged to 0.25 x 10⁶ cells/mL in 2 x T25 flasks. Every 24 hours, a sample of cells was mixed in a 1:1 ratio with TrypLE and incubated for 5 minutes prior to counting using a ViCell cell counter. A) Cell counts are plotted for each cell line. Data were plotted and fit using GraphPad Prism software. Datasets were fit to an exponential, Malthusian growth curve of the form $Y = Y0*e^{kx}$, where Y0 is the starting population, which was constrained to be equal to 0.25×10^6 cells/mL, k is the rate constant and x is the time in hours. R^2 values for each curve were 0.9612, 0.9657 and 0.8932 for WT, Cog4L36P and ΔCog4 cells respectively. Error bars, where visible, represent +/- 1 SE. B) Viability is plotted for each cell line. For both graphs, N=2, error bars are +/- 1 SE.

2.5.3.2 Cog4L36P cells have higher metabolic activity, despite lower proliferation

As well as monitoring cell numbers directly, colorimetric assays are often used to assess the metabolic activity of cells. By using dyes that cells can metabolise into coloured products, the relative amounts of metabolically active cells can be compared between cell cultures. Emily Priest, the undergraduate student supervised by the author, assessed the metabolic activity as well as the growth rates of the knock-in cell lines, by using an MTT (3- [4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide) assay. The MTT assay measures the rate at which cells reduce a tetrazolium dye to coloured formazan crystals. The precise location of formazan formation is uncertain, though the reduction reaction is primarily mediated by NAD(P)H-dependent oxidoreductase and dehydrogenase enzymes along the electron transport chain in the mitochondrial

membrane during aerobic respiration (Ghasemi et al., 2021). NAD(P)H associated with the cytoplasm and non-mitochondrial membranes can also reduce MTT (Berridge et al., 2005). Monitoring formazan formation during the MTT assay therefore serves to measure general metabolic activity in cells. In contrast to the observation that ΔCog4L36P cells proliferate more slowly than ΔCog4WT cells, we observed a significantly enhanced response to the MTT assay in the ΔCog4L36P cells (figure 2.13). This indicates that the ΔCog4L36P cells reduce MTT faster than ΔCog4WT cells, suggesting that the ΔCog4L36P cells are more metabolically active than ΔCog4L36P cells.

However, it was vital to repeat this assay using the CRISPR generated Cog4L36P cells, as it is uncertain whether altered metabolic activity is due to the mutation or due to the over-expression of Cog4 in the knock-in cells. A similar assay was therefore performed by the author, using the CRISPR-engineered Cog4L46P cells alongside both WT and ΔCog4 control cell-lines. For this assay, alamarBlue® reagent was used instead of MTT, as this also provides a readout of general metabolic activity but isn't toxic to cells. The active ingredient of alamarBlue® is resazurin, which is cell-permeable and non-fluorescent. Like MTT, resazurin is metabolised under reducing conditions by the cell during aerobic respiration, producing a product (resorufin) that is red in colour and highly fluorescent. The assay was performed on cells grown in suspension culture, to match the growth conditions used for recombinant expression of biologics.
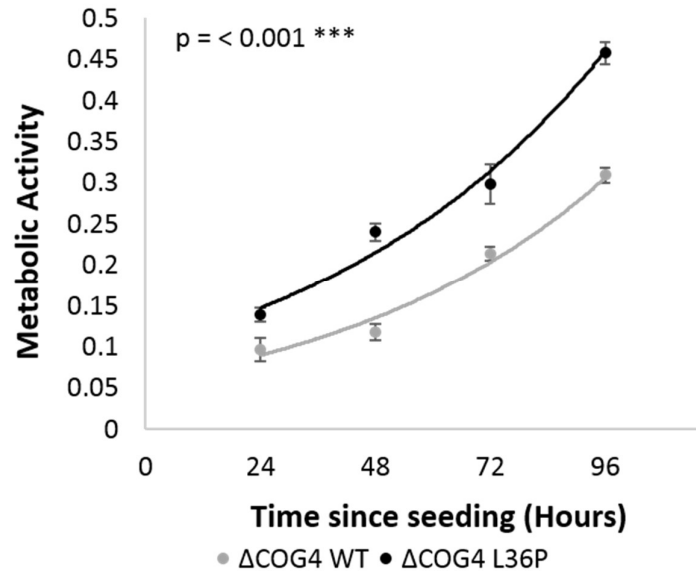
Figure 2.13 | ΔCog4L36P cells metabolise MTT faster than ΔCog4WT cells. Metabolic activity was monitored over 96 hours by project student, Emily Priest, using an MTT assay. Cells were seeded at 5,000 cells/well onto culture dishes, followed by treatment with 25 μl MTT after 24 hours, a 3-hour incubation, re-solubilisation of formazan in 100 μl acidic isopropanol, and determination of optical density at 570 nm. This was repeated at 48, 72, and 96 hours. There was a significant effect of genotype on metabolic rate (ANOVA; F = 80.4, d.f = 1,43, p < 0.001***). Results show the averages of each genotype (n=6).

In agreement with the data gathered by Emily Priest in the knock-in cells, results of the MTT assay did not correlate with cell number for the CRISPR engineered suspension cells (figure 2.14). Despite having a slower rate of proliferation than WT cells (figure 2.12a), Cog4L36P cells had a comparable response to the alamarBlue® assay to WT, suggesting that the Cog4L36P cells have higher metabolic activity per cell than WT. In contrast, ΔCog4 cells had slower proliferation and a lower response to the alamarBlue® assay. These data suggest that Cog4L36P cells are more metabolically active than WT or ΔCog4 cells in suspension culture.
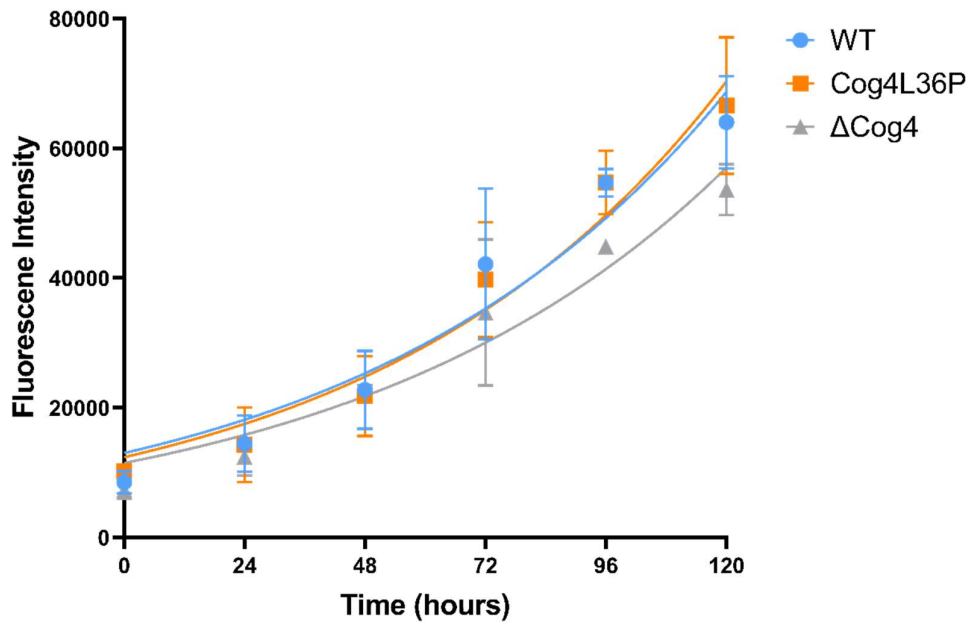
Figure 2.14 | Metabolic activity is altered in Cog4L36P and ΔCog4 cells. Suspension cells were cultured in shake flasks, starting at a density of 0.25 x 10^6 cells/mL. Every day, an aliquot of cells was removed from each flask, transferred to wells of a 12 well plate and incubated in medium containing 10% alamarBlue® reagent for 90 mins. After incubation, medium was harvested by centrifugation and transferred into wells of a black 96 well plate. Fluorescence intensity was measured using an excitation wavelength of 545 nm and emission wavelength of 600 nm using a CLARIOstar plate reader. Two biological replicates were performed, with one consisting of 9 technical replicates and the other, 4 technical replicates. Error bars represent +/- 1 standard error of the mean for the two biological replicates.

2.5.3.3 L36P mutation in Cog4 causes an adhesion defect

As well as observing that adherent Cog4L36P cells grew slower than WT, it was apparent that their morphology differed. Simply by observing the cells through a light microscope, it appeared that the Cog4L36P cells adhered more weakly to the tissue culture plate than WT cells. Their appearance was rounder and less elongated, suggesting that they have an adhesion defect.

To test how both cell surface adhesion receptor proteins and ECM components are affected by the Cog4L36P mutation, under the supervision of the author, an integrated Masters student, Maya Trivedy, analysed adhesion of ΔCog4L36P and ΔCog4WT cells

to both plastic and ECM components. This work was carried out prior to the development of the CRISPR engineered Cog4L36P cell line. Firstly, cells were trypsinised and seeded onto either uncoated tissue culture plates or plates coated in vitronectin (a purified ECM component) and incubated for 3 hours to allow enough time for ECM secretion and adhesion. All non-adhered cells were gently washed from the plates before adhered cells were detached using trypsin and counted using a hemocytometer. When seeded onto the plastic tissue culture plates, the ΔCog4L36P cells adhered significantly weaker than ΔCog4WT cells (ANOVA, $P < 0.0001$) (figure 2.15). The addition of vitronectin enabled more ΔCog4L36P cells to adhere, though still not to the same extent as WT. This suggests that both adhesion receptors and secreted ECM proteins are impacted by the L36P mutation.

In order to test the adhesive properties of ECM secreted by the cells, plates of each cell line were grown to confluency before removing the cells by treatment with EDTA and subsequent washing with PBS. It was hypothesised that treatment with EDTA, as opposed to trypsin, would detach cells whilst leaving secreted ECM components intact and adhered to the plates. Trypsin is a protease that digests both cell surface proteins and ECM components, whereas EDTA releases cells more gently by chelating calcium and magnesium, which are required for interactions between cell surface proteins and ECM components, whilst leaving all adhesion proteins intact. Released cells therefore retain their cell surface adhesion proteins but must secrete more ECM proteins upon seeding on a clean plate. The resulting plates, coated with either ΔCog4WT or ΔCog4L36P ECM, were used as vessels for subsequent adhesion assays with both cell lines. In both cases, ΔCog4L36P cells adhered more weakly than ΔCog4WT cells, confirming that adhesion receptors are affected by the Cog4L36P mutation (figure 2.15). Interestingly, plating cells onto ECM secreted by ΔCog4L36P cells weakened the adhesion of both cell types compared to their adhesion to plastic, suggesting that ECM proteins are also affected by the L36P mutation in Cog4 (figure 2.15).
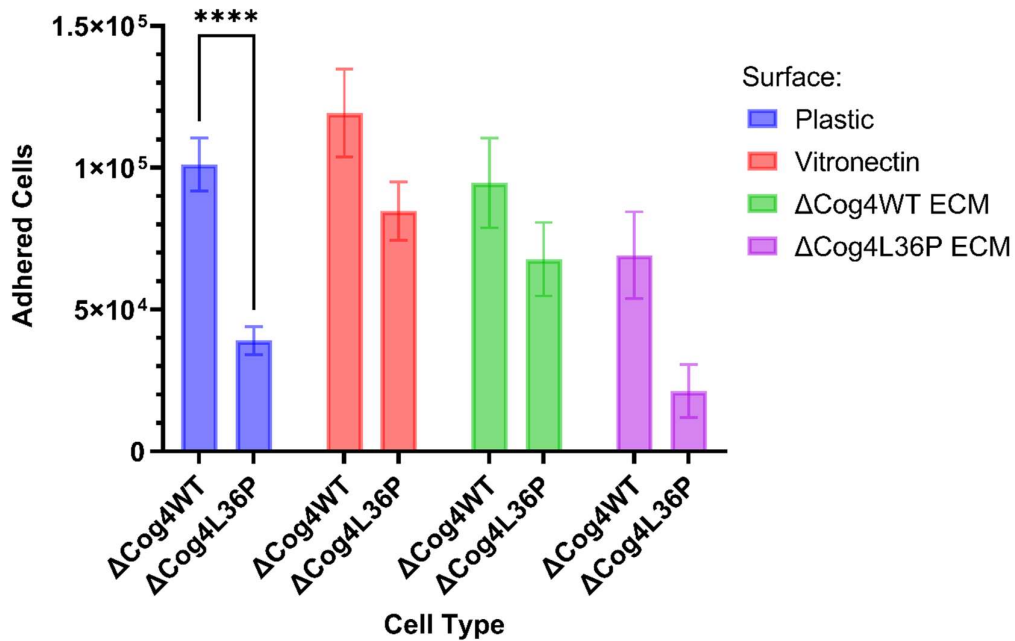
Figure 2.15 | Number of adhered cells per well. ΔCog4WT and ΔCog4L36P cells were seeded onto tissue culture plates, which had either been left untreated (plastic only) or coated with either vitronectin or ECM secreted by either ΔCog4WT or ΔCog4L36P cells. Cells were incubated on the plates for 3 hours, before washing off non-adhered cells with PBS. The remaining adhered cells were detached by trypsinisation, before counting them using a hemacytometer to estimate the number of cells in 0.5mL. The bars shown are averages (N=8 for plastic, 4 for vitronectin and 3 for ECM coated plates), with error bars representing +/-1 standard error of the mean. Statistical significance was calculated using Graphpad Prism software to perform a two-way ANOVA with Šidák's multiple comparisons test, comparing the mean numbers of ΔCog4WT and ΔCog4L36P cells on each surface, P < 0.0001****.

As with all other characterisation experiments, it was important to validate whether CRISPR-engineered cells behaved in the same way as did the knock-ins cell lines. Adhesion to plastic tissue culture plates was assessed using a similar 3-hour adhesion assay as carried out by Maya Trivedy, with the inclusion of ΔCog4 control cells. In light of the observation that ECM secreted by ΔCog4L36P cells had reduced adhesive properties compared to that secreted by ΔCog4WT cells, for the latest assay, cells were released with EDTA prior to seeding, as opposed to trypsin. Results of this assay suggested that adhesion of Cog4L36P cells to plastic was impaired compared to WT (figure 2.16). This could either be due to impaired secretion of ECM proteins in the

Cog4L36P cells or defective glycosylation of these proteins. The data also suggested that ΔCog4 cells adhered more weakly than did WT, though more ΔCog4 cells adhered than did Cog4L36P cells, suggesting L36P mutation in Cog4 causes a more severe adhesion defect than complete loss of Cog4. However, these differences were not statistically significant (ANOVA p > 0.1).
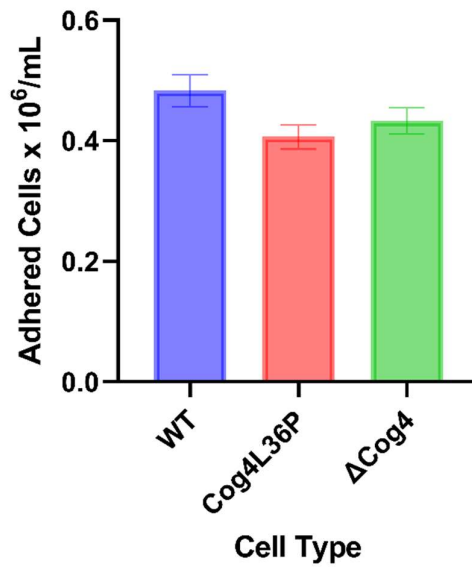
Figure 2.16 | Attachment of WT, ΔCog4 and Cog4L36P cells to plastic tissue culture plates following EDTA release. ΔCog4WT and ΔCog4L36P cells were seeded in triplicate onto untreated tissue culture plates. Cells were incubated on the plates for 3 hours, before washing off non-adhered cells with PBS. The remaining adhered cells were detached by trypsinisation, before counting them using a ViCell cell counter. Error bars represent +/- 1 standard error of the mean (n=3). Statistical significance was evaluated using one-way ANOVA with Tukey's multiple comparisons test (p > 0.1, not significant).

2.6 Conclusions and Future Plans

Prior to the start of this PhD, the Ungar lab identified and began characterisation of an L36P mutation in Cog4, which impairs its interaction with Rab30. In light of this impaired interaction and preliminary observations that a knock-in CHO-K1 cells expressing the mutation had an altered glycan profile, a Cog4L36P point mutant CHO-K1 suspension cell line was engineered using CRISPR-cas9. This chapter describes work to engineer and characterise the mutation, both *in vitro* to test the binding to Rab30 and *in vivo* to test its effects on various cell properties.

CRISPR/cas9 genome editing in CHO cells has previously been used to engineer glycosylation of mAbs by targeting glycosylation enzymes, or by altering the glycosylation sites on the mAbs themselves (Heffner et al., 2018). For example, inhibiting fucosylation by disrupting the FUT8 gene can enhance antibody dependent cell-mediated cytotoxicity in cancer cells (Wang et al., 2018, Louie et al., 2017). However, there are few examples in the literature of specific point mutations generated in CHO cells by CRISPR/cas9 for glycan engineering. Furthermore, estimations of editing efficiency using CRISPR/cas9 vary wildly, and there are no reports detailing the efficiency of using a combination of cas9 nickase and a HDR repair template to generate a point mutation in CHO cells. Therefore, it was of great interest to investigate this approach and report the efficiency for future benefit. A success rate of 1/11 was achieved for introducing a specific point mutation using the method presented here. It is hoped that this method can now be used as a template for generating future COG mutant cell-lines for glycan engineering. For example, our lab is also interested in generating cell lines expressing CDG-causing COG mutations, which are known to alter glycosylation (Ferreira et al., 2018, Ng et al., 2011).

Ultimately, the aim of this thesis is to test whether targeting the COG complex can improve the production of recombinantly expressed glycoprotein biologics. Characterisation of the Cog4L36P cell line is therefore beneficial not only for broadening our understanding of how the COG complex can indirectly play a role in multiple cellular processes, but also for testing important properties that could affect recombinant protein production.

When drawing conclusions from the characterization experiments described in this chapter, it is important to note that experiments were performed using only a single clone of each cell line. This was because a single clonal cell line of Cog4L36P was produced during CRISPR-Cas9n genome editing. It has long been known that CHO-K1 cells are naturally genetically plastic and prone to clonal variation (Konrad et al., 1977, Xu et al., 2011). Therefore, it is important to establish, in future, the extent to which clonal variation contributes to the differences observed between WT, Cog4L36P and ΔCog4 cells. This could be achieved by performing CRISPR-cas9n genome editing on a larger scale to increase the chances of isolating multiple Cog4L36P and ΔCog4 clones. Single cell clones of the WT parent line should also be generated, and experiments could be repeated using several clones of each cell line to establish how much variation in phenotypes there is between clones of the same genotype. This variation could be factored into analysis of characterization experiments to gain a better understanding of the true impact of Cog4 disruption on cell properties.

Despite being only a single base change, the effects of the L36P mutation in Cog4 are becoming strikingly apparent. Cog4L36P cells have higher metabolic activity, despite slower proliferation, and reduced adhesion compared to WT cells. Adhesion assays suggest that ΔCog4L36P cells secrete ECM that is less adhesive than that secreted by WT cells. Weaker binding of ΔCog4L36P cells to vitronectin compared to WT also suggests that cell surface adhesion molecules are affected by the mutation. To understand the nature of these defects, future work could include using PNGase F to release *N*-glycans directly from ECM laid down by each cell line and performing mass spectrometric glycan profiling to assess whether glycosylation of ECM proteins is altered in the mutant cells. A similar approach could be used to assess glycosylation of cell surface proteins, by releasing *N*-glycans from cells without lysing them.

In light of the discrepancy seen between cell proliferation and metabolic activity, it was also hypothesised that the Cog4L36P cells are metabolising faster than WT but may also be dying faster. In order to test this in future, apoptosis could be monitored in the cells. One way of monitoring apoptosis would be to culture the different cell lines in parallel and grow them beyond their viable concentration, whilst taking cell and medium samples every day. These samples could then be analysed by Western

blotting, using antibodies to test levels of early, mid- and late-phase apoptosis markers, such as Bcl-2, caspase-9 and PARP-1 respectively (D'Arcy, 2019). This would provide insight into how quickly each cell line enters and progresses though apoptosis.

In conclusion, a Cog4L36P CHO-K1 clonal cell line has been successfully engineered using CRISPR. Characterisation so far suggests these cells have properties that could be favourable for recombinant protein expression. Higher metabolic activity, specifically oxidative phosphorylation, has been linked to increased productivity of mAb-expressing cells (Link et al., 2004, Templeton et al., 2013). Weaker adhesion is also favourable, as biologics are produced in suspension culture, where minimising contact between cells or between the cells and the culture vessel means more cell surface is exposed for protein secretion.

Chapter 3: Whole Cell N-glycan Profiling of CHO-K1 Cell Lines with Altered Cog4 Expression

3.1 Introduction

3.1.1 Rationale for measuring whole cell *N*-glycans of mAb host cells

Monoclonal antibody (mAb) production is a $multi-billion global industry, which was valued at US$115.2 billion in 2018 and is expected to grow to $300 billion by 2025 (Lu et al., 2020). Currently, around 60% of mAbs are recombinantly expressed by Chinese hamster ovary (CHO) cells (Grilo and Mantalaris, 2019). As well as their excellent productivity and their capacity for correct protein folding, they are an expression system of choice due to their ability to perform human-like glycosylation (Brooks, 2004). However, costs are high and glycan heterogeneity poses a significant problem, since glycosylation can alter properties of therapeutic proteins (Sinclair and Elliott, 2005).

Glycan profiling of mAbs is routinely carried out to assess product quality. However, getting to that point requires selecting mAb-expressing clonal cell lines, scale up of those cell lines, before performing a production run in bioreactors, mAb protein purification, and mass spectrometric profiling of either the whole mAb or of the PNGase F-released *N*-glycans. A lot of work is therefore required before problems with mAb glycosylation can be detected, making discovery of new glycan engineered hosts expensive and time-consuming. Alternatively, whole-cell glycan profiling of potential hosts, before transfection with mAb plasmids, offers an attractive way of screening for desirable glycosylation characteristics in the early stages of a discovery experiment.

Glycomics techniques have improved considerably over the years, and the number of cells required for *N*-glycan analysis alone has been cut from ~100 million (North et al., 2010) to 0.5 million cells (Skeene et al., 2017, Abdul Rahman et al., 2014). A technique

developed in our lab has also enabled the use of a single filter to isolate *N*- and *O*-glycans, and the free protein (Skeene et al., 2017). Applying this technique to potential host-cell lines could therefore considerably cut costs of identifying cells with desirable glycosylation.

3.1.2 *N*-glycan profiling of cell lines with engineered Golgi organization

The approach presented in this thesis to engineer mAb glycosylation involves targeting the COG complex, an important regulator of Golgi vesicle tethering and glycosylation enzyme sorting, to generate glycan engineered mAbs. Previous evidence suggests that altering COG causes changes to glycosylation, both in the context of cultured cell lines (Bailey Blackburn et al., 2016, Fisher et al., 2019, Krieger et al., 1981, Kingsley et al., 1986) and in congenital disorders of glycosylation (Foulquier et al., 2006, Wu et al., 2004, Jaeken, 2011, Leroy, 2006), leading to the hypothesis that COG could be targeted for glycan engineering. Specifically, Cog4 is the focus, working with the CRISPR-engineered Cog4L36P mutant CHO-K1 cell line, which was introduced in the previous chapter of this thesis. Cog4's involvement in vesicle tethering and glycosylation enzyme sorting includes directly interacting with multiple Rabs, SNAREs and trafficking proteins (Willett et al., 2013b). These include Rabs 1a, 4a and 30 (Miller et al., 2013), multiple SNARE proteins, including STX5 and STX16 (Laufman et al., 2013b) and several other proteins, including the COPI coat protein (Suvorova et al., 2002). The importance of Cog4 for correct glycosylation is also highlighted in patients with Cog4-CDGs. For example, an R729W mutation in Cog4 disrupts Golgi structure and causes a reduction in galactosylation and sialylation (Reynders et al., 2009). Similar defects have also been seen in a patient carrying an L773R mutation in Cog4 (Ng et al., 2011).

In the previous chapter, it was shown that Cog4L36P mutant has a weakened interaction with Rab30. It was hypothesised that this would alter vesicle tethering, causing knock-on effects on glycosylation enzyme sorting in the Golgi. Alteration to the enzyme distribution would lead to changes in the cellular glycan profile and may also change the glycosylation of recombinantly expressed glycoproteins. Indeed, prior to the development of the CRISPR engineered Cog4L36P cell line, previous work in our lab with a Cog4L36P knock-in cell line suggested the mutation resulted in a glycan profile distinct from that of both WT CHO-K1 cells and ΔCog4 cells. However, it could not be

established whether this effect was due to the mutation itself or a result of over-expression of the Cog4 gene in the knock-in cell line. By using the CRISPR-cas9n engineered Cog4L36P cell line, it is more certain that observed changes to the glycan profile are due to the mutation, as opposed to expression levels, as Cog4 is expressed under its endogenous promoter and expression levels are comparable to those in the WT. By comparing the *N*-glycan profiles of Cog4L36P cells to those of both WT and ΔCog4 cells, a broader understanding of the role of Cog4 in glycan biosynthesis can be gained. Including ΔCog4 cells as a control has enabled understanding of the extent to which glycosylation is changed by specifically disrupting the Cog4-Rab30 interaction in the Cog4L36P cells compared to completely removing Cog4.

3.2 Aims

The aim of the work presented in this chapter was to analyse glycosylation in WT, Cog4L36P and ΔCog4 cells by permethylation and matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS). This approach has been used extensively in our lab (Abdul Rahman et al., 2014, Fisher et al., 2019, Wilson et al., 2016) and has been shown to allow relative quantification of glycans in a mixture (Wada et al., 2007). In light of the observed differences in the adherent properties of the cells, reported in the previous chapter, this chapter also includes a comparison of the *N*-glycan profiles of each cell line grown in both adherent and suspension culture. In this chapter, I have also begun to explore the potential for whole-cell glycan profiling in an industrial lab at GlaxoSmithKline, using a commercial kit by Waters for *N*-glycan release and labelling with a "RapiFluor" tag, coupled with HPLC-MS analysis. In combination with cell line characterization, mAb production and mAb glycan analysis, explored for Cog4L36P cells and described in other chapters, I can begin to understand how specific changes to glycosylation are linked to phenotypic changes and broaden understanding of the glycan biosynthesis pathway and how it is controlled.

3.3 Methods

N-glycans were prepared and analysed using FANGS, permethylation and MALDI-MS (figure 3.1a), or by using the Waters GlycoWorks RapiFluor *N*-glycan kit with HPLC-MS (figure 3.1b).
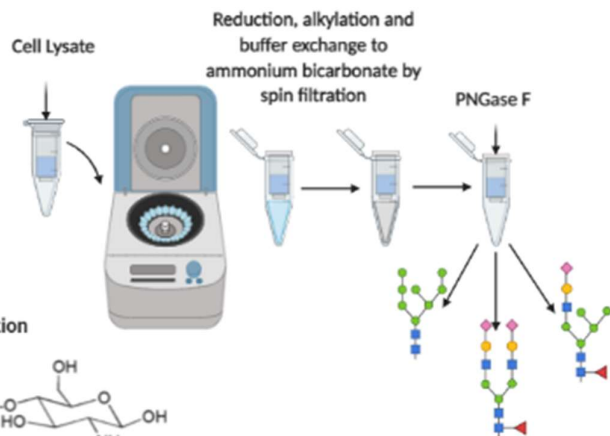
A
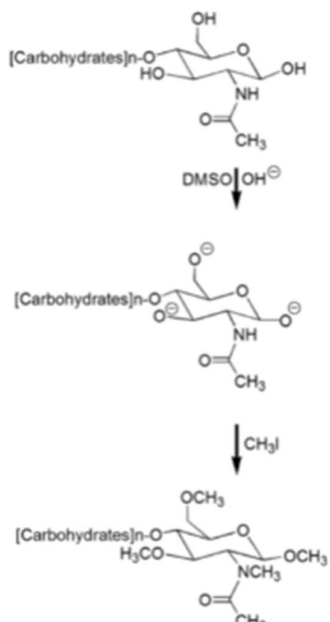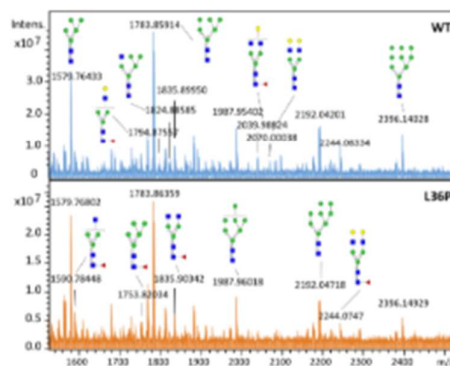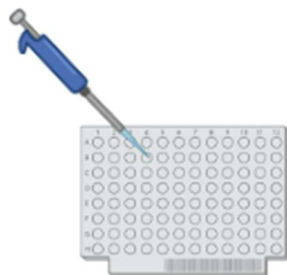
B
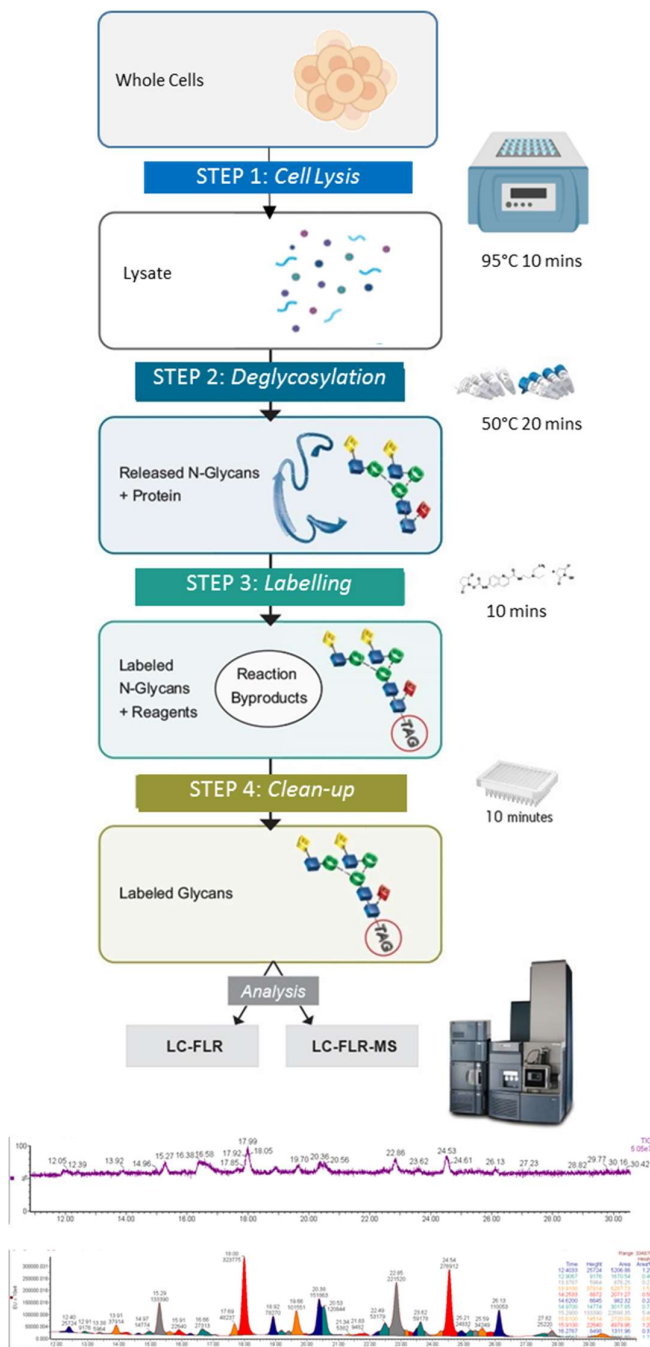
Figure 3.1 | Schematic representations of glycan profiling procedures, used in York and broadly in academia (permethylation/MALDI) vs. the Waters GlycoWorks RapiFluor N-glycan kit method (GlycoWorks/HPLC). A) 1. Release of N-glycans from whole cell lysates by FANGS, 2. Permethylation, 3. Spotting on a MALDI target plate with DHB matrix, 4. MALDI-MS for glycan profiling. B) Adapted Waters GlycoWorks RapiFluor N-glycan kit method, using whole cells instead of purified glycoprotein. This figure is adapted from a figure produced by Waters (Waters, 2017).

3.3.1 CHO-K1 Cell Lysis for *N*-glycan profiling

CHO-K1 cells grown in suspension were counted using a Vi-Cell counter prior to harvesting 9 million cells for each replicate. Cells were harvested by centrifugation at 1000 *x g* for 5 minutes. The pellets were washed five times with 5 mL PBS in 15 mL falcon tubes, centrifuging at 1000 *x g* for 5 minutes each time. After the fifth wash, the pellets were re-suspended in 1 mL PBS and transferred to a 1.5 mL Eppendorf® protein LoBind microfuge tube for centrifugation at 16000 *x g* for 5 mins.

Adherent CHO-K1 cells were grown to confluency in 10 cm plates and one confluent plate was used for each replicate. Medium was aspirated from each plate before gently washing cells five times with 5 mL PBS. After the final wash, plates were placed on ice and 1 mL PBS added to each. Cells were detached with a cell scraper and transferred to fresh 1.5 mL LoBind tubes for centrifugation at 16,000 *x g* for 5 minutes.

The volume of the pellets was then estimated by comparison to a known volume of liquid and pellets were re-suspended in an equivalent of 10 x the pellet volume of FANGS lysis buffer (FLB) (100 mM Tris-HCl, 4% w/v SDS, 100 mM DTT, pH 7.5). Lysates were heated at 97 °C for 5 mins and checked for viscosity by attempted aspiration into a 200 µL pipette tip. Viscous lysates were sheared by sonication at 70 % amplitude by delivering 3 x 1 second pulses with 10 second rests. Lysates were then centrifuged at 16,000 *x g* for 5 minutes and the supernatants transferred to clean Eppendorf® protein LoBind tubes and stored at -80 °C prior to *N*-glycan release.

3.3.2 Filter-aided *N*-glycan Separation

N-glycans were released following the filter-aided *N*-glycan separation (FANGS) protocol (Abdul Rahman et al., 2014). All centrifugation steps, including those for washes and buffer exchanges, were carried out at 14,000 x *g* for 7 minutes at room temperature. Cell lysates were re-suspended in urea Tris (UT) solution (8 M urea in 100 mM Tris/HCl pH 8.5, prepared fresh) using a ratio of 10:1 UT:lysate volume. This whole solution was passed through a centrifugal ultrafiltration device (Amicon Ultra-0.5, Ultracel-30 membrane, nominal mass cut-off 30 kDa, Millipore) in aliquots of approx. 500 µL at a time. The filter units were then washed twice with 250 µL UT. Proteins were alkylated with iodoacetamide treatment by the addition of 300 µL 40 mM iodoacetamide in UT and incubation in the dark at ambient temperature for 15 mins.

Filter units were centrifuged at 14,000 x *g* for 7 minutes to remove the iodoacetamide solution and washed with 250 µL UT. The buffer above the membrane was then exchanged for 50 mM ammonium bicarbonate solution (pH 7.5 to 8) by washing the filter 4 times with 250 µL volumes. Filter units were transferred to clean tubes and *N*-glycans were released by overnight treatment with 4 µL PNGase F (8 units) in 96 µL ammonium bicarbonate at 37°C. Filters were sealed with Parafilm™ to prevent evaporative loss.

The following day, *N*-glycans were retrieved following centrifugation at 14,000 *x g* for 7 minutes, followed by the addition of 2 x 300 µL HPLC water, centrifuging at 14,000 *x g* for 7 minutes after each addition for a total of 700 µL collected volume. FANGS-released *N*-glycans were transferred from the ultrafiltration device collection tubes to Eppendorf® protein LoBind tubes and either stored at -80 °C before drying or immediately transferred to clean glass tubes and dried using a SpeedVac vacuum centrifuge at 30 °C.

3.3.3 Permethylation

N-glycans were permethylated largely following the protocol described in (Abdul Rahman et al., 2014). Dry glycan samples were dissolved in excess (~600 µL) dimethyl sulfoxide (DMSO). Sodium hydroxide pellets were freshly ground in a glass mortar using a glass pestle and a small, heaped microspatula full of the powder (~25 mg) was added, followed by light agitation of the tube contents. Repeated additions of iodomethane were made by adding 10 drops (~200 µL), waiting 10 minutes before adding another 10 drops, waiting a further 10 minutes and finally adding 20 drops (~400 µL) and leaving for 20 minutes. A 100 mg/mL solution of sodium thiosulfate in HPLC-grade water was prepared fresh and to quench the reaction approximately 1.5 mL added to the tube, which was then vortexed. 1.5 mL of dichloromethane was added immediately afterwards, and the tube was vortexed once again. After allowing aqueous and organic layer to partition, the upper aqueous layer was removed. The organic layer was washed 3 times, each with the addition of 1.5 mL HPLC-grade water. The organic layer was dried under vacuum in a vacuum centrifuge. Dried, permethylated glycans were stored at -20 °C prior to mass spectrometric analysis.

3.3.4 MALDI target plate spotting

Permethylated *N*-glycans were dissolved in 50% acetonitrile (ACN) with 0.1% trifluoroacetic acid (TFA). For most samples, 20 µL of solvent was added. If weak signals were observed in the mass spectrum though, the sample was dried and dissolved in a minimum of 4 µL solution. If the spot appeared too concentrated and looked glassy in texture, the sample was diluted further with solvent up to a maximum volume of 160 µL prior to spotting. 1 µL of the permethylated *N*-glycan solution was spotted on a MALDI target plate and dried under vacuum. 1 µL of a solution of 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 50% ACN, 0.1% TFA was spotted on top of each glycan spot and dried under a warm lamp. Spots were recrystallized with the addition of 0.3 µL ACN, which air dried rapidly.

3.3.5 MALDI-MS

Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III time-of-flight instrument operated in reflectron mode, equipped with an Nd:YAG smart beam laser. Mass spectra were acquired over the *m/z* range 800-4500 at 100 Hz laser frequency. A total of 2400 shots were summed for each spectrum, with laser power manually adjusted over the range 40-80% to obtain the best signal to noise ratio (S:N) for glycan signals. Mass spectra were externally calibrated against an adjacent red phosphorus calibrant spot.

3.3.6 MALDI Mass Spectral Processing

Monoisotopic masses were obtained using the SNAP averaging algorithm, optimised for permethylated glycan units (C 10.5, O 5.3, H 18.5, N 0.5), with a minimum S:N of three. Bruker FlexAnalysis software (version 4.2) was used to perform spectral processing and peak list generation. A table containing the generated mass list, the full width half maximum (FWHM) of each signal and the area of each isotopic peak was exported as a .csv file. A Python algorithm, written by Peter Fisher, was then used to compare the peaks to a list of known CHO-K1 *N*-glycan masses, with a 1000 ppm error range. The algorithm also calculated the total peak intensity for each glycan from the FWHM and ion signal area in the .csv file. Using the algorithm for spectral processing gives a similar result to the manual quantification method used in Abdul Rahman et al. 2014. When comparing spectra from different cell lines, the total ion signal intensity

for each glycan was normalized to the total ion signal for only glycans that were common to all the cell lines used in the comparison. For normalization, common glycans were selected from the average of all the replicates for each cell line, as opposed to those that were common to all replicates. An arbitrary minimum of 15 *N*-glycan peaks was chosen for selecting quality spectra to include in the analysis. Glycan profiles were presented as bar charts, generated using Microsoft Excel, displaying average (N>=4) relative abundance for each glycan, with error bars representing +/- 1 standard error of the mean. To simplify the presented data, only glycans with a relative abundance of greater than 1% were included in the graphs.

3.3.7 GlycoWorks RapiFluor *N*-Glycan Sample Preparation

Whole-cell *N*-glycans were prepared using the Waters GlycoWorks RapiFluor *N*-glycan kit using a method based on the manufacturer's protocol. A schematic representation of the method is summarised in figure 3.1b. GlycoWorks reagents were prepared immediately prior to sample preparation, following the manufacturer's protocol. Rapigest solution was prepared by dissolving 20 mg Rapigest powder in 400 µL Rapigest buffer and 270 µL HPLC grade water by pipetting directly into the Rapigest vial and vortexing. PNGase F solution from the kit (80 µL) was diluted with 293 µL Rapigest buffer and vortexed to mix, as per the manufacturer's instructions. One vial of 55 mg RapiFluorMS label was dissolved in 670 µL dimethylformamide (DMF) and mixed by pipetting.

Cells were counted by mixing 500 µL culture with 500 µL TrypLE™ and incubating at 37°C for 5 mins prior to counting using a Vi-CELL cell counter. Two million cells were subsequently harvested by centrifugation at 1000 *x g* for 5 minutes. Cell pellets were washed three times with 1 mL PBS by re-suspension, centrifugation, removing the supernatant and repeating. Following removal of the PBS after the final wash, cell pellets were lysed by re-suspending in 100 µL Rapigest solution, heating at 95°C in a heating block for 10 minutes and sonicating for 10 minutes in a sonic bath running at approx. 40 kHz. *N*-glycans were released by adding 40 µL PNGase F solution and heating at 50°C for 20 minutes in a heating block. Samples were then allowed to cool to room temperature prior to labelling with 20 µL RapiFluor solution for 10 minutes at room temperature. The labelling step was quenched by adding 360 µL acetonitrile. Samples

were centrifuged at 1000 *x g* for 5 minutes to pellet cell debris prior to clean-up using a Waters GlycoWorks HILIC µElution plate.

Prior to sample clean up, µElution plate wells were washed by adding 200 µL HPLC grade water and passing it through the stationary phase into a waste container using a vacuum manifold. Wells were washed in the same way with 85% acetonitrile (ACN) before loading the samples under vacuum. Samples were then washed twice with 600 µL wash buffer consisting of 90% ACN, 1% formic acid. The waste tray was replaced with a collection plate and glycans were eluted with 3 x 30 µL GlycoWorks SPE elution buffer. Glycans were diluted with 310 µL GlycoWorks SPE diluent solution (a proprietary mixture of DMF and ACN) prior to analysis by HPLC-MS.

3.3.8 HPLC-MS Analysis of Labelled *N*-Glycans

The released, RapiFluor-labelled *N*-glycans were analysed by Kirsty Skeene at GlaxoSmithKline using HILIC UPLC with fluorescence (FLR) and mass detection using a quadrupole mass analyser. The analysis was performed on a Waters system equipped with ultra-high performance liquid chromatography (UHPLC, Waters, I-Class, FLR detection) and QDa mass detector (Waters ACQUITY QDa). Rapifluor-labelled *N*-glycans were loaded onto a HILIC column (ACQUITY UPLC Glycan BEH Amide Column, 130 Å, 2.1 mm × 150 mm, 1.7 µm, Waters), running at 0.4 mL/min at 45 °C column temperature. The mobile phases A and B were acetonitrile and 50 mM ammonium formate aqueous solution (pH 4.4), respectively. Analyte separation was accomplished by gradient elution. A 55-minute gradient program was run as follows: a linear gradient of mobile phase B from 25% to 46% was applied over 35 min and from 46% to 100% over the next 1.5 min. The column was maintained in 100% B for the next 3 min before returning to the starting gradient (25% B) for the remainder of the program to equilibrate the column. The mass spectrometer was run in positive mode, with capillary voltage set to 1.5 kV and cone voltage at 20 V. Spectra were recorded in the *m/z* range 500 – 1250. Excitation and emission of the FLR detector were set to 265 nm and 425 nm respectively.

3.3.9 HPLC-Mass Spectral Processing

LC-MS and fluorescence data were processed using Waters MassLynx™ software. The software automatically integrated peaks from the fluorescence chromatograms, which

were adjusted manually for peaks which were not fully resolved. Retention times from the total ion chromatograms (TIC) were aligned with fluorescence chromatograms. Mass spectra were combined across each retention time period in the TIC that corresponded to a peak in the fluorescence chromatogram, representative examples of which are shown in figures 3.1b and 3.9. Glycans were assigned following inspection of m/z values in the corresponding combined spectrum and comparison to a database of m/z values of RapiFluor tagged $[M+2H]^{2+}$ and $[M+3H]^{3+}$ signals for every known mammalian N-glycan. The database was generated in Microsoft Excel and m/z values were calculated by summing monosaccharide in-chain residue masses, the mass increment introduced by the bound RapiFluor tag (329 Da) and additional proton masses (2 or 3 Da) before dividing by 2 for $[M+2H]^{2+}$ ions or 3 for $[M+3H]^{3+}$ ions. Charge state and structural assignments were based on determining what glycan structure in the database, if any, would be represented by the measured m/z value resulting from that charge, and what structure gave the closest match between its calculated and the measured m/z values. It is not possible to determine charge state based on QDa mass spectra due to the inherently low mass resolution of this detector. If spectra indicated that a peak contained more than one glycan and the minor glycan signal was greater than or equal to half the intensity of the major signal, glycans were considered to be co-eluting and they were quantified as such.

For quantification, fluorescence peak areas were summed, and this sum was used to calculate a relative abundance percentage value for each glycan or pair of co-eluting glycans on the basis of the individual peak area as a percentage of the summed total peak area. When comparing glycan profiles generated using the two different methods presented in this chapter, the relative abundance for each glycan was normalized to the total ion intensity (permethylation/MALDI method) or total fluorescence signal (Glycoworks/HPLC method) for only glycans that were common to both profiles used in the comparison. For normalization, common glycans were selected from the average of all the replicates for each method, as opposed to those that were common to all replicates. Glycan profiles were presented as bar charts, generated using Microsoft Excel, displaying average (N>=3) relative abundance for each glycan, with error bars representing +/- 1 standard error of the mean.

3.4 Results

3.4.1 Evaluating the Effects of COG4 L36P Mutation or Knockout on Total Cell *N*-glycan Content in Adherent Cell Lines

To understand how much the Cog4L36P mutation impacts glycosylation, mass spectra of whole cell *N*-glycans extracted from WT, Cog4L36P and ΔCog4 adherent CHO-K1 cells were acquired (figure 3.2). It was hypothesised that the point mutation, and the resulting impairment to the interaction between Cog4 and Rab30, would lead to disruption of enzyme sorting and altered glycan processing and that this could alter the profile of *N*-glycans produced by the mutant cells. Comparing the glycan profiles of Cog4L36P cells with those of WT and ΔCog4 cells would allow me to understand the specific glycosylation changes caused by the mutation, and whether these are different to those caused by a complete lack of Cog4.

In all mass spectra, signals at between *m/z* 1200 and around 4300 were observed, which matched the expected *m/z* values for singly charged permethylated *N*-glycans, with no signs of contamination (figure 3.2). For each signal, glycan structures have been proposed based on knowledge of the *N*-glycan biosynthetic pathway and previously published CHO-K1 glycan profiles (Abdul Rahman et al., 2014, North et al., 2010). The signals with the highest intensities correspond to oligomannose type *N*-glycans. These signals start at ~*m/z* 1579 for $GlcNAc_2Man_5$ and repeat every 204 *m/z* units for each additional permethylated hexose residue, up to $GlcNAc_2Man_9Glc_1$ at ~*m/z* 2600. The most intense peaks correspond to the nominal monoisotopic masses for M+H+ signals of $GlcNAc_2Man_5$ and $GlcNAc_2Man_6$ (~*m/z* 1783), followed by $GlcNAc_2Man_7$ (~*m/z* 1987), $GlcNAc_2Man_8$ (~*m/z* 2192) and $GlcNAc_2Man_9$ (~*m/z* 2396), which are each around half the intensity of the Man_5 and Man_6 species. Multiple complex glycan peaks are also present, though these are less intense than those for the oligomannose glycans. The most prominent complex glycan peaks are at ~*m/z* 1835, ~*m/z* 2605 and ~*m/z* 2966, corresponding to $FucGlcNAc_2Man_3GlcNAc_2$, $FucGlcNAc_2Man_3GlcNAc_2Gal_2NeuAc_1$ and $FucGlcNAc_2Man_3GlcNAc_2Gal_2NeuAc_2$, respectively. Multiple other complex bi- and tri-antennary complex peaks are also present, though these are of much lower intensity.

146
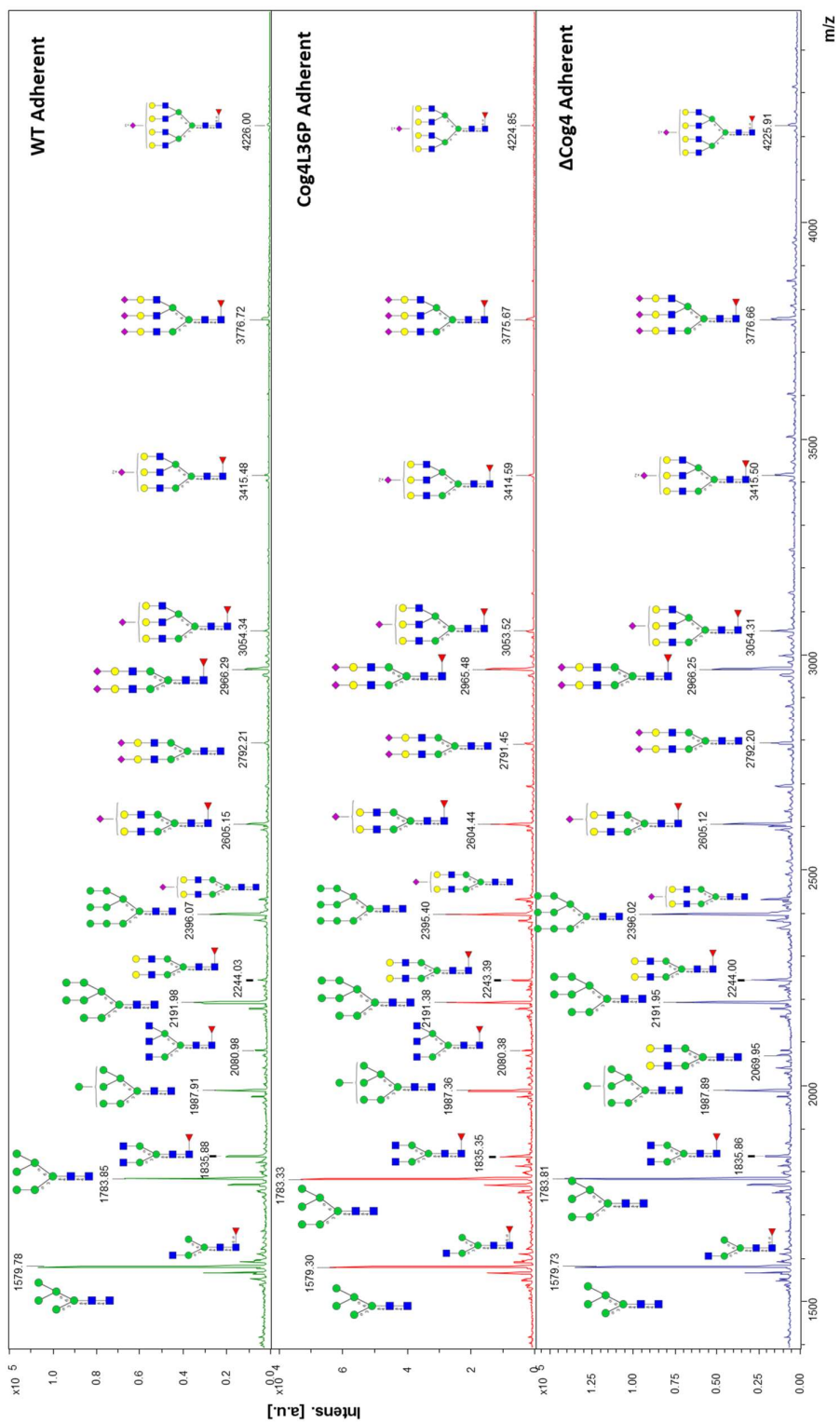
Figure 3.2 | Representative raw MALDI mass spectra for adherent WT, L36P and ΔCog4 N-glycans. Permethylated N-glycan peaks with signal:noise >3 are labelled with their calibrated m/z value and proposed glycan structure.

Similar glycan profiles, consisting of the same subset of *N*-glycans, are seen for each of the three cell lines. However, comparing the raw spectra by eye is not sufficient for quantitative comparison between the cell lines as each spectrum represents only a single cell culture plate. Multiple replicates are required to ensure natural variation between cell cultures is not mistaken for biological differences between cell lines. Therefore, multiple plates of each cell line were grown, *N*-glycans were harvested from each, spectra were recorded and used to calculate average relative abundances for all the glycans present in each (figure 3.3).

As with previous observations in our lab with Cog4L36P knock-in cells, the CRISPR-engineered Cog4L36P cells appear to have glycosylation distinct from that of both WT and ΔCog4 cells (figure 3.3). Starting with the oligomannose glycans, in Cog4L36P cells there are much higher levels of $GlcNAc_2Man_6$ but reduced levels of $GlcNAc_2Man_5$ and $GlcNAc_2Man_8$ compared to WT (figure 3.3). Differences in the levels of oligomannose glycans suggest altered levels, distribution, or function of mannosidase I (ManI), which acts in the *cis*-Golgi. This enzyme progressively trims alpha-1,2-linked mannoses from $GlcNAc_2Man_9$ to produce $GlcNAc_2Man_5$. In the Cog4L36P cells, it would appear that ManI is acting faster than in WT cells initially but reaches a bottleneck by the time it has trimmed the glycan down to $GlcNAc_2Man_6$. There are also altered levels of oligomannose glycans in the ΔCog4 cells: $GlcNAc_2Man_9$ and $GlcNAc_2Man_6$ are elevated compared to WT, and all other oligomannose glycans are less abundant. This suggests ManI is acting differently in these cells and also causing a bottle neck at the Man_6 species. Additionally, elevated $GlcNAc_2Man_9$ could indicate that mannosidase I processing in the ER is affected by knocking out Cog4 in CHO-K1 cells because $GlcNAc_2Man_9$ is formed in the ER and trimmed down to $GlcNAc_2Man_8$ before entering the *cis*-Golgi.

Cog4 is clearly important for the later steps of glycan processing too, as levels of all glycans terminating in GlcNAc are lower in both Cog4L36P and ΔCog4 cells than in WT cells (figure 3.3, figure 3.4a). This could be either due to less GlcNAc addition in the medial Golgi or more galactosylation in the trans-Golgi. N-acetylglucosaminyltransferase (Mgat) enzymes act in the medial Golgi and they are responsible for catalysing the addition of GlcNAc residues to growing glycan chains.

Initially, MgatI and mannosidase II (ManII) convert the glycan structure from oligomannose type to complex type. Other Mgat enzymes catalyse the addition of further GlcNAc residues, determining the number of branches the glycan has. Galactosyl- and sialyltransferase enzymes later catalyse the addition of galactose and sialic acid residues respectively. Based on these glycan profiles alone (figures 3.2 and 3.3) and without further modelling, we cannot say for sure which enzymes are impacted by impairment to Cog4; however, the sorting and distribution of all or multiple enzymes could be affected.

When the relative intensities for each glycan type in each cell sample are summed, broad effects on glycosylation can be seen more clearly (figure 3.4). All cell lines have a similar percentage of oligomannose glycans (figure 3.4), meaning the percentage of total complex glycans is also similar. There is also no significant difference in fucosylation between the cell lines (figure 3.4), suggesting that fut8, the enzyme catalysing core fucosylation, is unaffected by changes to Cog4. However, the types of monosaccharides making up the complex glycans differs between cell types (figure 3.4). Of these, the proportion of glycans terminating in GlcNAc is lower in Cog4L36P cells than WT, the proportion of galactose-terminating glycans is no different from that in WT cells and the proportion of sialylated glycans is slightly higher in Cog4L36P cells (figure 3.4). This suggests that glycan processing in the *trans*-Golgi, which is where galactose and sialic acid residues are added to *N*-glycans, is unaffected by the Cog4L36P mutation.

Of all the cell lines tested, ΔCog4 cells have the highest proportion of *N*-glycans containing sialic acid, but the lowest proportion of glycans terminating in GlcNAc (figure 3.4). It was expected that glycan processing would be negatively affected by knocking out Cog4, as previous observations in ΔCog4 HEK293T cells showed impairment in *cis*/medial Golgi glycosylation and thus reduced sialylation and fucosylation (Bailey Blackburn et al., 2016). It is therefore surprising to see an overall increase in the relative amount of all sialylated glycans, which have undergone more glycan processing steps, but no significant change in fucosylation.
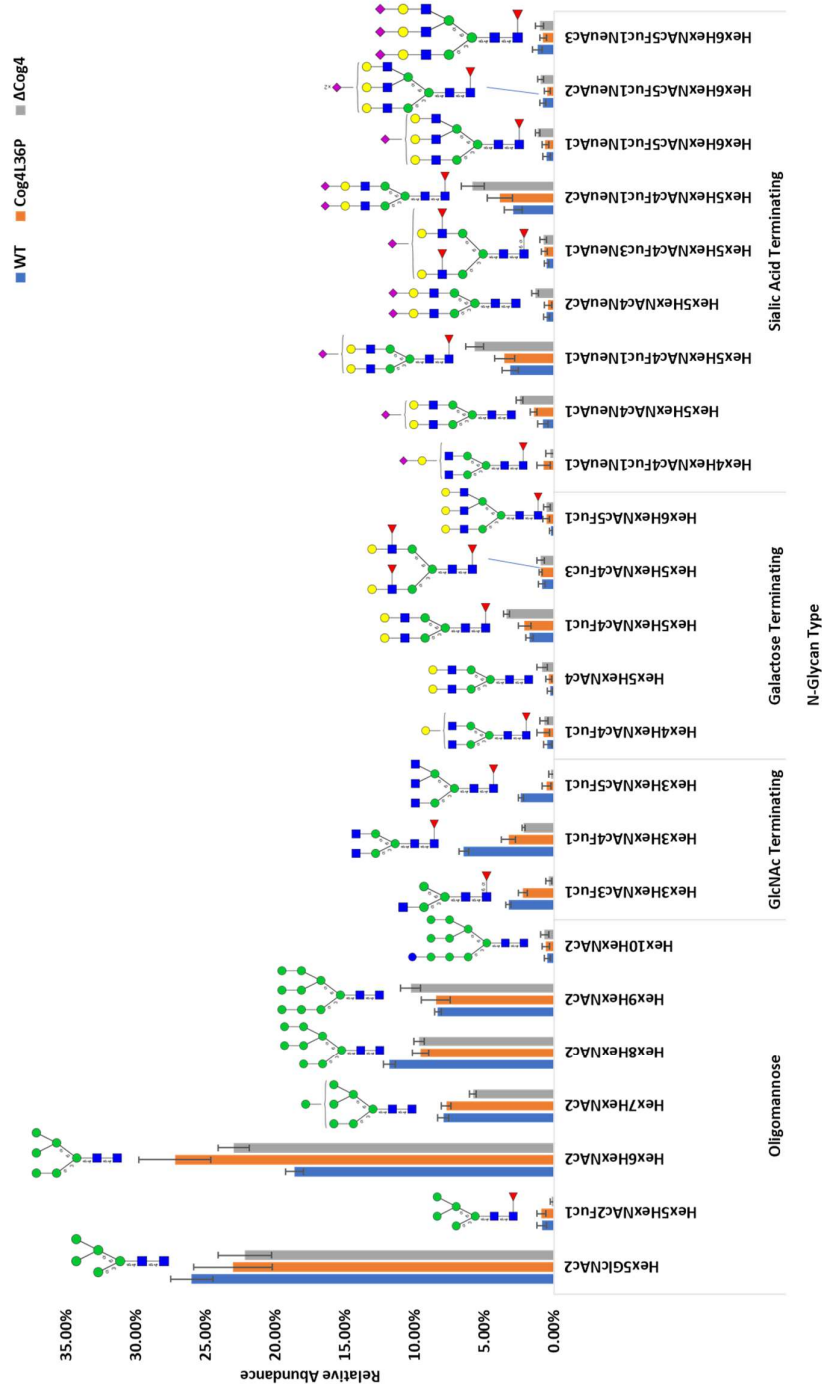
Figure 3.3 | Whole-cell N-glycan profiles of WT, L36P and ΔCog4 adherent cells. Bar chart showing the relative intensities of the MALDI-MS signals for all N-glycan species detected in each cell line. Spectra were only included for analysis if 10 or more glycans were detected, thus excluding poor quality data. The glycan intensities in each individual spectrum were normalized to the sum of intensities in each spectrum for the subset of glycans that are common to all three cell lines. The X-axis shows compositional names, with the predicted structures shown above each bar; structures were proposed based on knowledge of the biosynthetic pathway. Error bars represent +/- 1 standard error of the mean (WT n = 5, Cog4L36P n = 5, ΔCog4 n = 5).
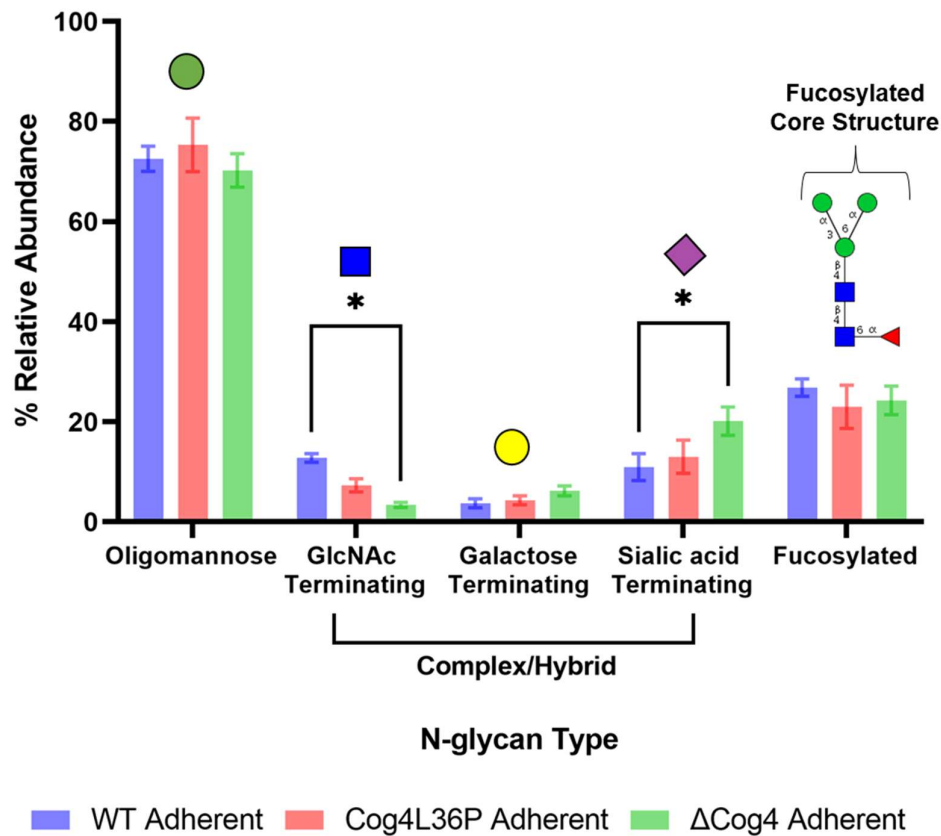
*Figure 3.4 | Relative abundances of each glycan sub-class in adherent WT, Cog4L36P and ΔCog4 cells. Bar chart showing average relative abundances of each subclass of glycans. Oligomannose glycans contain only mannose residues attached to the core (⬤). Of the complex/hybrid glycans, GlcNAc terminating refers to glycans with only GlcNAc residues attached to the core (◼), galactose terminating glycans have at least 1 galactose residue (◯) but no sialic acid and sialic acid terminating refers to glycans with at least 1 sialic acid (◆). Fucosylated glycans have core fucosylation. For all data shown here, error bars represent +/-1 standard error of the mean. WT n = 5, L36P n = 5, ΔCog4 n = 5. Relative abundances were calculated based on the sum of all glycans present in each sample. Statistical significance was assessed using Graphpad Prism software, by two-way ANOVA with Tukeys correction for multiple comparisons (P < 0.05 *).*

3.4.2 Evaluating the Effects of Cog4L36P Mutation or Knockout on Total Cell *N*-glycan Content in Suspension Cell Lines

Comparing the Cog4L36P cell glycan profile to those of both WT and ΔCog4 cells helps understand how altering Cog4 influences the glycan biosynthesis pathway. However, growth conditions can also heavily influence glycosylation. All glycan profiles described so far in this thesis are for adherent CHO-K1 cells grown in monolayer culture. It was important to characterize glycosylation in these cells to help understand the observed differences in their adherent and metabolic properties, which are described for all three cell lines in the previous chapter of this thesis. However, for industrial recombinant protein expression, CHO-K1 cells are typically grown in suspension culture. This creates very different growth conditions that influence glycosylation, as cell surface glycoproteins required for adhesion undergo drastic changes (Walther et al., 2016). Therefore, it was essential to also look at the *N*-glycan profiles of each cell line in suspension culture.

Firstly, glycan profiles for WT, ΔCOG4 and Cog4L36P suspension CHO cells were compared (figure 3.5). The raw mass spectra for these cell types are not shown here, as they broadly resembled those of their adherent counterparts (figure 3.2). Mass spectra for all three cell lines contained peaks corresponding to a mixture of oligomannose and complex permethylated *N*-glycans, with the former being the most prominent. As before, multiple biological replicates were gathered, for each cell line, and used to calculate relative abundances for all the glycans present in each (figure 3.5). Here, the relative abundance of most glycans in the Cog4L36P cells is somewhere between those of WT and ΔCog4. Levels of all oligomannose glycans, except $GlcNAc_2Man_5$ and $GlcNAc_2Man_7$, show a trend of decreasing from WT to Cog4L36P to ΔCog4. The opposite trend is seen for $GlcNAc_2Man_5$. Looking next at the complex glycans, there is a trend of increasing GlcNAc-terminating glycan abundance from WT to L36P to ΔCog4, with significant increases in the abundance of fucosylated GlcNAc-terminating glycan species in the ΔCog4 cells. This is the opposite of the trend seen in the adherent cells. Unlike in the adherent cells, more complex glycans are detectable in ΔCog4 cell samples than in the other two lines, with some glycans, such as $GlcNAc_2Man_3GlcNAc_3$, $FucGlcNAc_2Man_3GlcNAc_4$, $GlcNAc_2Man_3GlcNAc_3Gal$ and $FucGlcNAc_2Man_3GlcNAc_3Gal$ only detected in the ΔCog4 cells.
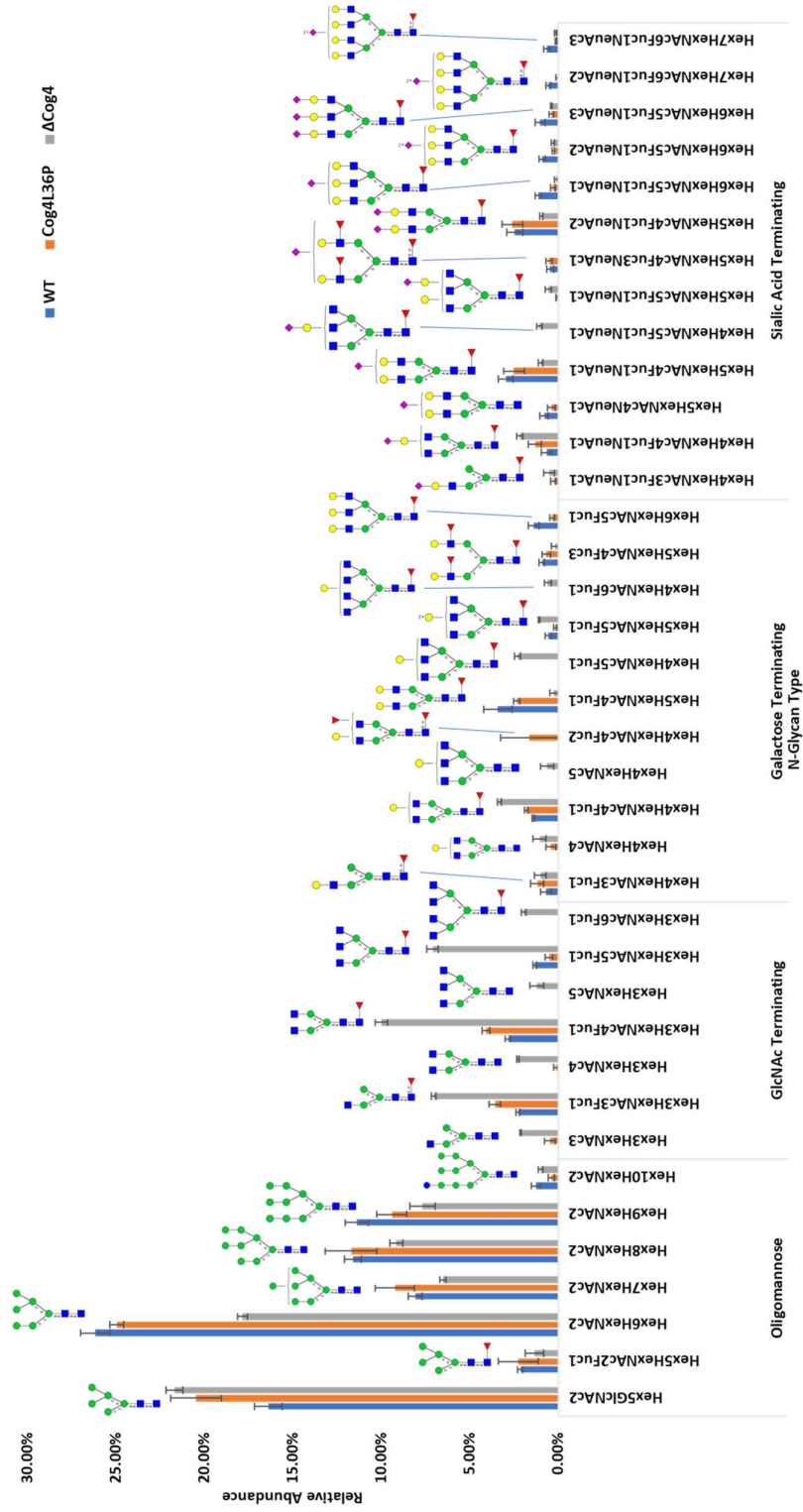
Figure 3.5 | Whole-cell N-glycan profiles of WT, Cog4L36P and ΔCog4 suspension cells. Bar chart showing the relative intensities of the MALDI-MS signals for all N-glycan species detected in each cell line. Spectra were only included for analysis if 10 or more glycans were detected, thus excluding poor quality data. Glycans with relative abundances less than 0.5% were excluded for the purpose of simplifying the data. The glycan signal intensities in each individual spectrum are normalized to the sum of intensities in each spectrum for the subset of glycans that are common to all three cell lines. The X axis shows compositions, with the predicted structures shown above each bar, which were drawn based on knowledge of the biosynthetic pathway. Error bars represent +/- 1 standard error of the mean (WT n = 5, Cog4L36P n = 5, ΔCog4 n = 4).

Again, grouping glycans into their subclasses shows trends in the data more clearly (figure 3.6). Like the adherent cells, suspension WT and Cog4L36P cells have a comparable ratio of oligomannose to complex glycans (figure 3.6). In contrast, ΔCog4 cells have much higher levels of complex glycans than WT or Cog4L36P cells when grown in suspension culture. Similarly, fucosylation is increased in the ΔCog4 cells (figure 3.6), as there are more complex glycans available for fucosylation, indicating that fut8 activity is not impaired by Cog4 dysfunction in suspension cells. This matches findings in the adherent cells. However, the majority of the complex glycans in the suspension ΔCog4 cells terminate in GlcNAc. Strikingly, the levels of *N*-glycans terminating in GlcNAc are more than three times higher in the suspension ΔCog4 cells than the WT. This could suggest that ManI, ManII and Mgat enzymes act faster in the ΔCog4 knockout suspension cells, causing a bottleneck in the medial Golgi where the amount of complex glycans terminating in GlcNAc is too high for complete processing by the available galactosyltransferase. Additionally, the levels of glycans terminating in galactose in all three cell lines are similar despite the overall increase in complex glycans in the ΔCog4 cells. This suggests that knocking out Cog4 in suspension cells impairs galactosyltransferase, which could also explain why there is an accumulation of GlcNAc-terminating *N*-glycans. Sialylation also appears to be affected by Cog4 dysfunction as the levels of sialylated glycans decrease from WT to Cog4L36P to ΔCog4 cells.
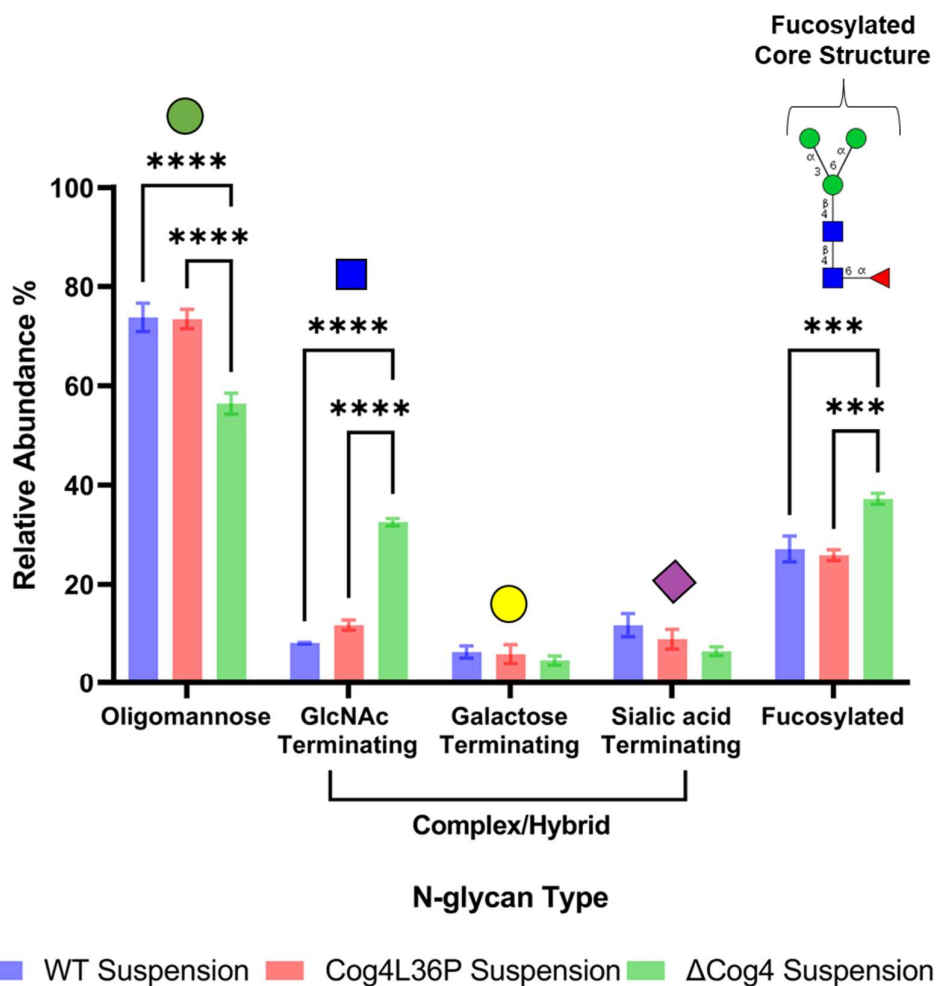
*Figure 3.6 | Relative abundances of each glycan sub-class in suspension WT, L36P and ΔCog4 cells. Bar chart showing average relative abundances of each subclass of glycans. Oligomannose glycans have only mannose residues attached to the core (⬤). Of the complex glycans, GlcNAc terminating refers to glycans with only GlcNAc residues attached to the core (■), galactose terminating have at least 1 galactose residue (⬤) but no sialic acid and sialic acid terminating refers to glycans with at least 1 sialic acid (◆). Fucosylated glycans have core fucosylation. For all data shown here, error bars represent +/-1 standard error of the mean. WT n = 5, Cog4L36P n = 5, ΔCog4 n = 4. Relative abundances were calculated based on the sum of all glycans present in each sample. Statistical significance was assessed using Graphpad Prism software, by two-way ANOVA with Tukeys correction for multiple comparisons (P < 0.001 ***, P < 0.0001 ****).*

3.4.3 Assessing the Impact of Adaptation from Monolayer Culture to Suspension Culture on Whole-cell *N*-glycan Content

By looking separately at *N*-glycan profiles of WT, Cog4L36P and ΔCog4 in monolayer (mCHO) or suspension culture (sCHO), it is clear that adaptation to suspension culture causes changes in the glycan profile. Next, I sought to look more in depth at the changes that happen to the glycan profiles of each individual cell line due to this adaptation. There are several reasons for growing suspension cells, including the ability to culture cells at a much higher density for increased productivity. A key difference between adherent cell medium and suspension medium used for recombinant protein expression is the lack of FBS, which contains calcium required for adherent culture, but is not ideal for soluble recombinant mAb expression due to the presence of bovine IgG in FBS. Cell lines were therefore gradually adapted from monolayer culture in Ham's F-12 medium with 5% FBS to suspension culture in FortiCHO medium lacking FBS. This medium is optimised for CHO-K1 cells producing recombinant proteins in suspension. Since the same parent cells were used in each cell culture environment, glycan profiles of monolayer CHO cells (mCHO) can be directly compared with suspension-cultured CHO cells (sCHO), enabling us to better understand how adaptation to suspension culture alters glycosylation. By plotting adherent vs suspension glycan profiles for each cell line (figure 3.7) and grouping each glycan sub-class for each cell line (figure 3.8) these changes can be seen clearly.

When WT CHO-K1 cells are adapted from adherent to suspension culture, it appears there is a shift in the ratio of $GlcNAc_2Man_5$ to $GlcNAc_2Man_6$, with less $GlcNAc_2Man_5$ in sCHO cells and less $GlcNAc_2Man_6$ in mCHO cells (figure 3.7a). There is also more $GlcNAc_2Man_9$ and $GlcNAc_2Man_9Glc$ in the suspension cells, suggesting that glycan processing in the ER, where glucose residues are removed and oligomannose glycans trimmed down to $GlcNAc_2Man_8$, is altered when the cells are grown in suspension. Elevated levels of these ER glycans could be due to faster, incomplete ER glycan processing, resulting in under-processed oligomannose glycans leaving the ER. Interestingly, levels of *N*-glycans terminating in GlcNAc are lower in WT suspension than they are in adherent cells (figure 3.8a), suggesting that these glycans may be present on adhesion related glycoproteins that are not expressed by the sCHO cells. This low level of GlcNAc-terminated glycans in the Cog4L36P adherent cells was also

observed, meaning their glycan profile is similar to that of WT suspension cells. The relative abundances of most sialylated complex glycans are similar in WT mCHO and sCHO and overall, the amount of sialylation is similar (figure 3.8a) whereas glycans terminating in galactose are more abundant in the sCHO cells.

Cog4L36P mCHO and sCHO cells were compared (figure 3.7b), following observation of weaker adhesion in the L36P mCHO than in WT cells. They also clump less than WT cells when grown in suspension, suggesting they are more suited to growth in suspension culture than WT. Compared to WT and ΔCog4 cells, there are far fewer differences between the mCHO and sCHO glycan profiles of L36P cells (figure 3.7). In fact, almost all glycans show comparable abundance in both the mCHO and sCHO cell data, with no major differences in any one particular glycan standing out. This suggests that adaptation from adherent culture involves fewer changes to glycoproteins in the Cog4L36P cells than in WT.

A

Figure legend: WT Suspension (orange), WT Adherent (blue)

Y-axis: Relative Abundance (30.00%, 25.00%, 20.00%, 15.00%, 10.00%, 5.00%, 0.00%, -5.00%)

X-axis: N-Glycan Type

Glycan categories: Oligomannose, GlcNAc Terminating, Galactose Terminating, Sialic Acid Terminating

Glycan labels: Hex5GlcNAc2, Hex5HexNAc2Fuc1, Hex6HexNAc2, Hex7HexNAc2, Hex8HexNAc2, Hex9HexNAc2, Hex10HexNAc2, Hex3HexNAc3Fuc1, Hex3HexNAc4Fuc1, Hex3HexNAc5Fuc1, Hex4HexNAc3Fuc1, Hex4HexNAc4Fuc1, Hex5HexNAc4Fuc1, Hex5HexNAc5Fuc1, Hex5HexNAc4Fuc3, Hex6HexNAc5Fuc1, Hex4HexNAc4Fuc1NeuAc1, Hex5HexNAc4NeuAc1, Hex5HexNAc4Fuc1NeuAc1, Hex5HexNAc4NeuAc2, Hex5HexNAc4Fuc1NeuAc2, Hex6HexNAc5Fuc1NeuAc2, Hex6HexNAc5Fuc1NeuAc2, Hex6HexNAc5Fuc1NeuAc3, Hex7HexNAc6Fuc1NeuAc2, Hex7HexNAc6Fuc1NeuAc3

158

B

Figure 3.7 | Whole-cell N-glycan profiles of WT, L36P and ΔCog4 adherent vs. suspension cells. Bar charts comparing the relative abundances of each glycan species in CHO-K1 cells grown in adherent vs suspension culture for: A) WT, B) Cog4L36P, C) ΔCog4 cells. This data is also presented in figures 3 and 5- see figure legends for more information about the data processing.

In contrast, adapting ΔCog4 cells from adherent to suspension culture results in far greater changes in glycan profiles than those in either WT or Cog4L36P cells. Levels of almost all the oligomannose *N*-glycans decrease when ΔCog4 cells are adapted to suspension culture (figure 3.7c, figure 3.8a). This decrease corresponds to an increase in the number and abundance of complex glycans expressed (figure 3.8a) and therefore an increase in the abundance of *N*-glycans with core fucosylation (figure 3.8b). The types of complex glycans present in ΔCog4 mCHO and sCHO cells appear to be vastly different (figure 3.7c), suggesting that their adaptation to the two different culture environments is very different.
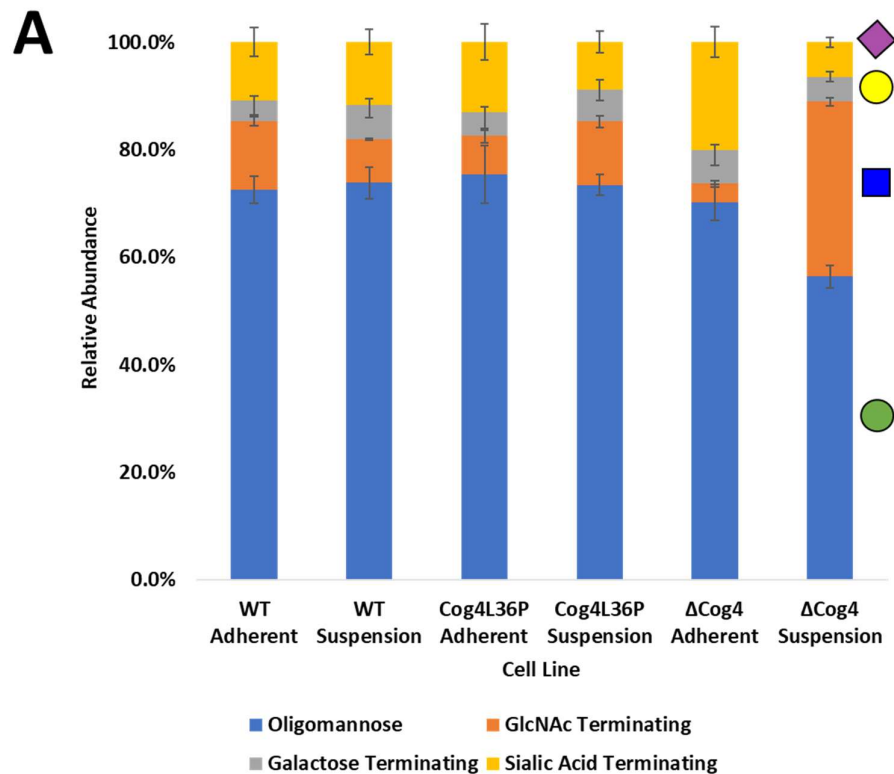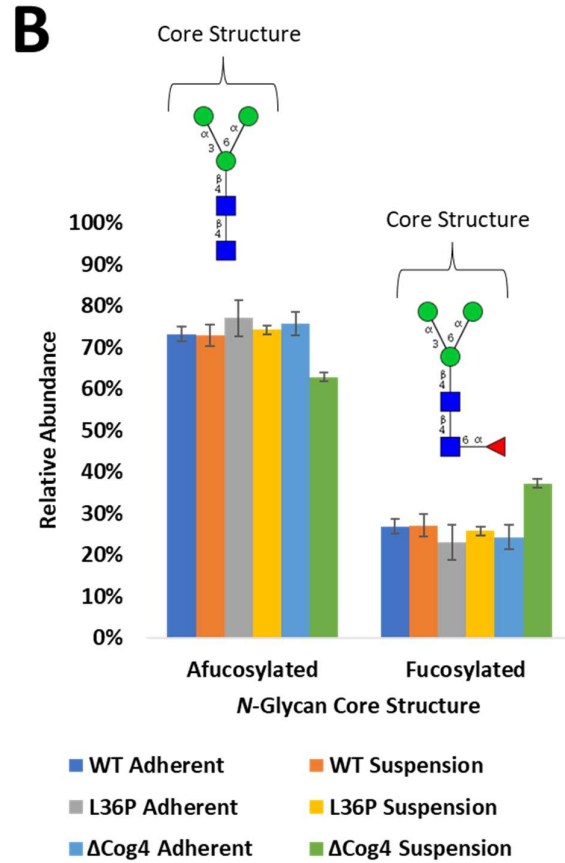
Figure 3.8 | Relative abundances of each glycan sub-class in all cells. A) Stacked bar chart showing average relative abundances of each subclass of glycans. Oligomannose glycans contain only mannose residues attached to the core (●). Of the complex glycans, GlcNAc terminating refers to glycans with only GlcNAc residues attached to the core (■), galactose terminating have at least 1 galactose residue (○) but no sialic acid, and sialic acid terminating refers to glycans with at least 1 sialic acid (◆). B) Bar chart showing the percentage of glycans with core fucosylation in the 3 cell lines. For all data shown here, error bars represent +/-1 standard error of the mean.

3.4.4 A Comparison of the GlycoWorks RapiFluor-MS *N*-glycan Sample Preparation Method, with HPLC Analysis vs. Permethylation with MALDI-MS Analysis and the Need for Industrial Whole-cell Glycan Profiling

In industrial mAb production labs, mAb glycosylation is tightly regulated. Since *N*-glycans can influence the efficacy and immunogenicity of mAbs (Louie et al., 2017, Shields et al., 2002, Sinclair and Elliott, 2005, Zheng et al., 2014, Mimura et al., 2018) , their *N*-glycan profiles are considered to be a critical quality attribute (Beck, 2013, Lauber et al., 2015, Mimura et al., 2018), and it is imperative that they can be evaluated quickly and efficiently. At GlaxoSmithKline (GSK), this is done using the Waters GlycoWorks RapiFluor-MS *N*-glycan kit (Lauber et al., 2015) to cleave and fluorescently label *N*-glycans from purified mAbs for quantitative analysis using high performance liquid chromatography (HPLC) and HPLC-MS (figure 3.1b). A similar approach is adopted across many pharmaceutical companies. An advantage of using this kit over conventional *N*-glycan sample preparation is that the same buffer can be used throughout the process, from protein denaturation to PNGase F treatment and labelling. This abolishes the need for buffer exchange and cuts sample handling time down from several days to just a few hours, saving time and minimising the potential for human error or sample contamination. Quantification can be achieved by measuring fluorescence of the RapiFluor tag, which binds once to each *N*-glycan structure at the reducing-terminal GlcNAc in the core structure, where the protein-linked glycan binds the amide nitrogen atom of an asparagine residue.

HPLC equipment is, however, expensive and a typical sample run takes up to an hour. Chromatographic resolution of the RapiFluor-tagged glycans can be incomplete especially when complex mixtures of glycans are being analysed, and most importantly, structural assignments are based on co-chromatography with authentic standards, preventing ready identification of novel or unexpected glycans. However, the limits of detection of fluorescence detectors are superior to those of mass spectrometers (Mechref et al., 2013, Lauber et al., 2015, Domann et al., 2007), offering a significant advantage of this approach. Furthermore, the approach is only well adopted for mAbs, which have relatively few glycosylation sites and typically display only a handful of glycan structures, making separation of the released glycans straightforward. To separate the larger range of glycans found in a whole-cell lysate, an even longer run

time would be needed to achieve separation, avoiding co-elution. Hyphenation of the HPLC system with mass spectrometry aids in obtaining molecular mass information on the eluting species, but the mass spectrometric fragmentation behaviour of the fluorescent derivative is largely structurally uninformative, so that the addition of a hyphenated mass spectrometer does not solve the problem of identifying novel glycan species or those for which authentic standards are not available.

By comparison, MALDI-TOF of FANGS-released permethylated glycans, the approach widely adopted in academia because of its ability to provide relative quantification at the same time as robust structural analysis (figure 3.1a), is cheaper and once the sample has been prepared and spotted, spectral acquisition can be achieved in seconds. However, in industry, automation of sample preparation is favoured, and labs are already set up for HPLC-MS, whereas MALDI instruments are not routine in industrial labs, so the glycan profiling method used in academia is not viable.

Whole-cell glycan profiling is not routinely done in industrial mAb production labs. However, when working with new cell lines, it is highly informative to gain an overview of how the mutations impact whole cell glycosylation before selecting cell lines as hosts for recombinant mAb expression. Therefore, the aim was to adapt this well-established mAb glycan profiling procedure for use on whole cells. For mAb glycan profiling with the Waters RapiFluor-MS *N*-glycan kit, the first step in the procedure is to denature the glycoprotein using Waters' Rapigest buffer, in which the sample is re-suspended and heated to at least 90 °C for 3 minutes. This buffer was used in an attempt to lyse whole cells. However, in the absence of information on the proprietary Rapigest buffer composition, it was not clear whether the detergent component would be strong enough to break down cellular membranes and release membrane-associated glycoproteins. Therefore, ten times the recommended RapiGest buffer volume was used as well as increasing the length of the heat denaturation step, introduced mechanical disruption using a sonic bath and incubated with PNGase F for longer than in the manufacturer's protocol.
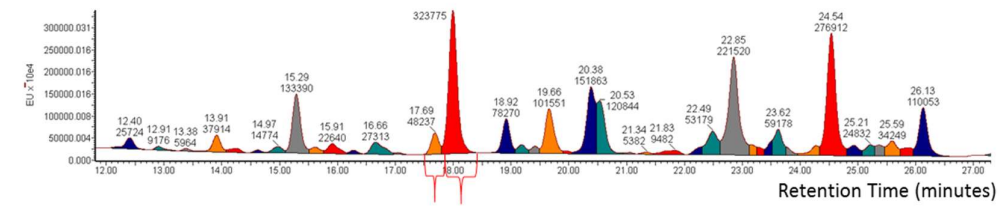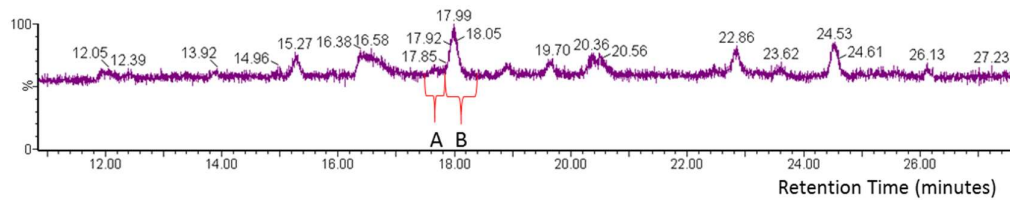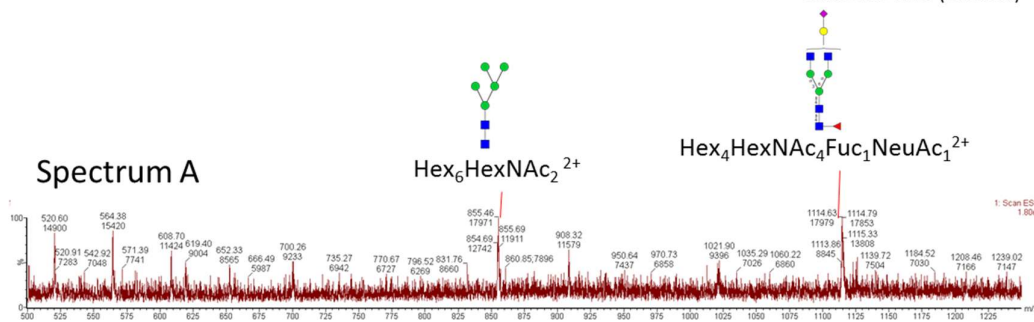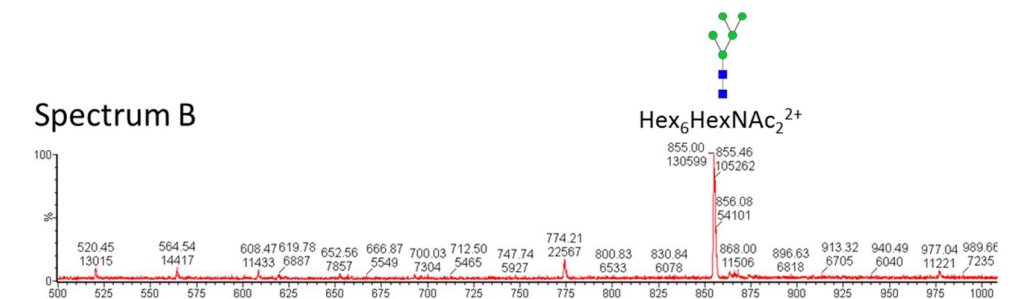
Figure 3.9 | Representative results from using HPLC-MS to measure whole cell N-glycans, prepared using Waters GlycoWorks RapiFluor-MS N-glycan kit. The top panels show a section of the fluorescence chromatogram for a WT CHO-K1 N-glycan sample, with peak areas each highlighted in a different colour. The corresponding total ion chromatogram (TIC) below. The bottom panels show spectra corresponding to the regions of the TIC that are highlighted as A and B. These regions were selected based on retention times of fluorescence peaks in the fluorescence chromatogram. Peaks were assigned based on the predicted mass of the charged (+2 for doubly charged or +3 for triply charged) glycan (sum of monosaccharide in chain residue masses) bearing the RapiFluor tag (mass increment 329 Da).

E.g. $Hex_6HexNAc_2 = (1378 + 2 + 329)/2 = 854.5$.

3.4.5 Optimisation of a New Whole-cell Glycan Profiling Sample Preparation Method Using Waters Glycoworks Kit

Using the Glycoworks method (figure 3.1b) on the WT CHO-K1 suspension cultured cells, multiple fluorescence peaks corresponding to *N*-glycans labelled with the RapiFluor tag (figure 3.9) were detected. Like the glycan profiles generated using the permethylation and MALDI approach, these included a mixture of oligomannose and complex glycans (figure 3.10). However, there were multiple glycans, including $GlcNac_2Man_6$ and $FucGlcNAc_2Man_3GlcNAc_2Gal_1NeuAc_1$ (figure 3.9) which co-eluted and these cannot therefore be distinguished for quantification. Nevertheless, we can compare glycan profiles generated using the Glycoworks method (figure 3.1a, figure 3.9) with those generated using the permethylation and MALDI approach (figure 3.1b), with the caveat that co-eluting glycans cannot be distinguished in the Glycoworks approach (figure 3.10).

Use of the permethylation and MALDI method results in a glycan profile dominated by oligomannose glycans, with multiple complex glycans detected at lower abundance. Oligomannose glycans are also highly abundant in the glycan profile generated using the Glycoworks method, however the relative levels of each are very different from those in the MALDI profile, with a particularly striking disparity in the relative abundance of $GlcNAc_2Man_5$ (figure 10). All the complex glycans detected by the MALDI method are less abundant than any of the oligomannose glycans, with the exception of $Fuc_1GlcNAc_2Man_5$. In contrast, the Glycoworks profile includes at least two complex glycans ($GlcNAc_2Man_3GlcNAc_4Gal_4$ and $FucGlcNAc_2Man_3GlcNAc_2Gal_2NeuAc_1$) that are more abundant than either $GlcNAc_2Man_5$ or $GlcNAc_2Man_7$. The presence of the former complex glycan is particularly interesting as this was not detected at all using the MALDI method. The Glycoworks method also detected two large complex glycans, $Hex_9HexNAc_8$ and $Fuc_1Hex_9HexNAc_8$ (albeit with other co-eluting glycans) (figure 3.10), that were not detected at all using the MALDI method (figure 3.5 and 3.7a). On the same note, the MALDI method detected multiple low-abundance complex glycans, such as $FucGlcNAc_2Man_3GlcNAc_3Gal_3$ and $FucGlcNAc_2Man_3GlcNAc_2Gal_2NeuAc_2$, that were not detected by the industrial method. In summary, despite both methods being used on the same cell line, from two different cultures grown in the same conditions, the two approaches result in very different glycan profiles (figure 3.10).
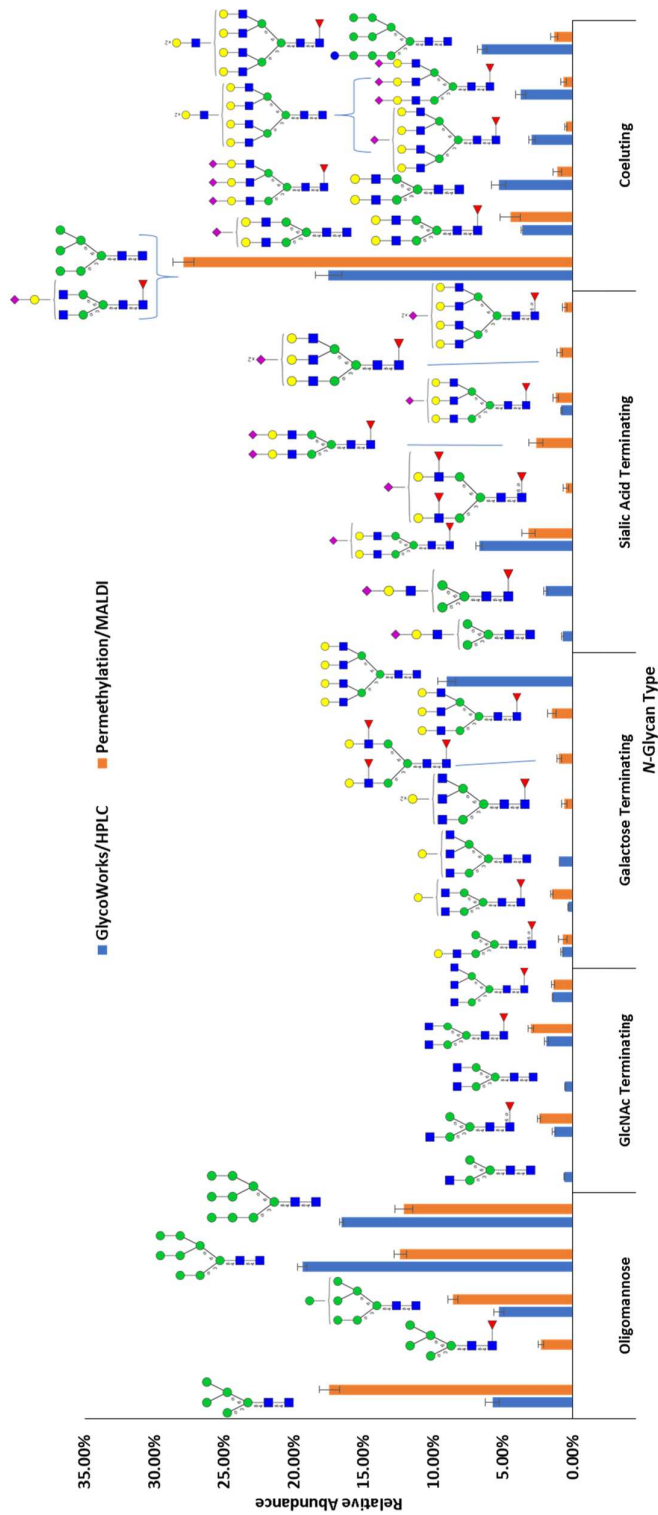
Figure 3.10 | Whole-cell N-glycan profiles for WT CHO-K1 suspension-cultured cells, obtained using either the Waters GlycoWorks RapiFluor N-glycan kit and HPLC detection (GlycoWorks/HPLC), or FANGS-release, permethylation and detection by MALDI-MS (Permethylation/MALDI). Bar chart showing the relative abundances of N-glycan species detected using each method. For the industrial sample profile, abundances were calculated based on peak area from the fluorescence chromatogram. For the academic sample profile, abundances were calculated based on the m/z peak intensities. Relative abundances were calculated only from glycans common to both profiles. Only glycans with relative abundances over 0.5% were included in the plot. Each bar represents the average (GlycoWorks/HPLC method N=3, permethylation/MALDI method N=5) and error bars represent +/- 1 standard error of the mean. Predicted structures shown above each bar, which were drawn based on knowledge of the biosynthetic pathway.

167

3.5 Conclusions and Future Work

Through generating and characterising a CHO-K1 cell line with a novel L36P mutation in Cog4, described in the previous chapter of this thesis, multiple phenotypic differences were observed compared to both WT and ΔCog4 cells. Briefly, these included weaker adhesion and enhanced metabolic activity. In light of these observations, as well as the disruption the mutation causes to the interaction between Cog4 and Rab30, how the mutation impacts glycan biosynthesis was investigated. Whole-cell glycan profiling on the Cog4L36P was carried out and compared to both WT and ΔCog4 cells. To help understand the link between glycosylation and the adhesion defect observed in the Cog4L36P cells, glycan profiles for all three cell lines cultured in both adherent and suspension conditions, were generated.

Higher relative levels of oligomannose glycans were observed and therefore lower relative levels of complex glycans overall in Cog4L36P adherent cell lysates than in those of the other two cell lines. This increase in oligomannose glycan content in Cog4L36P cells, and the discrepancy between the relative levels of $GlcNAc_2Man_5$ and $GlcNAc_2Man_6$ in WT and Cog4L36P cells, suggests that normal COG complex function is important for glycan processing in the *cis* Golgi, where oligomannose glycans are trimmed. It could also suggest that processing of oligomannose to complex glycans happens in the medial Golgi is impaired in the Cog4L36P cells. In contrast, in both mCHO and sCHO ΔCog4 cells, the proportion of many complex glycans was higher than in WT cells, which shows that knocking out Cog4 affects glycosylation in a different way than the Cog4L36P mutation does.

Glycosylation is known to be an important regulator of cell adhesion (Hall et al., 2014), so it is unsurprising that different glycan profiles in adherent-cultured and suspension-cultured cells were observed. This was particularly interesting for Cog4L36P cells, which have weaker adhesion than WT. In this case, the sCHO glycan profile resembled the mCHO profile more closely than those from either WT or L36P cells, suggesting the glycosylation of adhesion molecules is deficient in L36P cells.

These findings not only highlight the importance of Cog4 for normal glycosylation in CHO-K1 cells, but also the importance of growth conditions, which are known to affect glycosylation of recombinant proteins (Heffner et al., 2018, Lalonde and Durocher,

2017). Understanding the precise molecular mechanisms leading to glycosylation changes in Cog4L36P cells is, however, not trivial. Glycosylation is tightly regulated by a huge number of factors, including enzyme localization in the Golgi and substrate availability (Stanley, 2011). One possible way of dissecting the specific changes to the glycan biosynthetic pathway caused by Cog4 mutations is inputting these results into a computational model that predicts enzyme rates based on whole-cell glycan profiles (Fisher et al., 2019). This is something our lab would like to explore in future, but it is beyond the scope of this thesis. Other future work should include generating multiple clonal cell lines for each genotype and generating glycan profiles for each, to determine the extent to which natural clonal variation contributes to whole cell glycan profile changes observed between WT, ΔCog4 and Cog4L36P cells.

In this chapter, I have also begun to show the potential of the Glycoworks method, which is used by GlaxoSmithKline for *N*-glycan analysis of mAbs, for the analysis of whole-cell glycosylation. This method is typically used for mAbs but it has also been used to study other common glycoproteins, including fetuin, lactoferrin and ribonuclease B, and for the analysis of a cancer cell membrane proteins (Sun et al., 2017). It has also been used for *N*-glycan profiling of blood serum from CDG patients (Sturiale et al., 2019, Messina et al., 2021). However, to my knowledge, it has never been used for whole-cell glycan profiling. The Glycoworks/HPLC method was able to detect a broad range of oligomannose and complex glycans in WT CHO-K1 suspension cells, though overall the profile was quite different to that generated using the FANGS/MALDI method. The disparity in the glycan profiles produced by the two methods is most likely because of incomplete solubilisation of whole cells to generate lysates using Rapigest. This would result in the digestion of fewer glycoproteins, causing lower abundance glycans to be missed. Nevertheless, it is promising that the method can be used to detect such a broad range of *N*-glycans from whole cells.

Further optimisation of the Glycoworks/HPLC method is clearly required to both the cell lysis procedure and the HPLC separation, as low abundance glycans were not detected and absolute quantification is currently compromised due to the extent of co-elution. In future, it would be interesting to combine the initial lysis step of FANGS with the PNGase F treatment and RapiFluor labelling steps of the GlycoWorks method. This

would require a buffer exchange step, but it would use the FANGS approach which has optimised digestion of the cell membranes and glycoprotein solubilisation, whilst exploiting the enhanced limit of detection of the fluorescence detector used in the HPLC method, compared to MALDI. Additional optimisation of the HPLC separation is also required and a good start would be increasing the length of the run to help reduce co-elution.

In conclusion, glycosylation in the Cog4L36P cells has been identified to be distinct from that of both WT and ΔCog4 cells. Adaptation from adherent to suspension culture had little effect on the glycan profile of Cog4L36P cells, compared to more noticeable effects on the glycan profiles of WT and ΔCog4 cells. Also, an HPLC-fluorescence detection method for whole-cell glycan profiling has been investigated. A later aim of this thesis will also be to explore how whole-cell glycan profiles relate to glycan profiles of recombinantly expressed proteins.

Chapter 4: Examining the Potential of Cog4L36P Mutant Cell Lines as Host Cells for Glycoprotein Biologics

4.1 Introduction

4.1.1 Selecting Host Cells for the Production of Glycoprotein Biologics

Glycoprotein biopharmaceuticals have revolutionized treatment of cancers and autoimmune disease and the biologics industry is now worth hundreds of billions of dollars ($US), of which a significant proportion comes from mAbs alone (Walsh, 2014). Considering their high demand and the need to efficiently produce high titres, with stringent quality control measures in place, host-cell selection is an important consideration during mAb production (Lalonde and Durocher, 2017). The work described in this thesis focusses on CHO cells, which are the most common host cells for therapeutic mAb production (Brooks, 2004). The process of producing mAbs in CHO cells for therapeutic use involves making stable integrations of mAb-encoding DNA into the host cell chromosomal DNA (Lalonde and Durocher, 2017). Since this DNA can be incorporated anywhere within the genome, a mandatory step in gaining FDA approval for new mAb therapies is single cell cloning, to ensure that all progeny cells have the mAb plasmid integrated in the same place, minimising heterogeneity, and differences in expression. Typically, hundreds to thousands of clones are screened and their production rates monitored, to select the highest, stably producing cells for further development (Li et al., 2010).

Several techniques are exploited to select high producing clones. These include antibiotic selection and the use of metabolically selectable markers, which are well reviewed in (Li et al., 2010). However, clonal isolation and selection of stable mAb producing clones is time consuming and poses a major challenge in the production of therapeutic proteins. An alternative approach to testing new potential host-cell lines is

transient transfection, which skips the lengthy selection process for clones with stably integrated plasmid DNA (reviewed in Rita Costa et al., 2010). Unlike stably transfected DNA, transiently transfected DNA is maintained as an extra-chromosomal unit, which is rapidly lost. Nevertheless, transient transfection enables production of milligram to gram quantities of mAb in a short time (Liao and Sunstrom, 2006, Rosser et al., 2005). Although not suitable for large-scale production, transient transfection is very useful for high-throughput screening for host cells potentially capable of producing mAbs with desirable properties, including glycosylation (Lalonde and Durocher, 2017).

Glycosylation is a particularly crucial factor in biologic development, as non-human glycoforms can trigger unwanted immune responses in humans receiving the drug (De Groot and Scott, 2007) and the presence or absence of specific glycan substituents can influence therapeutic properties (Liu, 2015). For therapeutic mAb production, *N*-glycosylation of the IgG Fc region is considered a critical quality attribute (Beck, 2013, Lauber et al., 2015, Mimura et al., 2018). Selecting host cells that perform human-like glycosylation is therefore imperative, and this is a major reason for the popularity of CHO cells for mAb production, over other expression systems such as plants, bacteria, yeast and insects, which can all produce non-human glycoforms (Brooks, 2004). Human cells, such as human embryonic kidney 293 (HEK293) cells, are also used in the production of a few FDA approved biologics (Dumont et al., 2016). However, despite their capacity for human glycosylation, human cell biologic production lags behind that in non-human mammalian cells due to the risk of infection and transmission of viruses capable of infecting humans (Dumont et al., 2016). In contrast, CHO cells have reduced susceptibility to some viral infections compared to human or other mammalian cell lines, and risk to humans of potential viral contamination of CHO cell-produced biologics is minimised by the species barrier (Berting et al., 2010).

4.1.2 Rationale for Testing Cog4L36P Mutant CHO-K1 Cells as Biologic Host Cells

The ability to engineer glycosylation machinery of host cells to yield secreted glycoproteins with specific glycoforms having desirable properties, and to produce homogenous batches of biologics with minimised glycan heterogeneity is an attractive prospect for the biopharmaceutical industry (Gupta and Shukla, 2018). The work

presented in this thesis offers a new potential approach to altering glycosylation, by targeting the COG complex, an important coordinator of glycan processing enzyme sorting in the Golgi (Blackburn et al., 2019). Mass spectrometric whole-cell glycan profiling (chapter 2) revealed that both adherent and suspension-cultured Cog4L36P cells have a glycan profile distinct from either WT or ΔCog4 cells, suggesting that glycosylation of recombinantly expressed protein may be similarly altered in Cog4L36P cells when compared with WT.

In addition to altered glycosylation at a whole-cell level, the work presented in this thesis so far suggests that the Cog4L36P mutation confers properties that may be beneficial for biologic production. Cog4L36P cells exhibited enhanced metabolic activity compared to their WT parent cells (section 2.5.3). Higher metabolic activity, specifically oxidative phosphorylation, has been linked to increased productivity of mAb-expressing cells (Link et al., 2004, Templeton et al., 2013). Link et. al found that levels of production of recombinant MUC1 fusion protein in CHO cells was dependent on the rate of oxygen uptake rate in the cells, indicating a positive influence of the rate of oxidative phosphorylation on productivity (Link et al., 2004). Templeton et al. also used [13]C- labelling and metabolic flux analysis to show that for mAb-expressing CHO cells grown in fed-batch culture, peak antibody production corresponded to a highly oxidative state of metabolism (Templeton et al., 2013).

When grown in monolayer culture, it also became apparent that the Cog4L36P mutation caused defective adhesion, with evidence suggesting both cell surface adhesion molecules and ECM secreted by Cog4L36P cells were less adhesive than those associated with the WT parent cells. These findings were demonstrated by impaired binding of ΔCog4L36P cells to vitronectin and weaker binding of both ΔCog4WT and ΔCog4L36P to ECM secreted by ΔCog4L36P cells (section 2.5.4). Weaker adhesion is favourable for recombinant protein production, as biologics are produced in suspension culture, where minimising contact between cells or between the cells and the culture vessel is beneficial for recombinant protein secretion. The work presented in this chapter therefore aims to examine both production levels and glycosylation of a recombinantly expressed biologic: the readily available mAb, Herceptin.

4.1.3 Rationale for Engineering the Cog4L36P Mutation in an Industrial Legacy CHO-K1 Host Cell Line

Suspension cell culture is a requirement for industrial production of biologics. Growth in suspension offers the advantage of easier scale up in bioreactors and therefore enhanced production of recombinantly expressed biologics, as cells can grow at a much higher density than in monolayer culture (Wurm, 2004). Modern intensified fed-batch production runs can achieve viable cell densities of up to of 1×10^8 cells/mL (Rittershaus et al., 2022, Xu et al., 2020), with product titres exceeding 10 g/L (Handlogten et al., 2018, Ha et al., 2022). In contrast, growth in monolayer culture is usually limited by the surface area of the culture vessel. Micro-carrier beads can be used to increase the available surface area for cell attachment (Panchalingam et al., 2015), however these occupy a volume much larger than that of single cells in suspension, so adherent cultures can't reach the same densities as suspension cultures. Furthermore, the presence of animal serum, typically of foetal bovine origin, is routinely added to enhance growth of adherent cells and the calcium it contains promotes adhesion (Ham, 1965). Adaptation to serum-free medium used for suspension culture is advantageous for several reasons, including minimising risk of pathogenic contamination from animal-derived sera, limiting batch-batch variation and reducing cost (Sinacore et al., 2000). The absence of serum, which contains animal-derived secreted proteins including IgGs, in undefined quantities, also simplifies downstream processing of recombinantly expressed proteins, which do not have to be separated from similar serum proteins.

CHO cells, as originally isolated and cultured (Puck et al., 1958, Puck, 1985), are an adherent epithelial cell line, with elongated morphology often described as fibroblast-like (Puck et al., 1958). CHO-K1 cells, used for most work presented in this thesis, are a glycine-dependent clone of the original parental CHO cell line (Kao and Puck, 1968), which have become a standard in mammalian cell culture. Production of recombinant proteins in CHO-K1 and CHO-K1-derived cells requires their adaptation from adherent to suspension culture, a process involving gradually lowering the serum concentration in culture medium and transferring cells to shake-flasks (Sinacore et al., 2000).

174

It was useful and of interest to compare titres of the biologic, Herceptin, in WT and Cog4L36P suspension-adapted CHO-K1 cells, which were taken through the adaptation process in parallel and so subjected to the same growth conditions throughout their culture, enabling direct comparison between the cell lines. However, CHO cells are inherently genetically plastic and adaptation to suspension culture likely accelerates the accumulation of mutations and epigenetic change, which may lead to variation in production levels over time. Furthermore, these cells are unlikely to achieve titres of recombinantly expressed biologics comparable to industrially used production cell lines. To overcome this variation and to assess the effects of the Cog4L36P mutation in cell line capable of high titre biologic production, it was beneficial, and indeed timesaving, to engineer the mutation in a more established suspension-adapted production cell line. Industrial production cell lines are the product of a careful selection process to optimise growth of naturally adherent CHO cells in suspension culture, and to select for high titre biologic production. Therefore, in collaboration with GlaxoSmithKline, the Cog4L36P mutation was engineered into a proprietary, legacy (retired from production) suspension CHO cell line, provided by the company, and the effects of the mutation on production of Herceptin were investigated.

4.2 Aims

This chapter describes work to put the overall aim of this research, to develop a new method for potentially generating glycan-engineered biologics, to the test by assessing the production and glycosylation of biopharmaceuticals in Cog4L36P mutant suspension CHO-K1 cells compared to that in their WT parent cells. This involved monitoring the production of a readily available mAb, Herceptin, in Cog4L36P and WT cells by performing transient transfection and analysing the recombinantly produced protein. Another aim of the work presented in this chapter was to engineer the Cog4L36P mutation in an ex-production CHO-K1 suspension cell line, with properties similar to mAb-production host cells, provided by GSK. Levels of Herceptin production were also monitored in this cell line and compared to those of its WT parent cell line to test the potential of the Cog4L36P mutation for mAb production in cells optimised for high titre mAb production, resembling cells used during industrial production.

4.3 Materials

4.3.1 Cell Lines

ΔCog4WT and ΔCog4L36P cells (table 2.2) were adapted to suspension culture, as described in section 2.4.4, by Elizabeth Willows (Ungar group). Cog4L36P CHO-K1 cells were engineered by CRISPR-Cas9 and adapted to suspension culture as described in chapter 2, and WT CHO-K1 cells were also adapted to suspension culture as described in chapter 2. Industrial host WT CHO-K1 cells are a legacy production suspension cell line provided by GlaxoSmithKline and transferred to the University of York for work carried out in the Department of Biology.

4.3.2 Plasmids

The Herceptin plasmid was from Cobra Biologics. See appendix 1 for the heavy chain amino acid sequence. pcDNA3-eGFP plasmid was originally purchased from Clontech (pEGFP-C3). Before transfection, plasmids were purified by midi-prep as described in section 2.4.6.2.

4.4 Methods

4.4.1 Suspension Cell Culture

Suspension cells were cultured as described in section 2.4.2. Host CHO-K1 cells from GSK were cultured in CD CHO medium (Gibco) supplemented with 8 mM GlutaMAX™ (Gibco) (host suspension cell medium).

4.4.2 Suspension Cell Transient Transfections

Cells were transfected using one of two transfection protocols using FreeStyle™ MAX reagent, detailed below. Following transfection by either protocol, cells were incubated in transfection medium for 24 hours before changing the medium for 30 mL of the usual suspension medium used for cell culture (see sections 2.4.2 and 4.4.2 for media compositions) by pelleting the cells at 1,000 $x\ g$ for 5 minutes, aspirating the supernatant and re-suspending the pellet in fresh medium. Herceptin expression medium was harvested four days post-transfection. Medium was harvested by centrifuging the entire culture at 1000 $x\ g$ for 5 minutes and transferring the

supernatant to a clean tube before snap freezing in liquid nitrogen and storing at -80 °C.

4.4.2.1 FreeStyle MAX Transfection: Manufacturer's Protocol

Cells were transfected following the manufacturer's protocol. This involved the following stages: 24 hours before transfection, suspension cells were seeded at 1×10^6 cells/mL in FreeStyle^{TM} medium supplemented with 8mM GlutaMAX™. Cells were transfected using FreeStyle^{TM} MAX Reagent (Gibco), following the manufacturers guidelines. Briefly, cells were diluted to a density of 1×10^6 cells/mL in 30mL medium on the day of transfection. DNA-lipid complexes, consisting of 37 μg DNA and 37 μL transfection reagent, were prepared in OptiPRO^{TM} serum free medium (SFM), incubated at room temperature for 25 mins and added dropwise into each suspension cell flask.

4.4.2.2 FreeStyle MAX Transfection: Alternative Protocol

Following a placement carried out at GlaxoSmithKline, an alternative FreeStyle™ MAX transfection protocol was followed as this was found to result in superior transfection efficiency. In this method, cells were seeded at 1×10^6 cells/mL in 15 mL OptiPRO™ SFM supplemented with 8 mM GlutaMAX™ in 125 mL Erlenmeyer flasks on the day of transfection. FreeStyleMAX™ transfection reagent was diluted in OptiPRO™ SFM supplemented with 8 mM GlutaMAX™ by adding 25 μL to 600 μL medium per transfection. DNA was also diluted in the same medium by adding 18 μg DNA to 600 μL medium per transfection. Immediately after diluting the reagents, the diluted FreeStyleMAX™ mixture was added to the diluted DNA in a 1:1 (v:v) ratio and mixed by pipetting. These mixtures of transfection reagent and DNA were incubated for 15 minutes at room temperature before adding 1.2 mL of each transfection mixture to the relevant flask.

4.4.3 Monitoring GFP Expression by Flow Cytometry

Alongside transfections with either CRISPR reagents or Herceptin plasmid, separate control transfections were performed using an eGFP plasmid. Transfection efficiency was estimated based on the percentage of green fluorescent protein (GFP) expressing cells. GFP expression was monitored by flow cytometry using either a CyAn™ ADP (Beckman Coulter) or a CytoFlex LX flow cytometer (Beckman Coulter) in semi-

automatic (tube) mode. 100 µL samples of GFP transfected cells were transferred to microfuge tubes and diluted in 300 µL PBS. For measurements on the CyAn™ ADP cytometer, DAPI was also added at a concentration of 1 µg/mL. Tube contents were aspirated into the flow cytometer at a rate of approximately 60 µL per minute. GFP fluorescence was excited by the 488 nm laser. On the CytoFLEX LX cytometer GFP fluorescence was detected with the 525/40 reflective band pass (rBP) emission filter and avalanche photodiode detector: on the CyAn™ ADP cytometer GFP fluorescence was detected with a 530/40 BP emission filter and photomultiplier detector (PMT). Viable cells were identified (above a forward scatter threshold) using a forward scatter (FS) / side scatter (SS) dot plot on the CytoFLEX or by the exclusion of DAPI from the cell on the CyAn™ ADP cytometer. DAPI was excited by the 405 nm laser and emitted fluorescence was detected with a 450/50 BP emission filter. Cellular doublets were excluded from analysis gates using a SS-height / SS-area dot plot or a FS-height / FC - area dot plot. GFP cells were then identified using a FS / GFP dot plot or GFP histogram that was gated on single cells. 10,000 events (total cell count) were recorded for each sample.

4.4.4 CRISPR-cas9 Genome Editing

Industrial host WT CHO-K1 cells were transfected with the same CRISPR-cas9 reagents used to engineer the adherent Cog4L36P CHO-K1 cell line. Design and preparation of these reagents is detailed in section 2.4.5.

4.4.4.1 Industrial Host Cell CRISPR-cas9n XFect Transfection

Industrial host WT CHO-K1 cells were transfected with CRISPR-cas9n reagents using XFect transfection reagent (ClonTech), following the transfection and puromycin selection protocols described in section 2.4.5.3. Transfection using XFect reagent was carried out in triplicate wells of a 6-well plate.

4.4.4.2 Industrial Host Cell CRISPR-cas9 FreeStyle™ MAX Transfection

Industrial host WT CHO-K1 cells were also transfected with CRISPR-cas9 reagents utilising the Freestyle™ MAX GSK transfection protocol (section 4.4.2.2), in triplicate wells of a 6 well plate. The method was scaled down for transfection of cells in a 6-well suspension culture plate, in 1 mL OptiPRO medium (all quantities for transfection in 125mL shake flasks were divided by 15 for scaling down). This required 1.2 µg DNA per

transfection, which consisted of 315 ng of each gDNA-Cas9n plasmid and 570 ng ssODN template. 24 hours after transfection, cells were pelleted and re-suspended in 2 mL CD CHO suspension culture medium, supplemented with 8 mM glutamax and 15 µg/mL puromycin. Cells were left in puromycin selection medium for 72 hours before single cell cloning.

4.4.4.3 Flow Cytometric Single Cell Cloning of CRISPR-cas9-Transfected Industrial Host Cells

Four days after transfection, triplicate wells from each transfection were pooled and counted by diluting a 300 µL sample with 300 µL TrypLE before analysis using a Vi-CELL™ XR (Beckman Coulter). Half of each pooled transfected cell culture (3 mL) was centrifuged at 1,000 $x\ g$ for 5 mins, re-suspended in 2 mL host suspension cell medium with sterile filtered DAPI stain added at a concentration of 1 µg/mL and transferred to sterile, lidded 5 mL flow cytometry tubes. The remaining cells were kept in culture, topped up to 10 mL volume with host suspension cell medium and allowed to grow to a density of 2 x 10^6 cells/mL before freezing following the protocol outlined in section 2.4.3.

Transfected cells were cloned to single cells by Karen Hogg (York Technology Facility). Single cell sorting was performed by flow cytometry on a CytoFLEX SRT (Beckman Coulter), fitted with a 100-micron nozzle and operating in single cell mode with 0.5 drop envelope. Cells were identified (above a forward scatter threshold) using a forward scatter (FS) / side scatter (SS) dot plot. Doublets were excluded from the sort gates from a SS-height / SS-area dot plot: doublets have a larger signal area than single cells. Live single cells were identified as DAPI negative by 405 nm laser excitation (70 mW), 450/45 reflective band pass emission filter and avalanche photodiode detector using a FS / DAPI dot plot gated on single cells. A single DAPI negative cell per well was sorted into the inner 60 wells, 96-well plate (Sarstedt 83.3924.500) containing 100 µL medium per well consisting of host suspension cell medium with 60% v/v spent medium (harvested from 3-day-old host-cell culture and sterile filtered). In between each transfected cell sample, ten back-flushes were performed, and the system was flushed with autoclaved sheath fluid to verify the sample line as clear before subsequent samples were sorted to avoid carry over.

4.4.4.4 Expansion and Genotyping of CRISPR-Cas9n-Transfected Clones

Following single-cell sorting, 96-well-plate clones were examined under a microscope for established colonies. When > 50% of the well surface was covered with cells, the entire well contents were transferred to wells of a suspension 24-well plate and topped up to 1 mL volume with host suspension cell medium. After 1 week in 24-well-plate culture, 500 µL of the culture was harvested at 800 $x\,g$ for 5 minutes and DNA extracted using QIAGENs DNeasy blood and tissue kit (section 2.4.6.2) for genotyping using the methods described in section 2.4.6. Briefly, PCR reactions were run, and DNA purified (section 2.4.6.5) and analytical restriction digests carried out (section 2.4.6.3). Analytical DNA gels were run to compare digested and non-digested DNA from each clone (section 2.4.6.4) and PCR products were sent for sequencing using Eurofins TubeSeq service with CRISPR Rev primer (table 2.3). The remaining volume of cell culture was transferred to a well of a 6-well plate and topped up to 2 mL volume with host suspension cell medium.

Following verification of successful CRISPR editing, Cog4L36P-positive clones were allowed to grow to approx. 2×10^6 cells/mL in 6-well plates before expansion into T25 suspension flasks, which were topped up to 7 mL volume with host suspension cell medium. Cells were again allowed to grow to approx. 2×10^6 cells/mL before expansion into 30 mL culture volume in 125 mL shake flasks. Five back-up frozen stocks of each cell line were made, as described in section 2.4.3, before the cells were used for experimentation.

4.4.5 Enzyme Linked Immunosorbent Assay (ELISA)

Herceptin-expression-medium samples (600 µL volume) were collected every day post-transfection by 5 min 1000 x g centrifugation and stored at -80 °C prior to thawing, dilution with PBS and analysis by ELISA. Pierce^{TM} protein A-coated white 96-well plates (ThermoFisher), which were pre-blocked with SuperBlock blocking buffer (ThermoFisher), were washed three times with PBST. Serial dilutions of human IgG standards (Sigma Aldrich I3506), ranging from 5 ng to 100 ng in PBS, and 100 µL diluted Herceptin-expression-medium samples were added to plates in triplicate. Plates were sealed and incubated for 1 h at RT on an orbital shaker. Wells were washed three times with PBST before adding HRP-conjugated chicken polyclonal antibody recognising

human IgG (Abcam 112454), diluted 1:1000 in SuperBlock blocking buffer, containing 0.05% Tween-20. Plates were sealed and incubated as before. Wells were washed three times with PBST before adding Supersignal ELISA pico luminol detection agent (ThermoFisher 1859677). Plates were read on a CLARIOstar (BMG LABTECH) plate reader at 425 nm. Measurements from medium samples were calibrated against a standard curve, which was plotted in Microsoft Excel from human IgG standard measurements.

4.4.6 Herceptin Purification

Four days post-transfection, medium from Herceptin-transfected cells was harvested by centrifugation at 1000 x g for 5 mins at 4 °C, before passage through a 0.2 µM syringe filter. Herceptin was purified using the Akta 3D system (GE Healthcare). A 1 mL HiTrap^{TM} Protein L column (GE Healthcare) was pre-equilibrated with 20 mM sodium phosphate (binding buffer), pH 7.2. Filtered Herceptin medium was loaded onto the column, before washing with binding buffer and elution with 0.1 M glycine, pH 2.5 into 1 M Tris-HCl, pH 8, to neutralize, using a ratio of 1:4 Tris:eluate volume. Solution containing purified Herceptin was concentrated to approximately 100 µL volume into 20 mM ammonium acetate using a Vivaspin 4 centrifugal concentrator (Sartorius) at 4000 x g. The Herceptin in the concentrated solution was reduced by adding 500 µL 2 mg/mL DTT in 20 mM ammonium acetate and incubating at room temperature for 15 minutes. Following reduction, the buffer was exchanged for 0.5 mg/mL DTT in 20 mM ammonium acetate (reducing buffer) by washing the filter three times with 500 µL reducing buffer and concentrating to approximately 100 µL.

4.4.7 ESI-MS of Herceptin Heavy Chain

MS analysis of reduced Herceptin, to determine heavy chain glycoforms, was performed by Elizabeth Willows, using a method established by Dr. Adam Dowle (York Technology Facility). Positive-ion ESI mass spectra were obtained using a Bruker maXis HD™ LC-MS/MS system recording mass spectra, without LC hyphenation. Instrument control, data acquisition and processing were performed using Compass 1.7 software (microTOF control, Hystar and DataAnalysis, Bruker Daltonics). Samples were infused at 3 µL per minute dissolved in aqueous 50% acetonitrile, 1% formic acid. This solution was made by adding an equal volume of acetonitrile and the required volume of formic

acid to the concentrated Herceptin heavy chain sample in aqueous solution. Source settings were ion spray voltage: 2000 V, dry gas: 4 L/min, dry gas temperature: 250 °C. The ion acquisition *m/z* range was 900–5000, and the spectral acquisition rate was 0.1 per second. Ion transfer settings were tuned for optimisation of high mass signals, specifically: isCID energy, 90.0 eV; transfer time, 75.0 µs; collision cell RF 1500.0 Vpp; quadrupole ion energy, 18 eV; quadrupole RF, 1134.3 Vpp. The instrument was externally calibrated over the acquisition *m/z* range using Agilent Tuning Mix (G1969-85000), with an enhanced quadratic fit applied to the data.

4.4.8 MS Data Analysis

Spectra were summed before MaxEnt deconvolution to yield neutral average masses, specifying a mass range of 15,000-90,000 Da and a mass resolution of 2000. Peaks were picked using Sum Peak with a signal to noise threshold of 1 and an absolute threshold of 100. Heavy chain peaks were identified based on a distinct glycosylated heavy chain signal pattern in the spectra: an intense signal around the expected heavy chain *m/z* value, with less intense signals of higher *m/z* in increments of 162 Da, characteristic of additional hexose residues on an *N*-glycan chain. Glycoforms were assigned by subtracting the known molecular weight of Herceptin heavy chain, 49,252 Da, subtracting 128 Da for assumed loss of the terminal lysine residue and assigning the *N*-glycan structure with the closest molecular weight to the remaining unassigned *m/z* value.

4.5 Results

4.5.1 Investigating the Effects of Cog4L36P Mutation on Production of Recombinantly Expressed Herceptin

Characterisation of Cog4L36P and ΔCog4L36P cells (chapter 2) suggested that the mutation may confer favourable properties for biologic production, including reduced adhesion and increased metabolic activity. Therefore, it was of interest to test the production levels of a model biologic by Cog4L36P cells. For this purpose, the therapeutic mAb, Herceptin, was chosen as a plasmid was readily available and it has previously been well-characterized.

Prior to adaptation of CRISPR-engineered Cog4L36P CHO-K1 adherent cells to suspension culture, the ΔCog4L36P adherent CHO-K1 cell line and its ΔCog4WT counterpart were adapted to suspension culture and transfected with Herceptin by Elizabeth Willows (Ungar group). Herceptin levels were monitored by protein A ELISA, with quantitation based on comparison to a linear standard curve of human IgG protein. Initially, this experiment was performed over a seven-day period, however the cell cultures became too dense and less viable over the final few days, resulting in lower production and some Herceptin degradation. Therefore, further experiments were limited to a four-day period.

When Herceptin production was monitored in this way in ΔCog4WT and ΔCog4L36P cells, Herceptin levels were found to be consistently higher in ΔCog4L36P cultures than in ΔCog4WT (personal communication from Elizabeth Willows): a result with promising implications for the biopharmaceutical industry. It was therefore imperative to repeat this experiment using CRISPR-engineered Cog4L36P cells, to assess whether the effect was due to overexpression of Cog4 in the knock-in cells or due to the Cog4L36P mutation.

Suspension-adapted CHO-K1 Cog4L36P and WT cells, cultured in OptiCHO medium, were therefore also transfected with the Herceptin plasmid and production was monitored over a four-day period. Additionally, separate cultures of each cell line were transiently transfected with GFP, enabling flow-cytometric monitoring and estimation of transfection efficiencies. For Herceptin-transfected cells, cell number and viability

were monitored daily. By quantifying the Herceptin ELISA and correcting for viable cell number (figure 4.1a) and transfection efficiency (figure 4.1b), total daily Herceptin concentration (figure 4.1c) was also expressed as production in pg per transfected cell (figure 4.1d). Although this calculation relies on the assumption that the transfection efficiency of the GFP plasmid is representative of that of the Herceptin plasmid, which may not be completely accurate, the relative transfection efficiencies of WT and Cog4L36P cells are likely to be the same for both plasmids. This enables at least relative quantitation between the cell lines. Results showed that Cog4L36P cells produced markedly more Herceptin than did the WT, in agreement with previous findings in the knock-in cell lines. Every day following transfection, Cog4L36P cells produced around double the amount of Herceptin per cell compared to WT (figure 4.1d). Furthermore, transfection efficiency tended to be higher in Cog4L36P cells than WT (figure 4.1b), so the amount of total Herceptin the Cog4L36P cells produced daily was more than four times higher than the amount produced by WT cells (figure 4.1c). This mutant cell line could therefore be of great interest for therapeutic mAb production, where maximising antibody titre is critical.
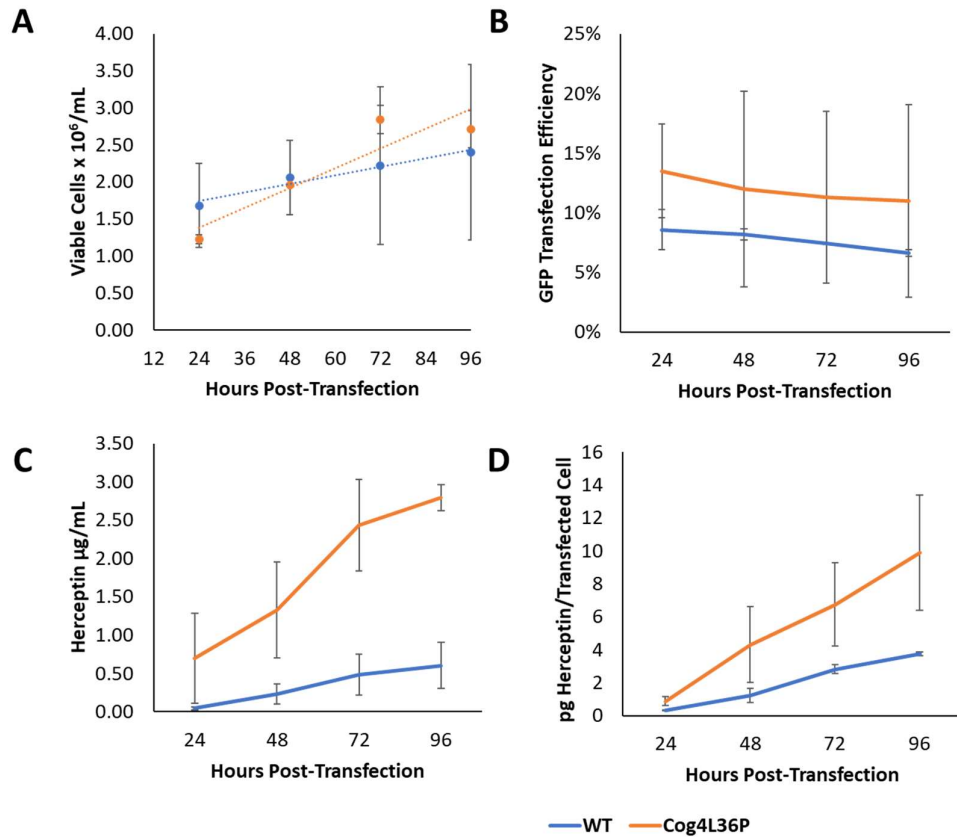
Figure 4.1 | Transient transfection graphs for suspension-adapted WT CHO-K1 and Cog4L36P CHO-K1 cells transfected with Herceptin. Cells were transfected with either Herceptin or GFP using the manufacturer's FreeStyle™ MAX transfection protocol and monitored every 24 hours for 4 days. Error bars represent +/- 1 standard deviation. Every day, a media sample was harvested by centrifugation at 1000G for 5 mins and used to measure Herceptin concentration. A protein A sandwich ELISA was performed with each sample using α-(hu)IgG-HRP antibody. Concentration was calculated by calibration to a standard curve of purified Herceptin standards. Cell numbers obtained from daily Vi-Cell measurements and flow cytometry GFP transfection efficiency values were used to calculate Herceptin concentration in pg/transfected cell. A) Daily viable cell counts following transfection, as monitored by counting using a ViCELL™ XR cell counter (Beckman Coulter). B) Daily GFP Transfection efficiency, as monitored by flow cytometric analysis of GFP transfected control cells. C) Herceptin production (µg/mL), as monitored by ELISA every day for four days following transfection. D) Herceptin production (pg) per transfected cell. Each data point represents the average of two biological replicates, with error bars showing +/- 1 standard error of the mean. Both experiments were carried out within four months of adaptation from adherent to suspension culture.

185

4.5.2 Evaluating the Effects of Different Cell Culture Media and Passage Number on Herceptin Production by WT and Cog4L36P CHO-K1 Cells

WT and Cog4L36P CHO-K1 cells were originally adapted from adherent to suspension culture in OptiCHO medium. This medium was readily available in the labs and routinely used for culture of other suspension CHO cell lines by colleagues. However, persistent cell clumping was observed, and manual disruption was required to separate cells. Clumping minimises the exposed surface area of cells so is detrimental for secretion of recombinant protein. To attempt to overcome cell clumping, different media were researched and FortiCHO was selected as the manufacturer recommended this for recombinant protein production in CHO-K1 cells. Cells were passaged from OptiCHO into FortiCHO medium and allowed to reach viability above 95% before proceeding with more experiments. Clumping was minimised, however subsequent Herceptin transient transfection sampling and analysis revealed a decrease in Herceptin production in the Cog4L36P cells and WT cells expressed considerably more Herceptin than Cog4L36P cells in the new medium (figure 4.2a).

Several possibilities were explored to understand the discrepancy in Herceptin production in the two types of media. Cells were transferred back into OptiCHO medium one week prior to transient transfection with Herceptin to establish whether the growth medium was impacting recombinant protein expression. This failed to restore Herceptin expression levels to those observed before the medium change (figure 4.2b). The possibility of a cell mix up was also ruled out by purifying DNA from both cell-lines and running the same test used to identify original Cog4L36P CRISPR clones. The cells were positively identified as the correct cell-type, based on SacI sensitivity of the WT Cog4 sequence and resistance to SacI digestion in the Cog4L36P cells (figure 4.2c). This led to the hypothesis that cell aging over the course of six months between the first experiments in OptiCHO and later experiments in FortiCHO caused the reduction in recombinant protein expression. Younger cell pellets, which were freshly adapted from adherent to suspension culture, were thawed into FortiCHO medium and the transfection experiment was repeated. This time, the result matched the original observation in the younger cells in OptiCHO medium: the Cog4L36P cells express considerably more Herceptin than do the WT (figure 4.2d).
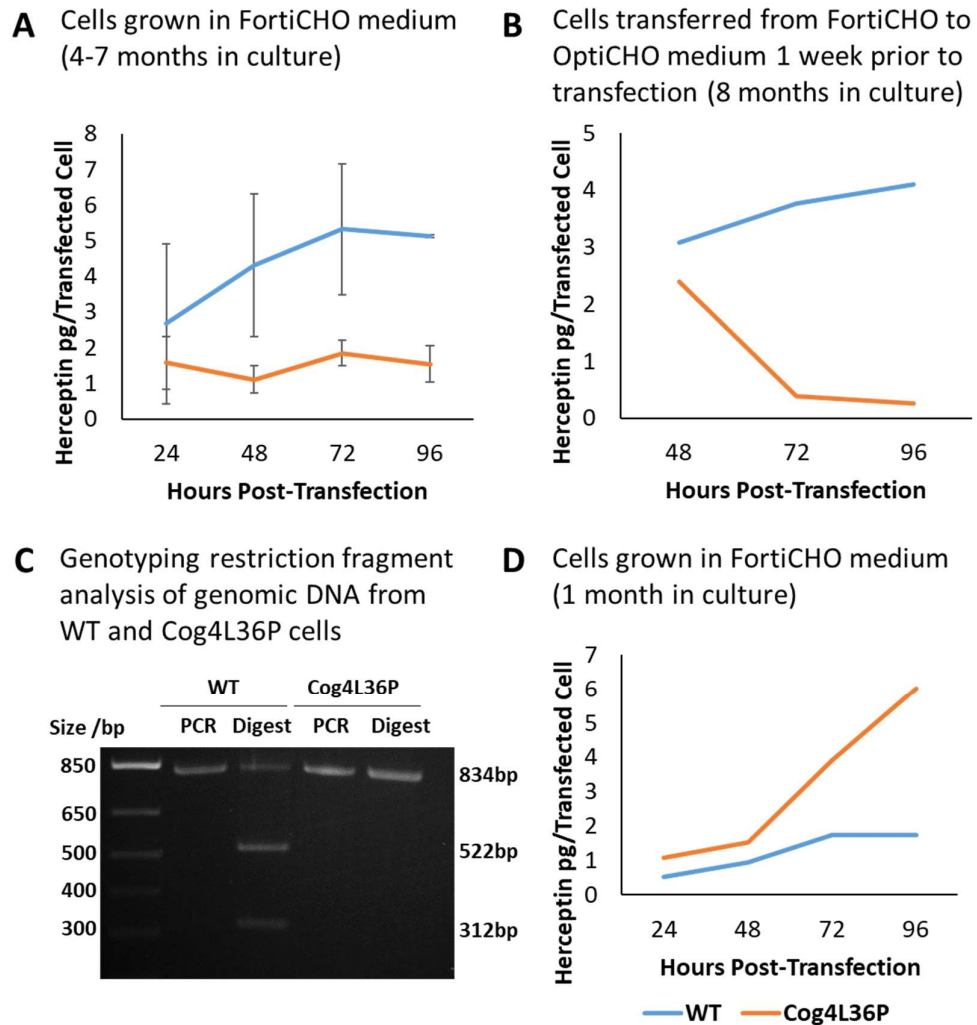
Figure 4.2 | Investigating Herceptin production discrepancies in WT and Cog4L36P CHO-K1 cells. Cells were transfected with either Herceptin or GFP and monitored every 24 hours for 4 days. Error bars represent +/- 1 standard deviation. Every day, a media sample was harvested by centrifugation at 1000G for 5 mins and used to measure Herceptin concentration. A protein A sandwich ELISA was performed with each sample using α-(hu)IgG-HRP antibody. Concentration was calculated by calibration to a standard curve of purified Herceptin standards. Cell numbers obtained from daily Vi-Cell measurements and flow cytometry GFP transfection efficiency values were used to calculate Herceptin concentration in pg/transfected cell. A) Cells grown in FortiCHO medium (4 months-7 months post-suspension adaptation) N=3. B) Cells transferred from FortiCHO to OptiCHO medium 1 week prior to transfection (age 8 months) N=1 C) Genotyping restriction fragment analysis of genomic DNA from WT and L36P cells. D) Cells grown in FortiCHO medium (age 1 month) N=1.

4.5.3 Testing the Effects of Cog4L36P Mutation on Glycosylation of Recombinantly Expressed Herceptin

Whole-cell glycan profiling of Cog4L36P suspension-adapted cells revealed glycosylation distinct from either WT or ΔCog4 cells (chapter 3). Considering this finding, glycosylation of Herceptin recombinantly expressed by WT and Cog4L36P cells was studied using ESI of intact protein. Production and glycosylation of Herceptin produced by ΔCog4WT and ΔCog4L36P cells was also examined by Elizabeth Willows (Ungar group) using the same approach. After transiently transfecting cells and monitoring Herceptin production for four days, Herceptin IgG was purified using a protein L chromatography column on an FPLC system. Purified Herceptin from each cell line was then concentrated, reduced, and analysed using a maXis UHR-TOF mass spectrometer. However, it became apparent that 30 mL cultures of WT cells produced an insufficient quantity of Herceptin for intact MS analysis as spectra did not show any signals for Herceptin produced by WT cells. This was likely because these cells already produced less Herceptin, and protein was lost due to non-specific binding to the protein concentrator membranes used following purification, resulting in the amount being below the limit of detection for the mass spectrometer. Sufficient Herceptin was however, obtained from ΔCog4WT, ΔCog4L36P and CRISPR-engineered Cog4L36P cells.

Comparison of heavy chain (HC) spectra for Herceptin produced by transiently transfected ΔCog4WT and ΔCog4L36P cells revealed very similar data, suggesting the same glycoforms are present in each (figure 4.3). Representative deconvoluted spectra for Herceptin preparations taken from a single transient transfection experiment are shown in figure 4.3 and additional spectra were recorded for two other Herceptin preparations from replicate transient transfections in each cell line. Overall, the results confirmed the same glycoforms of Herceptin were produced by each cell line in the same relative ratios. For each cell line, three major heavy chain glycoforms were observed in the deconvoluted spectra at approximately 50562, 50724 and 50885 Da, each differing by the approximate mass of a single hexose monosaccharide (+ 162 Da). Based on a reduced heavy chain mass of 49252 Da and loss of a C-terminal lysine residue (-128 Da), the difference between the heavy chain protein mass (49124 Da) and the observed deconvoluted mass for the most intense peak was 1438, 1600 for the next peak, and 1762 for the next. For each of these mass increments, a potential glycan

structure was proposed, FucGlcNAc$_2$Man$_3$GlcNAc$_2$ (G0F) (1,444 Da), FucGlcNAc$_2$Man$_3$GlcNAc$_2$Gal (G1F) (1606 Da) and FucGlcNAc$_2$Man$_3$GlcNAc$_2$Ga$_{l2}$ (G2F) (1768 Da): each is within 6 Da of the calculated masses, which is within the measurement error of the used instrument. Thus, despite differences in the whole cell glycan profiles of each cell line (chapter 3), no difference was observed between glycan profiles of the Herceptin heavy chains produced in each of the cell lines (figure 4.3).
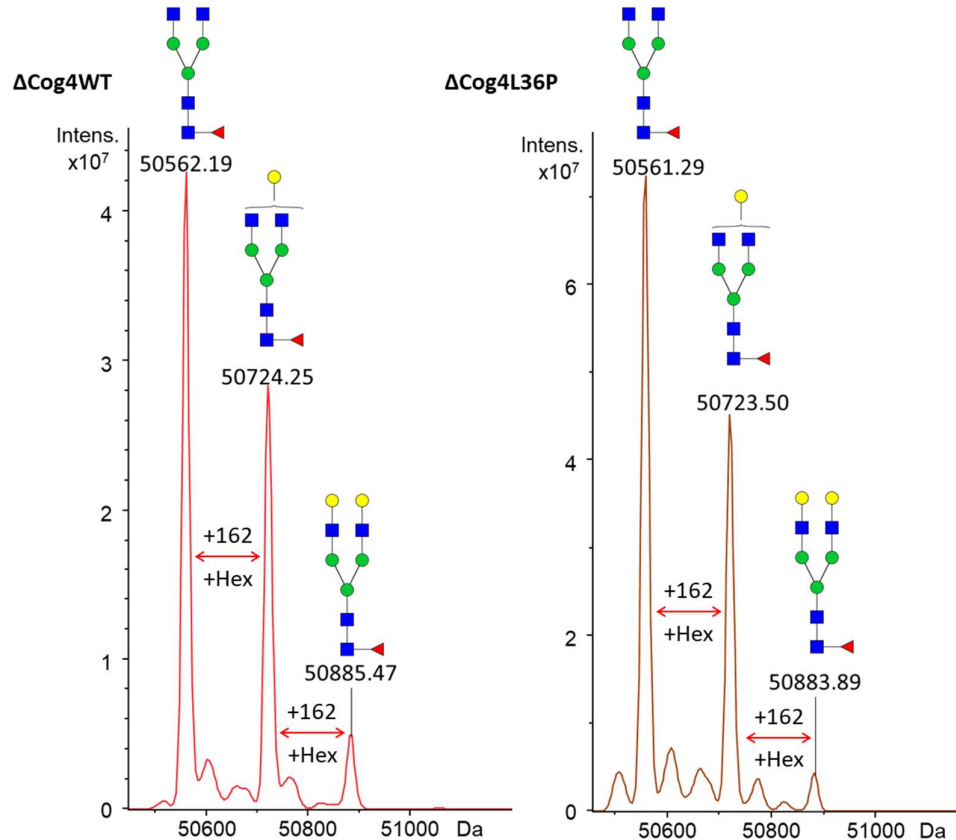
Figure 4.3 | Mass spectrometric analysis of glycosylated heavy chain of Herceptin recombinantly expressed by transiently transfected ΔCog4WT and ΔCog4L36P. Deconvoluted mass spectra of Herceptin heavy chain expressed by ΔCog4WT and ΔCog4L36P CHO-K1 cells. Cells were transiently transfected with Herceptin for 4 days prior to harvesting using a protein L column on a UPLC protein purification system. Purified Herceptin was concentrated and reduced using a VivaSpin 4 centrifugal concentrator device. A sample of the resulting preparation was analysed using a maXis UHR-TOF mass spectrometer. Proposed glycan structures were assigned based on a Herceptin heavy chain mass of 49252 Da with loss of its terminal lysine residue (-128). Glycan masses are within 6 Da of the observed m/z values.

It would be beneficial to analyse HC glycosylation of Herceptin recombinantly expressed by WT and CRISPR-engineered Cog4L36P cells too as Cog4 is expressed at different levels in the knock-in (ΔCog4WT and ΔCog4L36P) cells. Unfortunately, insufficient quantities of Herceptin produced by WT cells prevented comparison using the analytical methods applied here. However, comparable deconvoluted spectra were obtained for Herceptin produced by ΔCog4L36P and CRISPR-engineered Cog4L36P cells (figure 4.4). These Herceptin samples were concentrated and analysed in parallel, enabling direct comparison between the MS results. Although the deconvoluted spectra aren't as well resolved as those presented in figure 4.3, very similar signals are present for both cell lines, suggesting the same glycoforms are present in each (figure 4.4). Therefore, the tentative assumption could be made that Cog4 levels do not affect glycosylation of recombinantly expressed Herceptin and WT and ΔCog4WT cells may produce similar Herceptin glycoforms too.
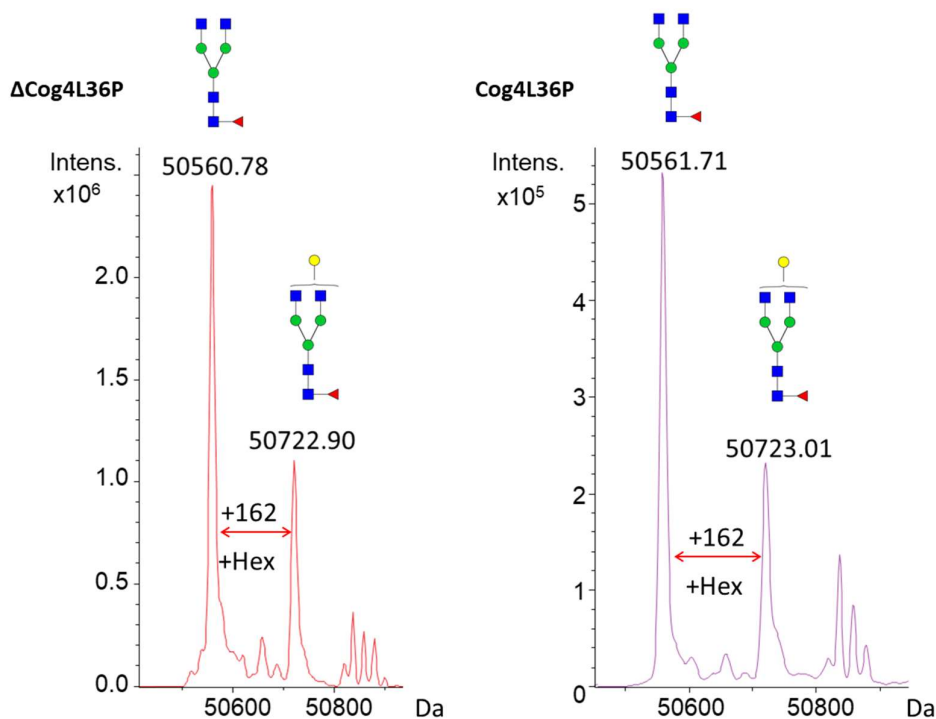
Figure 4.4 | Deconvoluted mass spectra of glycosylated heavy chain of Herceptin recombinantly expressed by ΔCog4L36P (left panel) and CRISPR-engineered Cog4L36P cells (right panel). Proposed glycan structures were assigned based on a Herceptin heavy chain mass of 49252 Da with loss of the terminal lysine residue (-128 Da). Glycan masses are within 8 Da of the observed molecular weights.

4.5.4 Boosting Transfection Efficiency to Improve Herceptin Titre

Transient transfection is a valuable method for high throughput screening of host cells to identify desirable properties. However, production levels are heavily influenced by transfection efficiency. For all transfections performed using FreeStyle™ MAX reagent following the manufacturer's protocol, the transfection efficiencies were below 20%, dropping as low as 4% according to GFP measurements. Increasing the transfection efficiency would be the simplest way to increase Herceptin titre, to aid upstream analysis of mAb properties such as glycosylation. During a placement with the Biopharm cell line development team at GSK, an alternative protocol for FreeStyle™ MAX transfection was utilised. The main difference between the manufacturer's protocol and the alternative protocol is that the cell culture medium is replaced with OptiPRO SFM prior to transfection, and cells remain in this medium overnight, whereas the manufacturer's protocol involved transfecting cells in FreeStyle™ medium.

This alternative transfection protocol was used to transfect WT and Cog4L36P CHO-K1 cells growing in FortiCHO medium, to establish whether there was any improvement in transfection efficiency following the manufacturers protocol. Cells were cultured for a similar length of time as those used for the earlier transfection presented in figure 4.2d. The alternative method increased the transfection efficiencies from >15% to ~25% for WT cells and from >10% to ~17% for Cog4L36P cells (figure 4.5a). The resulting Herceptin production levels were considerably higher for the alternative method than for the manufacturer's method (figure 4.5b). Interestingly, the calculated Herceptin production levels in pg/transfected cell were also highest for the alternative method (4.5c), suggesting that each transfected cell can produce more Herceptin following transfection using the alternative method. This transfection also validated earlier findings that Cog4L36P cells produce more Herceptin than do WT cells. The alternative transfection method was therefore adopted for all subsequent transfections.
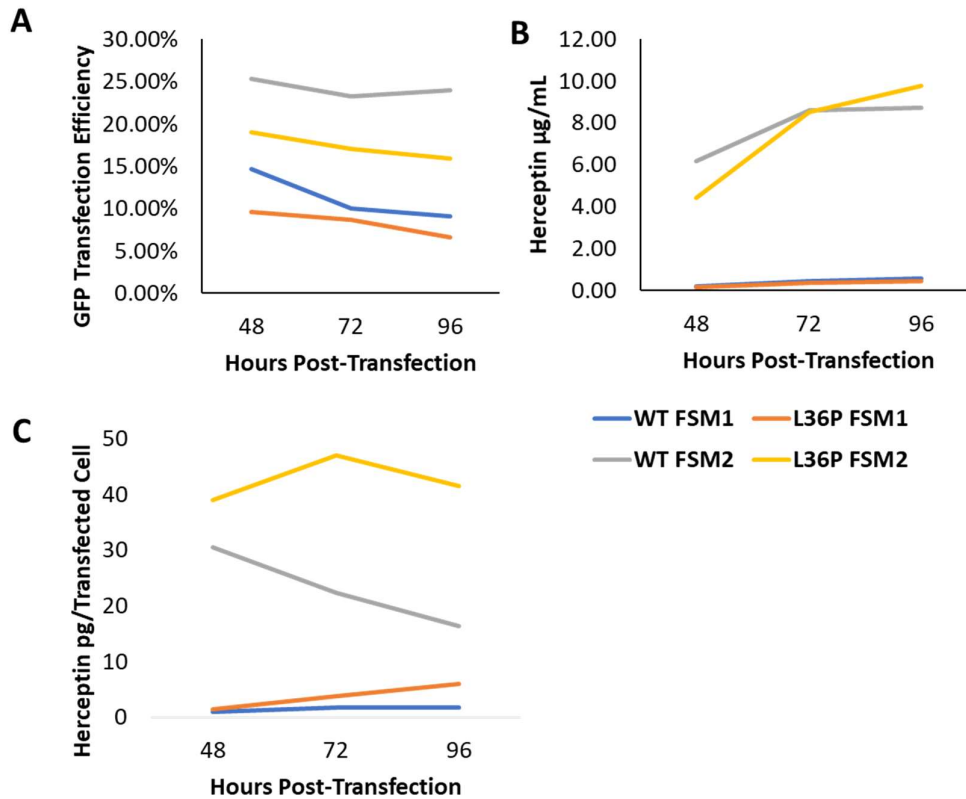
Figure 4.5 | Comparison of Herceptin production and GFP transfection efficiency in WT and Cog4L36P cells transfected with FreeStyle™ MAX following either the manufacturer's protocol (FSM1) or the alternative protocol (FSM2). A) GFP transfection efficiency 2-4 days post-transfection. B) Herceptin production in µg/mL 2-4 days post-transfection. C) Herceptin production in pg/Transfected cell 2-4 days post-transfection. N = 1 for every experiment.

4.5.5 Engineering the Cog4L36P Mutation in an Industrial Host CHO Suspension Cell Line

Enhanced production of recombinantly expressed protein would be hugely beneficial for the biologics industry, where maximizing titre is key to cutting production costs. However, the suspension-adapted cell lines developed at the University of York were adapted over a much shorter timescale than industrial host suspension cells, retaining some undesirable adherent properties such as cell clumping. Furthermore, industrial host cells undergo stringent clonal isolation, expansion, and screening for high-level producers. This is particularly important for stably transfected cells, which can have biologic encoding plasmids integrated anywhere in their genome, and integrated site

can have a huge impact on transcription, translation and therefore production levels. For transient transfection, it's also valuable to select for clones with generally favourable properties, such as ability to grow at higher density. Since only one confirmed Cog4L36P clone was identified out of all the clones screened during CRISPR-cas9n clonal isolation, there were no other Cog4L36P clones available for comparison of favourable properties. Therefore, it was of benefit to engineer the mutation in an industrial host cell line, to establish the effect of Cog4L36P mutation on production of biologics by an already high-producing cell line.

The industrial host cell line used in this investigation was a proprietary legacy CHO suspension cell line provided by GlaxoSmithKline for use in the Department of Biology at the University of York. PCR and sequencing of the region of Cog4 containing the leucine 36 residue carried out in another proprietary (current) production CHO cell line at GSK confirmed that the region of interest shared the same genomic DNA sequence as in the CHO-K1 WT cell line (data not shown). As the legacy host cells came from the same parent cells as the current production cell line, it was assumed that the Cog4 sequence would also be the same for all cell lines. This meant the same protocol and reagents as developed for CHO-K1 WT cells (chapter 2) could be used for the legacy host CHO cells too.

To maximise chances of a successful hit, two different transfection protocols were followed. These were the XFect (Clontech) transfection protocol, as used for transfecting adherent WT CHO-K1 cells previously, with minor modifications for use on suspension cells, and the alternative FreeStyle™ MAX transfection protocol. As previously described (chapter 2), cells were placed under puromycin selection for 72 hours before clonal isolation. Cells from triplicate wells for each transfection reagent treatment were pooled and counted, revealing sufficiently high numbers of viable cells (> 1 million cells for each reagent) for single cell cloning by flow cytometry. Previously, when WT CHO-K1 adherent cells were transfected, the number of surviving cells was too low to consider flow cytometric cloning, so clones were isolated by serial dilution. Flow cytometric cloning is preferable for ensuring only single cells are exported into each well as doublets are effectively excluded, whereas serial dilution relies on manual identification of single cell clones.

4.5.6 Identifying the Cog4L36P Mutation in an Industrial Host CHO Suspension Cell Line

A total of 78 clones were identified from 1440 96-well plate wells after two weeks of growth. Initially, these were screened using restriction fragment analysis (figure 4.6), as described in chapter 2. Initial screening showed that some clones, including F2B8, F2E9, F7B8, F7C9, F9E4 and F10D8 gave PCR fragments that were either partially or completely digested by SacI, suggesting they retained the WT Cog4 gene sequence with the restriction site (figure 4.6). Several clones, including F3F7, F7E5, F8C5, F9E3, F10F5 and X5E8, appeared to resist SacI digestion, as the DNA product did not change size after restriction digest (figure 4.6). These were therefore identified as potential Cog4L36P clones and sent for sequencing, as insensitivity to SacI digestion suggested that CRISPR edited DNA may have been repaired by incorporation of the ssODN repair template by HDR, which contains a silent mutation abolishing the SacI restriction site. However, sequencing revealed the region of DNA containing the leucine 36 site and the SacI restriction site was missing from most of clones (figure 4.7), except for F8C5, which contained an insertion, and F10F5, which gave inconclusive results. Their DNA had most likely been repaired by non-homologous end joining, creating a small knock-out that removed the restriction site, yet was indistinguishable from the WT digestion pattern by DNA gel.
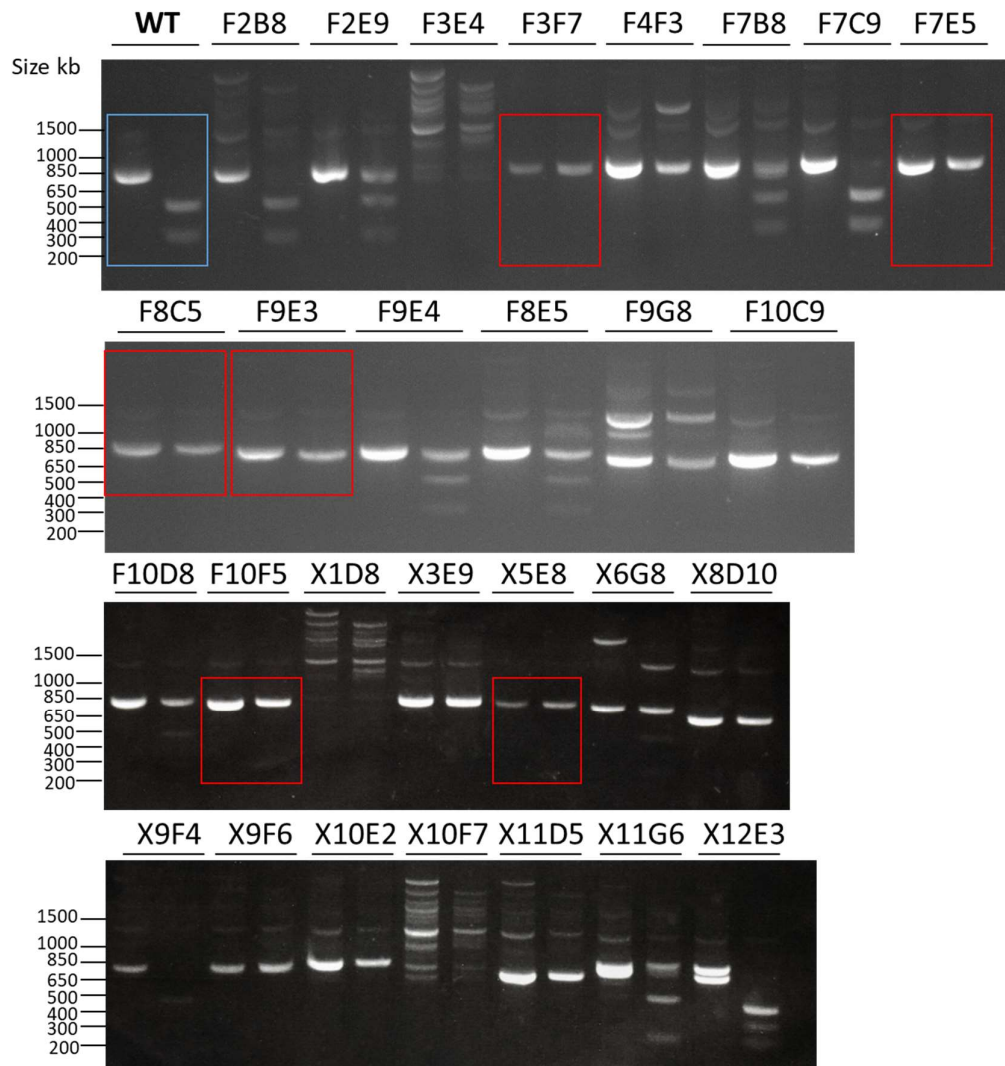
Figure 4.6 | Restriction fragment analysis of CRISPR clones. Cell lines grown to a density of approximately 1 x 10^6/mL in 24-well plates were screened for CRISPR HDR editing using restriction fragment analysis. An 834bp strand of genomic DNA, centred asymmetrically on a single SacI restriction site was amplified and PCR purified. DNA was digested overnight with SacI and analysed on a 1.5% agarose gel. Suspected positive clones are boxed in red and the WT control is highlighted in blue. As an additional control for the PCR reaction, ddH2O was also used as a template to rule out contamination. For each cell line, an undigested PCR product is run alongside a SacI digested sample of the PCR product. Clones were named based on the type of transfection reagent used (X for XFect, F for FreeStyle™ MAX), followed by the plate number and well position.

Figure 4.7 | Sequence alignment for the WT Cog4 genomic sequence and suspected Cog4L36P positive clones. The WT Cog4 genomic sequence is displayed at the top, with sequences for several clones identified as potentially Cog4L36P by restriction fragment analysis (figure 4.6. DNA mismatches are highlighted in yellow, and gaps are indicated by dashes. The Leucine 36 codon is highlighted in blue. Sequencing results for F10F5 were inconclusive as the trace ended before the leucine 36 codon. Sequences were aligned using SnapGene software.

Restriction fragment analysis revealed that non-homologous end joining occurred frequently in industrial host CHO cells edited by CRISPR-cas9n, resulting in an abundance of false positive DNA gel results. Therefore, for subsequent screening, PCR products obtained from genomic DNA isolated from each clone were not subject to restriction digest, and instead sent for sequencing following purification of DNA from the PCR reaction mixture. Sequencing results showed that some clones, including X9F9, had WT Cog4 DNA and some, including F9F8, F2B8 and F6F5 had DNA deletions (figure 4.8). F9F8 and F6F5 were missing large sections of DNA, whereas F2B8 was only missing a few bases. Interestingly, F9F8 also contained a CTG->CTT silent mutation downstream of the leucine 36 codon, suggesting the ssODN repair template had been partially incorporated as this silent mutation was designed into the repair template to abolish a potential PAM recognition sequence and prevent unwanted cas9n binding after editing. One clone, X2F5, was positive for the Cog4L36P mutation (figure 4.8). X2F5 contained four mutations in total, three of which were silent mutations upstream of the leucine 36 codon. The repair template contained five mutations in total, indicating that in the X2F5 clone, HDR editing had occurred by a crossover that occurred upstream of the fifth mutation in the template. For subsequent experiments, clone X2F5 Cog4L36P was compared to its WT parent.

196
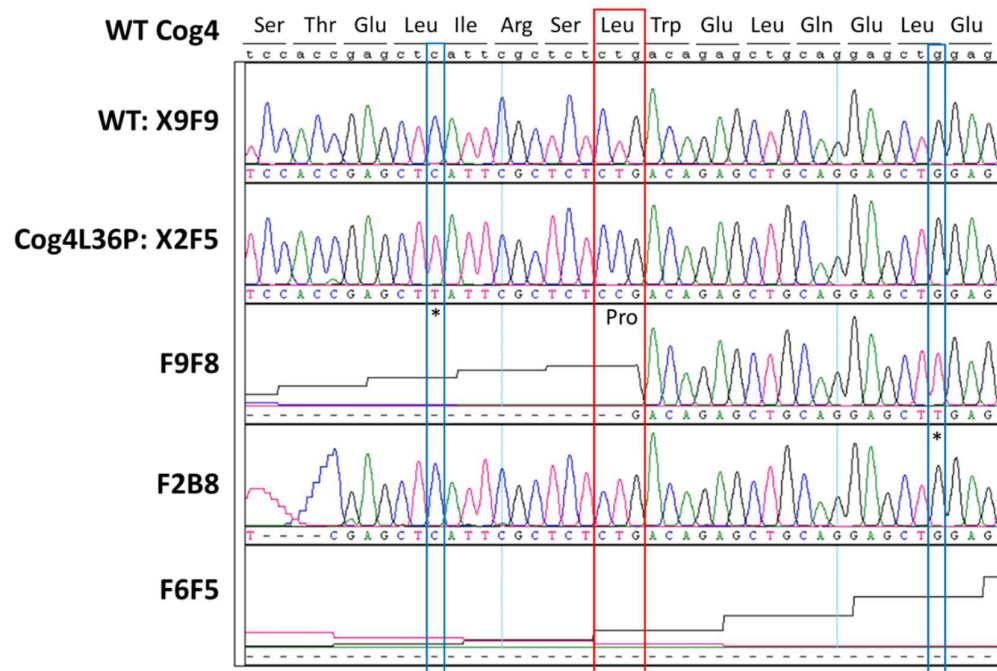
Figure 4.8 | Sequence trace alignment for select CRISPR-edited clones. The WT Cog4 genomic sequence is displayed at the top in lower case, with corresponding amino acids labelled above each codon. Sequence traces for several CRISPR-edited clones are aligned below, with the name of each clone from which the DNA was isolated written to the left of each trace. Blue boxes highlight sites targeted by silent mutations engineered into the ssODN repair template. Asterisks indicate sequences containing these silent mutations. The leucine 36 codon is highlighted with a red box, with the proline codon in the Cog4L36P mutation labelled underneath the Cog4L36P sequence. Sequences were aligned using SeqMan Pro software.

4.5.7 Testing the Effect of Cog4L36P Mutation on Herceptin Production by Industrial Host CHO Cells

Both WT and Cog4L36P industrial host cells were transiently transfected with a Herceptin plasmid and production was monitored by ELISA, over a period of four days, replicating the experiments performed in the suspension-adapted CHO-K1 cells. Following transfection, cells were counted every day and the transfection efficiency was estimated based on flow cytometric analysis of separate GFP-transfected control cultures of each cell line. These results were combined with results of ELISA analysis of daily samples taken from each Herceptin-transfected culture to calculate daily Herceptin production in pg/transfected cell. Measurements of transfection efficiency

and amount of Herceptin produced per transfected cell taken 24 hours after transfection were excluded from the analysis, as the calculated Herceptin production was abnormally high (higher than most literature reported values) and there was massive variation between replicates. This suggested that at 24 hours post-transfection, the GFP transfection efficiency was not representative of the Herceptin transfection efficiency in these cell lines using the alternative FreeStyle™ MAX protocol. The results show that over the course of the transfection, cells displayed a linear growth curve and Cog4L36P cells proliferated more slowly than did the WT cells (figure 4.9a). Slower proliferation of Cog4L36P cells compared to WT matches previous observations in WT and Cog4L36P suspension adapted CHO-K1 cells. The amount of Herceptin produced everyday overall was higher in WT host cells than in Cog4L36P cells (figure 4.9c). However, when production was adjusted for the different cell counts and transfection efficiency, which was similar in both cell lines (figure 4.9b), Herceptin production was highest in WT cells at 48 hours after transfection, but there was no difference between the cell lines at 72 or 96 hours.
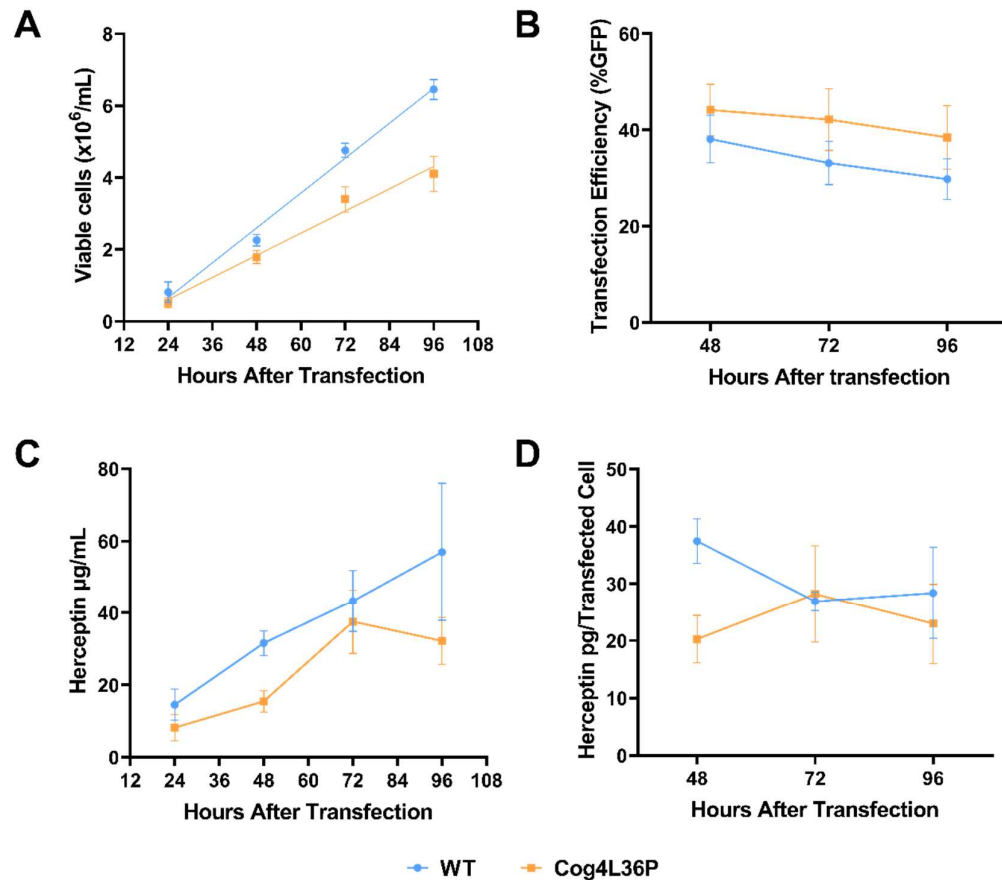
*Figure 4.9 | Transient transfection graphs for industrial host WT CHO and Cog4L36P CHO cells transfected with Herceptin. Cells were transfected using the GSK FreeStyle™ MAX transfection protocol. A) Daily viable cell counts following transfection, as monitored by counting using a ViCELL™ XR cell counter (Beckman Coulter). B) Transfection efficiency between 2-4 days after transfection, as monitored by flow cytometric analysis of GFP transfected control cells. C) Herceptin production (µg/mL), as monitored by ELISA every day for four days following transfection. D) Herceptin production (pg) per transfected cell, calculated based on measured Herceptin concentration (C), viable cell count (A) and transfection efficiency (D). Error bars represent +/- 1 standard error of the mean (n=3). Statistical significance for each dataset was assessed by two-way ANOVA using GraphPad Prism software. There was a significant effect of genotype on growth rate (A) (P < 0.0001****). There was no significant effect of genotype on transfection efficiency (B) or Herceptin production (C and D).*

4.6 Conclusions and Future Work

Maximising production of biologics whilst minimising their glycan heterogeneity is crucial for the biopharmaceutical industry. For mAbs, as well as other glycoprotein biologics, controlling glycosylation is critical as glycans can influence mAb properties (Beck, 2013, Lauber et al., 2015, Mimura et al., 2018). With this in mind, Cog4L36P cells, which were shown to have enhanced metabolic activity and an adhesion defect (chapter 2) as well as altered whole-cell glycosylation (chapter 3), were tested as host cells for the production of the mAb, Herceptin. This chapter presents the results of testing production levels and glycosylation of Herceptin in WT and Cog4L36P cells. Production of Herceptin was achieved by transient transfection, as a high throughput alternative to stable integration of the Herceptin plasmid.

Prior to testing Herceptin production in CRISPR-engineered cells, ΔCog4WT and ΔCog4L36P cells were transiently transfected, leading to the observation that the mutant cells produced more Herceptin than did the WT. This was a highly promising result, and it was imperative to test whether the CRISPR-engineered Cog4L36P mutation conferred the same enhancement. Initially, results showed an increase in Herceptin production in Cog4L36P cells compared to WT, suggesting that the enhancement was due to the mutation rather than overexpression of Cog4. This effect appears to be dependent on the age of the cell culture, however genetic plasticity is a known property of CHO cells so it is common practice to keep passage numbers low when working with clonal CHO cell lines (Chusainow et al., 2009).

Genetic plasticity is an inherent property of CHO cells, which is readily exploited to select for specific traits in clonal variants of CHO cells that have differences in their genetic/epigenetic makeup. In freshly suspension-adapted CHO cells, cells are under increased pressure to adapt to their new environment, so it was somewhat unsurprising that expression of Herceptin changed as suspension-adapted Cog4L36P cells aged. It was therefore important to test the mutation in a true production line, which had previously been selected for high titre biologic production. This production line is an established suspension cell line, so unlike freshly adapted suspension cells, adaptation to growth in suspension culture isn't a driver of accelerating random mutations.

The same CRISPR-cas9n engineering strategy as used for adherent CHO-K1 cells was also employed to engineer the Cog4L36P mutation in legacy industrial host CHO cells, as the cell lines shared the same genomic Cog4 sequence. This not only saved time and minimised costs, but also meant comparison between the CHO-K1 cells and the industrial host CHO cells after CRISPR-cas9n editing was more valid as cells were subject to similar treatment and had incorporated the same silent mutations. Previously, a success rate of 1/11 was achieved for engineering the Cog4L36P mutation in CHO-K1 cells and restriction fragment analysis was a reliable screening approach. However, the industrial host CHO cells appeared to be more prone to repair of cas9n induced DNA breaks by non-homologous end joining, resulting in a higher frequency of DNA deletion and false positive restriction digest patterns. Overall, one Cog4L36P positive cell line was identified out of over 70 clones screened.

When WT and Cog4L36P industrial host CHO cells were transiently transfected with Herceptin and subject to the same ELISA analysis as the CHO-K1 cells, the host cells produced a considerably higher titre of Herceptin. For CHO-K1 cells, even after using a more efficient transfection protocol, the amount of Herceptin produced overall by WT and Cog4L36P cells was in the 5-10 µg/mL range, whereas industrial host CHO cells achieved titres up to 60 µg/mL. This was likely due to slower proliferation and lower tolerance to high cell density in the suspension-adapted CHO-K1 cells compared to the industrial host cells. Industrial host cells are selected partly on their ability to tolerate high cell density and remain viable, whereas the suspension-adapted cells developed by the author were not. However, both CHO-K1 and industrial host CHO cells achieved similar Herceptin production per transfected cell following a more efficient transfection method. Although young cultures of Cog4L36P CHO-K1 suspension-adapted cells produced more Herceptin per cell than their WT parent cells, there was no difference in production levels per cell between the WT and Cog4L36P industrial host cells. Cog4L36P cells proliferated more slowly than did the WT cells, so overall the production of Herceptin was lower, however during industrial scale production of mAbs, proliferation rate is less important as cells are grown to a certain density and production is maintained in a fed batch.

In addition to monitoring production levels of Herceptin in WT and Cog4L36P CHO-K1 cells, the work presented in this chapter also began to assess glycosylation of recombinantly expressed Herceptin. Although Cog4L36P cells had a distinct whole-cell glycan profile compared to WT cells, analysis of intact glycosylated Herceptin heavy chain revealed no detectable changes to the type or relative abundance of glycoforms produced by the mutant cells. However, intact glycoprotein analysis is typically only used in industrial production of mAbs as a first screen, as detection and quantitation of low abundance glycoforms is inaccurate (Reusch et al., 2015). Furthermore, conclusions from the HC glycosylation analysis presented in this chapter can only tentatively be drawn as the spectra were gathered for Herceptin expressed by ΔCog4WT and ΔCog4L36P cells, which express Cog4 at different levels compared to WT and CRISPR-engineered Cog4L36P cells. To ascertain whether the Cog4L36P mutation can influence glycosylation of recombinantly expressed biologics, considerably more investigation is needed, which was unfortunately not possible within the restricted timescale of this PhD.

In future, *N*-glycans from purified Herceptin produced by WT and Cog4L36P cells should be cleaved prior to MS analysis, as this allows accurate quantitation of the relative abundance of different glycan structures and enables more sensitive detection of low abundance glycans (Sinha et al., 2008). Utilising Waters' RapiFluor-MS *N*-glycan labelling kit, which involves PNGase-F *N*-glycan release and *N*-glycan reducing terminal fluorescent tagging, for analysing Herceptin glycosylation would be one approach. This method is routinely used by GSK and other pharmaceutical companies for mAb glycan profiling, where the repertoire of expected *N*-glycan structures is known. Alternatively, PNGase-F released glycans from recombinantly expressed Herceptin could be permethylated and analysed by MALDI-MS, which would be an advantageous approach if mutant cells produce unexpected glycoforms.

In future, perhaps aided by the more efficient FreeStyle™ MAX transfection method utilised for some work presented in this chapter, higher quantities of biologics can be produced by the cells. This would enable the same HC glycosylation ESI-MS analysis to be performed in the WT and CRISPR-engineered Cog4L36P CHO-K1 cells as for the ΔCog4WT and ΔCog4L36P cells (figure 4.3), or for glycans to be cleaved prior to MS

analysis. Glycan profiles should also be obtained for Herceptin produced by WT and Cog4L36P industrial host CHO cells. Since these cells produce considerably more Herceptin than the suspension-adapted CHO-K1 cells, it is anticipated that it will be easier to obtain glycan profiles for the Herceptin they produce, and the results will be more relevant to industry.

The results presented in this chapter concerned a single mAb, Herceptin. The repertoire of possible *N*-glycan structures on the mAb Fc region is sterically limited (section 1.3.3.2), so it may be the case that Herceptin glycosylation is unaffected by changes to glycosylation machinery in Cog4L36P cells for this reason. It would therefore be beneficial to compare production and glycosylation of alternative glycoprotein biologics, with more complex glycosylation patterns, in Cog4L36P and WT cells. For this reason, transient transfections were also performed with a plasmid encoding the glycoprotein cytokine drug, human BMP2. This plasmid, which contains a poly-His tag for detection, has previously been used by colleagues in studies of adherent cells with mixed success and it was known to be difficult to express. Unfortunately, levels of BMP2 expressed by WT and Cog4L36P cells, of both the suspension-adapted CHO-K1 lineage and the industrial host cells, were undetectable by western blotting and below the detection limit for ELISA. In future, more optimisation of expression of this plasmid is needed, or an alternative glycoprotein biologic could be tested. Alternative mAb expression plasmids could also be tested in each cell line, to assess whether any of the results presented in this chapter are common all mAbs or specific to Herceptin. Ideally, a mAb plasmid conjugated to a fluorescent tag should be tested, enabling direct quantitation of transfection efficiency as opposed to tentative estimation based on a separate GFP transfection, which may not be accurate.

Another future experiment that could be performed using WT and Cog4L36P cells would be to stably integrate a biologic plasmid and assess production levels over time. For each cell line, multiple clones should be isolated and tested, to account for differences attributed to clonal variation. This experiment would be an industrial standard approach to ascertaining whether the Cog4L36P mutation can be beneficial for biologic production in general. Testing multiple clones of each genotype would shed light on the extent to which clonal variation contributes to the observed differences.

In conclusion, the results presented in this chapter began to highlight the potential of the Cog4L36P mutation for enhancing recombinant expression of Herceptin. However, minimising the effects of genetic plasticity, which are accelerated in the process of suspension-adaptation, is key to obtaining reproducible expression levels. In CHO-K1 cells at low passage numbers, transient transfection followed by ESI-MS analysis of HC glycoforms suggests the Cog4L36P cells produce more Herceptin per cell than their WT parent, with unaltered glycosylation. Although the same effect could not be replicated for Herceptin production in transiently transfected industrial host cells, more robust conclusions could be drawn by more thorough investigation, involving stable plasmid integration, selection of more clones and ideally multiple biologics, followed by cleaving *N*-glycans prior to MS analysis.

Discussion

Protein glycosylation is a highly prevalent, yet often overlooked, covalent modification present on the majority of proteins. During the process of glycoprotein synthesis, the addition and modification of *N*-glycans at Asn-X-Ser/Thr consensus sequences is intricately linked to protein folding (section 1.1.5), and the structure, function and stability of glycoproteins can be heavily influenced by the types of glycans present (section 1.1.7). Correct glycosylation is therefore important for maintaining normal physiology and changes are associated with a range of diseases. Many signalling molecules with critical roles in immunology and development, including antibodies, cytokines, and hormones, are glycoproteins that don't function correctly without their attached glycans. There is rapidly growing interest in these molecules by the pharmaceutical industry for their therapeutic use in the treatment of cancers and autoimmune diseases (section 1.3): an industry worth hundreds of billions of dollars. Glycosylation of biotherapeutics is considered a critical quality attribute and it is thus of great interest and importance to understand the glycan biosynthetic pathway to achieve control over biotherapeutic properties and to minimise batch to batch variation resulting from heterogenous glycosylation.

The conserved oligomeric Golgi complex plays an essential role in controlling glycosylation (Shestakova et al., 2006). Non-templated glycan biosynthesis is determined by the order in which glycoproteins encounter glycosylation enzymes inside the Golgi apparatus. Maintaining enzyme levels and distribution within the Golgi is dependent on tightly regulated vesicle tethering interactions, coordinated by the COG complex, which ensure Golgi resident proteins are delivered to their required cisterna. The importance of COG for correct glycosylation is highlighted in patients with congenital disorders of glycosylation, caused by mutations in COG subunits, which result in serious consequences for glycan processing, manifesting as a range of pleiotropic symptoms (Freeze et al., 2014). In each CDG case, different COG mutations alter glycosylation in different ways, highlighting the different roles of each subunit and their importance for glycosylation.

The research presented in this thesis aimed to explore the potential of targeting the COG complex for aiding the production of biologics. It was hypothesised, based on evidence that specific mutations in COG alter glycosylation in specific ways, that COG mutation in host cells for recombinant expression of biologics could offer a new approach to engineering biologic glycosylation. Testing this hypothesis involved engineering a specific Cog subunit mutation into CHO-K1 cells, a popular expression system for mAb biologics, using CRISPR-cas9. These cells were characterised *in vivo* to understand the effects of the mutation on cell properties and to assess their suitability as host cells for recombinant protein expression. Whole-cell glycan profiling was also carried out to elucidate the impact of the mutation on glycosylation. Finally, the effects of the mutation on the production and glycosylation of the mAb, Herceptin, were investigated.

5.1 An L36P Mutation in Cog4 Impairs its Interaction with Rab30

The Cog4L36P mutation was originally identified based on an impaired interaction with Rab30, in a study aiming to elucidate the importance of specific vesicle tethering interactions for glycosylation and membrane trafficking (personal communication from Prof. Dani Ungar). The impaired interaction was identified in a yeast-two-hybrid assay; however, the assay involves binding between Rab30 and Cog4 alone, without the rest of the COG complex. Since Cog4 in its native environment is always associated with the other lobe A COG subunits, it was important to test the binding between Rab30 and Cog4 in the context of the whole complex. For this purpose, a Rab30-GST pull-down assay was performed, to test the binding of COG complex-associated Cog4, expressed in CHO-K1 cells, to GST-tagged Rab30. Results of this assay, presented in chapter 2, confirm that the mutation impairs Rab30 binding.

The structural basis for this weakened interaction is unknown, as only the structure of the C-terminal region of Cog4 has been solved so far (Richardson et al., 2009). However, it has been well-established that the N-termini of COG subunits form coiled coil (CC) domains, which likely act as Rab binding sites, and there is a predicted CC domain in the first 100 residues of Cog4 (Whyte and Munro, 2001, Fukuda et al., 2008). The interaction between Cog4 and Rab30 was first identified in a large-scale screen to identify novel Rab effectors, in which the authors also demonstrated that the N-

terminus of Cog4 (residues 1-186) was sufficient to bind Rab30 (Fukuda et al., 2008). The Cog4L36P mutation studied in this investigation is within this region and is likely involved in the CC α-helical structures, which are known to be rich in leucine residues (O'Shea et al., 1991). Furthermore, leucine to proline mutations, which involve a switch from a flexible hydrophobic residue, to a fixed, cyclic structure, are known to alter CC structures by introducing kinks at the N-terminal side of the substitution site (Chang et al., 1999). It is therefore a reasonable speculation that the L36P mutation in Cog4 may impair Rab30 binding through disruption of a potential binding site in the CC domain at the N-terminus.

Interactions between COG subunits and Rab proteins have been well documented (Miller et al., 2013, Fukuda et al., 2008). However, the functional roles of COG-Rab interactions, including the specific interaction between Rab30 and Cog4, have yet to be determined beyond speculation. The COG complex is localized to the Golgi apparatus and has no spatial or functional preference for any specific cisterna, meaning COG is involved in all Golgi retrograde trafficking pathways (Vasile et al., 2006). However, several studies have highlighted potentially distinct functions of COG lobe A and lobe B, with lobe A subunits necessary for early/medial Golgi protein trafficking and lobe B subunits primarily involved in late Golgi protein trafficking (Oka et al., 2004, Oka et al., 2005, Peanne et al., 2010). Indeed, each subunit of the COG complex has a distinct set of interacting partners, strongly pointing to a functional role of the COG complex as a hub of interacting vesicle tethering proteins, with specific combinations of interactions, including COG-Rab interactions, driving targeted vesicle tethering (reviewed in Fisher and Ungar, 2016). The interaction between Cog4 and Rab30 may therefore be an important component of a particular combination of tethering protein interactions. Rab30 is localized to the Golgi but has no *cis-trans* preference (Kelly et al., 2012), suggesting that the interaction between Rab30 and Cog4 alone is insufficient to drive targeted vesicle tethering. Several interacting partners for Rab30 do however have *cis/trans* preference, including GM130, which is a *cis*-Golgi localized golgin, and TGN localized golgin-97 and golgin-245 (Sinka et al., 2008). Disruption of the Cog4-Rab30 interaction may therefore have knock-on effects on vesicle tethering and therefore on the retrograde trafficking of glycan-processing enzymes in the Golgi, thus altering glycosylation.

5.2 The Impact of Cog4L36P Mutation in CHO-K1 Cells on Glycosylation

Considering the hypothesis that disrupting the Cog4-Rab30 vesicle-tethering interaction would impact glycosylation enzyme sorting, it was important to establish the impact of the Cog4L36P mutation on glycosylation. Mutations in Cog4 have previously been shown to alter glycosylation, for example reduced galactosylation and sialylation is seen in patients with CDGs caused by Cog4R729W mutation (Reynders et al., 2009) and Cog4L773R mutation (Ng et al., 2011). Additionally, the development of COG subunit knock-out HEK293T cells revealed deficiencies in *cis*/medial Golgi glycosylation as well as altered sialylation and fucosylation (Bailey Blackburn et al., 2016, Blackburn and Lupashin, 2016). HEK293T cells lacking COG subunits also exhibited changes to Golgi morphology, retrograde trafficking, and sorting (Bailey Blackburn et al., 2016, Blackburn and Lupashin, 2016), highlighting the link between retrograde trafficking and glycosylation, and the importance of every COG subunit for these processes. Together, these studies add weight to the hypothesis that disrupting the Cog4-Rab30 interaction, which is likely important for vesicle tethering during retrograde trafficking, may affect glycosylation.

Prior to the development of the CRISPR engineered Cog4L36P cell line, previous work carried out in the Ungar lab using ΔCog4L36P (knock-in) cells suggested the mutation resulted in a glycan profile distinct from that of both WT CHO-K1 cells and ΔCog4 cells. However, as the knock-in cells overexpress Cog4, it was important to perform whole-cell glycan profiling of CRISPR-engineered Cog4L36P cells, which express mutant Cog4 under its endogenous promoter, to establish whether the changes were due to Cog4 overexpression or due to the mutation. Additionally, since WT, ΔCog4 and Cog4L36P CHO-K1 cells were originally grown in adherent culture and adapted to suspension culture for biologic production, it was valuable to compare glycan profiles of each cell line in both adherent culture and after suspension adaptation.

Whole-cell *N*-glycan profiling, presented in chapter 3, revealed altered glycosylation in adherent-cultured Cog4L36P CHO-K1 cells, which had a glycan profile distinct from those of both WT and ΔCog4 cells. Differences included a much higher relative abundance of $GlcNAc_2Man_6$ in Cog4L36P cells compared to WT and ΔCog4 cells, and

lower abundance of GlcNAc-terminating glycans in both Cog4L36P and ΔCog4 cells compared to WT. There was also a difference between the relative levels of $GlcNAc_2Man_5$ and $GlcNAc_2Man_6$ in WT and Cog4L36P cells, with adherent Cog4L36P cells having a higher abundance of the Man_6 species compared to WT, but a lower abundance of the Man_5 species. This suggests that *cis*-Golgi glycosylation was affected by the mutation, as this is the site of oligomannose glycan trimming.

At the time of writing, there are no published glycan profiles for ΔCog4 or mutant Cog4 CHO-K1 cell lines for comparison. However, the COG complex was originally identified in adherent CHO-K1 cell lines lacking Cog1 or Cog2, termed ldlB and ldlC cells (Krieger et al., 1981, Podos et al., 1994, Chatterton et al., 1999, Ungar et al., 2002), which do have extensively characterized glycosylation defects (Kingsley et al., 1986). Kingsley et al. treated glycoproteins expressed by ldlB and ldlC cells with glycosidases, showing they had normal oligomannose *N*-glycan synthesis but further processing to complex glycans was defective (Kingsley et al., 1986). Abdul-Rahman and colleagues, who published their results in their paper describing the FANGS method, also generated whole-cell *N*-glycan profiles of ldlB and ldlC cells (Abdul Rahman et al., 2014). These results indicated that processing to complex glycans was defective in the cells lacking Cog1 or Cog2, in line with similar observations by Kingsley et al. Similarly, higher levels of oligomannose glycans overall and therefore lower levels of complex glycans overall were observed in Cog4L36P adherent CHO-K1 cells than in WT cells, suggesting that normal COG complex function is important for glycan processing after the *cis/*medial Golgi. However, in both adherent and suspension-cultured ΔCog4 CHO-K1 cells, the relative abundance of many complex glycans was higher than in WT cells, which could mean that knocking out different COG subunits results in different changes to glycosylation.

Whole-cell glycan profiling has also previously been carried out on other adherent mammalian cell lines, HeLa and HEK293T, lacking Cog4 (Bailey Blackburn et al., 2016, Fisher et al., 2019), allowing comparison of the results presented in this thesis for ΔCog4 adherent cells to those previous publications. No significant differences in the proportion of fucosylated complex *N*-glycans were observed in ΔCog4 or Cog4L36P adherent CHO-K1 cells compared to WT, in contrast with previous work in HEK293T

cells lacking Cog4, which had a lower abundance of fucosylated glycans than did WT (Bailey Blackburn et al., 2016). However, reduced fucosylation in ΔCog4 HEK293T cells compared to WT was found to be due to different abundance of complex glycans available for fucosylation, as opposed to impairment of the fucosyltransferase enzyme, Fut8, which was shown, through computational modelling, to be unaffected by Cog4 knockout (Fisher et al., 2019). The results presented in this thesis are therefore in agreement with the previous literature that Cog4 disruption does not alter Fut8 activity.

Bailey Blackburn *et al.* observed a decrease in sialylation in ΔCog4 HEK293T cells with respect to WT cells, which contrasts with my findings in both Cog4L36P and ΔCog4 adherent cells. Interestingly, the glycan profiles presented in this thesis showed a lower proportion of galactosylated and sialylated complex glycans in the Cog4L36P and ΔCog4 suspension cells than in either their adherent parent cells or WT cells (both adherent and suspension-cultured), which does match the findings by Bailey Blackburn et al. Another discrepancy between my results and those of Bailey Blackburn *et al.* was that knocking out Cog4 in HEK293T cells resulted in elevated abundance of $GlcNAc_2Man_5$ but lower abundance of $GlcNAc_2Man_6$ compared to WT (Bailey Blackburn et al., 2016), whereas the opposite trend was observed in the glycan profiles of ΔCog4 and Cog4L36P adherent CHO-K1 cells. However, adaptation to suspension culture resulted in increased abundance of $GlcNAc_2Man_5$ and reduced $GlcNAc_2Man_6$ levels in ΔCog4 and Cog4L36P CHO-K1 cells compared to WT, in agreement with findings in HEK293T cells lacking Cog4 (Bailey Blackburn et al., 2016).

A possible explanation for the similarities between the effects of disrupting Cog4 in suspension CHO-K1 cells and HEK293T cells on glycosylation is that HEK293T cells are known to have weaker adhesion than CHO-K1 cells, so glycosylation of adhesion molecules in HEK293T cells may be more closely related to that of CHO-K1 cells grown in suspension than in adherent culture. Indeed, there is evidence that the same adhesion molecule, the extracellular domain of a protein called GlialCAM, exhibited different glycosylation when expressed in CHO-K1 and HEK293T cells grown in adherent culture (Gaudry et al., 2008). Cog4L36P adherent-cultured cells adhere more weakly than WT CHO-K1 cells, which may also explain in part why their adaptation to

suspension culture caused fewer changes to their glycan profile compared to suspension-adaptation of WT and ΔCog4 cells.

5.3 Suspension Adaptation Alters Glycosylation

Adaptation to suspension culture is known to cause significant changes to cell surface adhesion-related glycoprotein expression (Klingler et al., 2021, Walther et al., 2016), so it was expected for suspension adaptation to alter the whole-cell glycan profile. Indeed, all CHO-K1 cell lines tested in this work had altered glycan profiles following their adaptation to suspension culture. WT CHO-K1 cells had altered levels of several oligomannose glycans following suspension adaptation, reduced abundance of glycans terminating in GlcNAc and an increase in galactose-terminating glycans. Adapting ΔCog4 cells to suspension culture caused marked differences to their whole cell glycan profile, including reduced oligomannose glycan abundance and an increase in the number and abundance of complex glycans, with the majority terminating in GlcNAc. Cog4L36P cells had fewer sialylated glycans overall following suspension adaptation and more glycans terminating in GlcNAc compared to their adherent glycan profile. However, most Cog4L36P glycans had comparable abundance in both adherent and suspension culture, with no major differences in any one particular glycan standing out. This indicates that adaptation from adherent culture requires fewer glycosylation changes in the Cog4L36P cells than in WT or ΔCog4 cells, which could suggest that glycosylation of adhesion molecules is deficient in L36P cells. Overall, these results highlight the impact of adhesion and suspension adaptation on glycosylation, with different responses in Cog4L36P and ΔCog4 cells demonstrating a potential link between Cog4 and glycosylation of glycoproteins involved in adhesion.

Previously, researchers have compared whole-cell glycan profiles between adherent CHO cells (mCHO) and suspension CHO cells (sCHO) from the same parent cell line (North et al., 2010). They found the cells had similar glycan profiles in both adherent and suspension culture but observed decreased sialylation in the sCHO cells. Decreased sialylation was also observed following suspension adaptation of both Cog4L36P cells and ΔCog4 cells. However, both adherent and suspension-cultured WT CHO-K1 cells had similar levels of sialylated glycans, in contrast to findings by North *et al.* This discrepancy could be because the CHO cells studied by North *et al.* were CHO Pro[-5] cells

rather than CHO-K1, a different subclone of the original CHO cell line, which may have different properties to CHO-K1 cells due to clonal variation. Indeed, cell surface glycosylation of both CHO-K1 and CHO Pro^{-5} adherent-cultured cells was investigated in a study looking at viral attachment to cell surfaces, revealing that CHO Pro^{-5} cells had significantly less terminally sialylated glycans on their cell surface than CHO-K1 cells (Gillespie et al., 2016). Additionally, the medium used by North *et al.* for suspension culture was the same as that used for adherent culture, with 10% FBS present (North et al., 2010). The presence of FBS has previously been shown to affect glycosylation of recombinantly expressed mAbs (Costa et al., 2013), so whole cell glycosylation might be affected by FBS too. This may explain why glycan profiles of the CHO-K1 cells studied here showed more differences after suspension adaptation than those of CHO Pro^{-5} cells, as CHO-K1 cells were adapted from adherent growth in medium containing 10% FBS to growth in serum-free suspension medium.

To date, most studies investigating the effects of suspension adaptation on CHO cell glycosylation have focussed on studying the impact on glycosylation of recombinantly expressed protein (Costa et al., 2013, Novo et al., 2014), making comparison of my results to existing literature challenging. However, discrepancies between the results presented in this thesis and those published by North *et al.* highlight the glycosylation differences between clonal CHO cell lines and the importance of medium composition on whole-cell glycosylation.

5.4 Physiological Impact of the Cog4L36P Mutation

To assess the physiological impact of Cog4L36P mutation, its impaired interaction with Rab30, and resulting changes to glycosylation, characterisation of various cell properties was carried out. This included analysing proliferation, metabolic activity, and adhesion of Cog4L36P cells compared to WT (chapter 2). Despite being only a single base change, the effects of the Cog4L36P mutation were striking. Both adherent and suspension cultured Cog4L36P cells had slower proliferation than their WT parent cells, yet higher metabolic activity. Cog4L36P cells also had an adhesion defect, with evidence suggesting that their cell surface adhesion molecules and their secreted ECM are impaired in some way compared to those of WT cells.

5.4.1 Relationship Between the COG Complex and Adhesion

Thus far, no link between Cog4 or any other component of the COG complex and adhesion has been published. However, there are several plausible explanations as to how disrupting Cog4 may lead to reduced adhesion. Adhesion of CHO and other fibroblast-like cells to surfaces and to other cells involves an intricate network of interactions between secreted extracellular matrix proteins, including collagens and fibronectin, and cell-surface adhesion molecules, including integrins and cadherins (reviewed in Parsons et al., 2010). Glycoconjugates called glycosaminoglycans (GAGs), which include heparan sulfates (HS), also decorate the cell surface and interact with other cell-surface adhesion molecules and ECM proteins to regulate adhesion (Vuoriluoto et al., 2008). Cell surface adhesion molecules span the cell membrane and interact with intracellular cytoskeletal proteins and components of signalling cascades. Many ECM proteins and nearly all cell surface adhesion molecules are glycosylated and changes to their glycosylation can alter their adhesion (section 1.1.7.4), so changes in glycosylation caused by Cog4L36P mutation could influence adhesion if glycosylation of cell-surface adhesion molecules is affected. Alternatively, trafficking of cell surface adhesion or secretion of ECM proteins could be altered in Cog4L36P cells, possibly because of altered glycosylation. For example, the importance of glycosylation for trafficking of adhesion molecules was highlighted by Isaji *et al.,* who showed that abolishing an *N*-glycosylation site on α5β1 expressed by CHO cells prevented correct folding and ER exit (Isaji et al., 2006). Glycosylation and cell-surface expression of integrins in WT and Cog deficient CHO-K1 cell lines therefore warrants further investigation.

Although there are no published examples of adhesion defects in COG deficient cell lines, a direct link between Cog4 and cell surface molecules involved in adhesion has been recently uncovered. Subunits of the COG complex, including Cog4, were associated with components of the HS synthesis pathway in two separate loss-of-function based genome wide screens of genes associated with cell survival following viral infection (Petitjean et al., 2020, Riblett et al., 2015). Both studies identified mutants of several components of the COG complex as well as components of the HS pathway that were enriched following treatment with viral dsRNA. Petitjean *et al.* also showed that Cog4 knockout led to a decrease in extracellular HS, which conferred

resistance to viral infection and increased cell survival (Petitjean et al., 2020). HS has specifically been shown to mediate adhesion in CHO cells (Franco et al., 2001, Vuoriluoto et al., 2008). GAG synthesis in HEK293T cells lacking COG subunits has also been shown to be differentially affected depending on which COG subunit is knocked out (Adusumalli et al., 2021). Additionally, a very recent publication specifically implicated Cog4 in HS expression at the cell surface in fibroblasts derived from patients with Saul-Wilson Syndrome (SWS), a CDG caused by a heterozygous, dominant variant (p.G516R) in Cog4 (Xia et al., 2021). In this case, SWS fibroblasts displayed an accumulation of glypicans, a type of HS, at their cell surface, most likely as a result of abnormal trafficking or turnover, as opposed to transcriptional regulation (Xia et al., 2021). This finding was replicated in a cellular model for SWS expressing Cog4G516R (Sumya et al., 2021). Together, these findings suggest that targeting COG subunits can result in specific changes to GAG function, offering a credible mechanistic explanation for the observation of impaired adhesion in Cog4L36P cells, which warrants further investigation.

5.4.2 Relationship Between the COG Complex, Proliferation and Metabolic Activity

Previous characterization of HEK293T cells lacking COG subunits revealed that no subunit was essential for cell growth, as growth curves for all COG subunit-knockout cells were no different from those of WT cells (Bailey Blackburn et al., 2016). This is consistent with results presented in this thesis demonstrating that the rate of proliferation of ΔCog4 adherent CHO-K1 cells was comparable to that of WT cells. However, slower proliferation of Cog4L36P compared to WT indicates that the Cog4L36P mutation affects cell behaviour by mechanisms distinct from those of knocking out Cog4. Indeed, evidence of an adhesion defect and slower proliferation in Cog4L36P cells is consistent with well-established links between adhesion and proliferation. Adhesion molecules, such as integrins, also initiate intracellular signalling cascades that can promote proliferation (reviewed in Schwartz and Assoian, 2001). The adhesion defect observed in Cog4L36P cells may therefore be a causative factor for slower proliferation.

Existing evidence for the relationship between adhesion and proliferation in CHO cells includes the finding that overexpressing β1 integrin in CHO cells conferred increased proliferation in response to fibronectin (Pasqualini and Hemler, 1994). A later study showed that treatment with fucan B, a sulfated L-fucose-rich polysaccharide, impairs adhesion of CHO-K1 cells by binding to fibronectin (Rocha et al., 2005). Fucan-induced adhesion impairment was later shown to cause an antiproliferative effect on CHO cells, due to activation of integrin α5β1, which resulted in activation of the FAK/RAS/MEK/ERK signalling pathway, promoting G_1 cell cycle arrest (Nobre et al., 2013). Impaired proliferation observed in Cog4L36P cells may therefore be linked to altered adhesion. Other lines of evidence linking adhesion to proliferation include the observation that the cyclinE-cyclin dependent kinase (CDK) complex, which is required for the transition between G_1 and S phases of the cell cycle, is activated in adherent fibroblasts but not in suspension-cultured fibroblasts (Fang et al., 1996). Moreover, CDKs are downstream of signalling pathways that are under the control of integrins, highlighting another link between integrin-mediated adhesion and proliferation (Schwartz and Assoian, 2001). This is consistent with the observation that proliferation of all tested cell lines was slower in suspension culture than adherent culture (chapter 2).

The results presented in this thesis also suggest that Cog4L36P cells have higher metabolic activity, as evidenced by their ability to reduce MTT and alamarBlue® at a higher rate than WT cells, despite slower proliferation. A possible explanation for this could be that Cog4L36P cells are proliferating and metabolizing faster than WT but cell death may be occurring at a higher rate. Adhesion has also been linked to cell death through the integrin-mediated activation of apoptosis-inducing signalling cascades (reviewed in Ruoslahti and Reed, 1994), so the adhesion defect in Cog4L36P cells might also influence cell survival. Further investigation of apoptosis in WT and Cog4L36P cells is therefore recommended.

5.5 Potential of Cog4L36P CHO Cells as Hosts for Recombinant Expression of Glycoprotein Biologics

A key aim of this research was to explore the potential of Cog4L36P cells as hosts for the recombinant expression of glycoprotein biologics, with a particular focus on mAbs.

The observation that the glycan profile of Cog4L36P mutant cells was distinct from that of WT and ΔCog4 cells suggests that glycosylation of recombinantly expressed proteins might be altered in the cells too. Given the influence of glycosylation on important properties of biotherapeutics, the ability to engineer glycosylation is highly attractive to the biopharmaceutical industry. Herceptin glycan profiling results, shown in chapter four, suggest that the Cog4L36P mutation does not alter glycosylation of the mAb Fc region. However, the MS technique used to assess intact heavy chain glycosylation is not quantitative and low abundance glycans are difficult to pick out after deconvolution due to the formation of multiple differently charged forms. Therefore, it is not possible to conclude that the Cog4L36P mutation does not affect glycosylation of recombinantly-expressed mAbs at the level of minor glycoforms that are below the limit of detection in either the WT or the Cog4L36P mutant products. In future, following up on all the recommendations for future work presented in chapter four, including cleaving *N*-glycans from Herceptin and derivatisation prior to MS analysis, would help address this uncertainty.

Enhancing cell productivity is also a major goal in the production of biologics. Characterisation of Cog4L36P cells suggested that some of their properties, particularly enhanced metabolic activity, and reduced adhesion, may be favourable for biologic production in suspension culture. The results presented in this thesis also show that production of Herceptin is increased in Cog4L36P suspension-adapted CHO-K1 cells, which is hugely promising. However, it is important to consider natural variation when interpreting these results, as CHO cells are known to be genetically plastic and exhibit clonal variation (Xu et al., 2011). Phenotypic variation can be observed within a single population of CHO cells: an observation first acknowledged in the 1970's for CHO cells cultured on agar plates (Konrad et al., 1977). On agar, Konrad et al. found that CHO cells formed distinctly shaped colonies and subsequent single cell cloning of cells from these colonies revealed that subcloned cells shared the same morphology as their parent colony, suggesting that differences were due to genetic variation (Konrad et al., 1977). More recently, Davies et al. also found that un-transfected, clonally derived CHOK1SV cells (a derivative of CHO-K1) exhibited significant variability in their proliferation rates, total protein biomass, ability to transiently produce recombinant proteins, and in their cell surface *N*-glycosylation (Davies et al., 2013).

Given that such variability can be observed from clones originating from a single population, it could be argued that phenotypic differences observed between the clonal cell lines studied in this thesis could be in part due to natural variation, independent of Cog4. However, the observation that both ΔCog4L36P and CRISPR-engineered Cog4L36P cells exhibited enhanced metabolic activity and produced more Herceptin than WT control cell lines (ΔCog4WT and WT) adds weight to the theory that the differences are due to the Cog4L36P mutation. A more recent experiment in the Ungar lab also supports the hypothesis that the L36P mutation in Cog4 is beneficial for mAb production. In this experiment, performed by John Petrie, suspension CHO cell cultures stably expressing Herceptin were transiently transfected with either an empty pcDNA plasmid, a pcDNA plasmid containing WT Cog4 DNA or a pcDNA plasmid containing Cog4L36P DNA. Herceptin production was monitored for a week and the results showed that production was highest in the cells transiently transfected with the pcDNA Cog4L36P plasmid. This result was reproducible and statistically significant (personal communication from Dani Ungar), suggesting that expression of Cog4L36P enhanced Herceptin production. Considering increased Herceptin production has now been observed in three different CHO cell lines expressing Cog4L36P, it is unlikely that clonal variation alone can account for enhanced production. Nevertheless, in future, the extent to which clonal variation contributes to the observed differences could be examined by generating multiple clones of each genotype, as described in section 2.6.

Results presented in this thesis also suggest that Herceptin production in Cog4L36P cells may be dependent on cell culture age, as Cog4L36P cells that had been in culture for several months produced less Herceptin than Cog4L36P cultures at a lower passage number. However, genetic plasticity is a known phenomenon in CHO cells, and even clonally derived recombinant CHO cell lines are known exhibit changes in biotherapeutic production over time (Ko et al., 2018, Barnes et al., 2003, Frye et al., 2016). Therefore, it is common practice to routinely return to frozen stocks of a cell line and limit the number of passages. Using this approach may be expected to ensure consistently high titres of Herceptin in Cog4L36P cells. If the use of quantifiable glycan profiling techniques can also show that there is indeed no change to mAb glycosylation in Cog4L36P cells, this would be encouraging, as enhanced production of mAbs with unaltered glycosylation would be hugely beneficial.

The mechanism by which Cog4L36P mutation enhances production of Herceptin requires further investigation. One hypothesis is that anterograde trafficking and secretion are faster in Cog4L36P cells. This could be because Golgi morphology may be altered in Cog4L36P cells in such a way that proteins spend less time in the Golgi and are secreted faster. Altered Golgi morphology in Cog4L36P cells is plausible, since Rab30 and Cog4 are both essential for maintaining the morphological integrity of the Golgi (Kelly et al., 2012, Bailey Blackburn et al., 2016) and the Cog4L36P mutation disrupts the Cog4-Rab30 interaction. Alterations to Golgi morphology are also known to influence trafficking and secretion. For example, mutation or knock-out of COG subunits has been shown to disrupt the structure of the Golgi (Bailey Blackburn et al., 2016, Blackburn et al., 2018), causing defective trafficking and secretion. Previous publications generally show a trend of impaired trafficking and secretion in COG deficient and COG-CDG-derived cells (Bailey Blackburn et al., 2016, Blackburn et al., 2018, Blackburn et al., 2019, D'Souza et al., 2019). However, mutation of Cog subunits is not always detrimental to Golgi morphology or secretion. For example, in a recently developed human retinal pigmented epithelial (RPE1) cell line expressing Cog4G516R, the mutation associated with SWS, Golgi morphology and secretion of almost all tested protein was normal and comparable to that in WT cells, but secretion of ERGIC53 and SIL1 was enhanced (Sumya et al., 2021). These proteins normally reside in the ER-Golgi intermediate compartment (ERGIC) and ER respectively, but impaired retrograde trafficking causes their aberrant secretion instead of recycling to their proper location.

Despite improved production of transiently expressed Herceptin in Cog4L36P cells compared to WT, production levels, which were in the 5-10 µg/mL range, were still much lower than those reported for industrial ex-production host cell lines, which produced up to 60 µg/mL (mg/L) Herceptin. In comparison, current production cell lines, which stably express mAbs, can achieve product titres exceeding 10 g/L (Ha et al., 2022, Handlogten et al., 2018). WT and Cog4L36P CHO-K1 cells presented in this thesis were adapted to suspension culture after using CRISPR-cas9n to engineer the Cog4L36P mutation, following a gradual adaptation process. Similarly, it used to be common practice for adherent CHO cells to be stably transfected to produce recombinant proteins before adapting them to grow in suspension culture (e.g. Hayter et al., 1991). However, the adaptation process takes weeks to months, and freshly adapted cells

have low tolerance to metabolic by-products released by cells, including ammonia and lactic acid, which limits them to grow at low densities ($<2 \times 10^6$ cells/mL) (Sinacore et al., 2000), thus limiting the titre of recombinantly expressed protein. During the process of adaptation presented in this thesis, it was also observed that freshly adapted cells showed a tendency to clump together. Increasing tolerance to growth-inhibiting substances, enabling cells to grow at higher densities, requires further selection (Rodrigues et al., 2013, Sinacore et al., 2000).

Given the known limitations of freshly suspension-adapted cells and the fact that further clonal isolation of Cog4L36P suspension-adapted cells was not possible in the timescale of this PhD, it was hugely beneficial to test the potential of the Cog4L36P mutation in a production CHO cell line, which had already been selected on the basis of enhanced production of recombinant proteins. Engineering the mutation into a legacy-production CHO suspension-cell line provided by GSK enabled testing of whether Cog4L36P mutation could enhance production levels beyond the level achieved by selection of high producing WT clones. Results of a transient transfection of WT and Cog4L36P legacy production CHO cells with Herceptin revealed no difference in Herceptin production. The discrepancy between the effect of the mutation in freshly suspension-adapted CHO-K1 cells and in the legacy production cells could be explained by a potential secretion bottleneck in WT CHO-K1 suspension-adapted cells, which is overcome by the Cog4L36P mutation. Since the legacy production cells were selected for their high titre production of mAbs, an increase in secretion conferred by the Cog4L36P mutation may have negligible impact if a secretion bottleneck was overcome by other means.

5.6 Recommendations of Future Research Related to This Research

Extensive further experimentation is required to elucidate the full potential of Cog4L36P cells for biologic production. Recommendations for many of these experiments are summarised in section 4.6. Here, further recommendations are proposed based on discussion of potential mechanisms underlying observed phenotypic changes and enhanced mAb production conferred by the Cog4L36P mutation.

Following the hypothesis that Cog4L36P mutation disrupts Rab30 binding through the introduction of a kink in a CC Rab-binding domain, the structural requirements for the Cog4-Rab30 interaction could be explored further. This could be investigated through targeted mutagenesis within the CC domain, for example mutating Leu36 to a more similar amino acid such as valine or isoleucine that would cause a less severe distortion to the tertiary structure, and so would allow distinction of whether primary or tertiary structure is more important for Rab30 binding. Other nearby leucine residues could also be mutated to proline and their Rab binding could be investigated, to determine whether specific regions of the CC structure act as binding sites for specific Rabs.

In light of the novel observation that a mutation in Cog4 causes an adhesion defect, further investigation is needed to understand the underlying mechanism. Experiments proposed in section 2.6 included glycan profiling of secreted ECM and cell surface proteins to establish whether glycosylation changes are responsible for reduced adhesion. Additionally, proteomic analysis to determine levels of ECM and cell surface proteins in each cell line is also recommended. This would inform whether expression, processing or trafficking of cell surface adhesion molecules or secretion of ECM proteins is affected by the mutation. Following recent publications implicating COG in HS synthesis, analysis of cell surface GAGs on Cog4L36P cells should also be performed. GAGs could be analysed using similar methodologies to those employed by the Lupashin group in their characterisation of cell surface GAGs in COG knockout HEK293T cells (Adusumalli et al., 2021), which included metabolic ^{35}S labelling of GAGs in addition to gel filtration analysis.

Further investigation is also needed to understand the mechanism of increased Herceptin secretion in Cog4L36P cells. One hypothesis is that the rate of trafficking is altered by the mutation. This could be investigated by monitoring trafficking of VSVG protein through both WT and Cog4L36P cells. A mutant form of this protein, VSVG[ts045] is temperature-sensitive for exit from the ER, meaning its transport from the ER to the cell-surface can be synchronized with a temperature ramp to the permissive temperature (Bergmann, 1989). GFP-tagged VSVG [ts045] can be visualized by fluorescence microscopy (Presley et al., 1997) and the amount of fluorescence in each cellular compartment can be quantified in order to establish the trafficking rate. This

experiment was performed comparing Cog4L36P and WT cells, and differences between the cell lines were observed, notably faster trafficking from the ER to the Golgi and shorter time spent in the Golgi in Cog4L36P cells compared to WT cells (personal communication from Ungar/Bryant group colleague, John Petrie). More biological repeats are needed for this experiment, in addition to comparison to trafficking in ΔCog4 cells, to establish whether Cog4L36P mutation causes a trafficking abnormality distinct from that caused by knockout of Cog4. Trafficking abnormalities can also be caused by morphological changes to the Golgi, so electron microscopic comparison of Cog4L36P and WT cells could be performed to visualise the Golgi complex in each cell line.

Considering the observation that Cog4L36P mutation alters the whole-cell glycan profile, potentially without alteration of mAb Fc glycosylation, it would be highly informative to use computational modelling to predict mAb glycan profiles from whole-cell glycan profiling data. Following quantifiable glycan profiling of Herceptin heavy chain secreted by WT and Cog4L36P cells, and identification of low abundance glycoforms that were missed by ESI-MS analysis, these results could be input into the model to improve understanding of which specific glycan processing steps are affected by Cog4L36P mutation. Development of this model, which builds on a previous model for predicting whole-cell glycosylation, developed by Ungar group colleague Peter Fisher (Fisher et al., 2019), is underway by current PhD student, Ben West.

Expanding on the work presented in this thesis, which suggests that targeting COG complex subunits could be beneficial for biologic production, other COG subunit mutations could be explored in production CHO cells. For example, CDG-causing COG subunit mutations could be tested for their effects on glycosylation of recombinant protein, as these are known to cause a range of glycosylation changes. This PhD has only scratched the surface of investigating the potential for targeting the COG complex to aid the production of biologics.

References

AALTO, M. K., KERÄNEN, S. & RONNE, H. 1992. A family of proteins involved in intracellular transport. *Cell,* 68**,** 181-182.

ABDUL RAHMAN, S., BERGSTROM, E., WATSON, C. J., WILSON, K. M., ASHFORD, D. A., THOMAS, J. R., UNGAR, D. & THOMAS-OATES, J. E. 2014. Filter-aided N-glycan separation (FANGS): a convenient sample preparation method for mass spectrometric N-glycan profiling. *Journal of Proteome Research,* 13**,** 1167-76.

ADUSUMALLI, R., ÅSHEIM, H.-C., LUPASHIN, V., BLACKBURN, J. B. & PRYDZ, K. 2021. Proteoglycan synthesis in conserved oligomeric Golgi subunit deficient HEK293T cells is affected differently, depending on the lacking subunit. *Traffic (Copenhagen, Denmark),* 22**,** 230-239.

AGHAMOHSENI, H., OHADI, K., SPEARMAN, M., KRAHN, N., MOO-YOUNG, M., SCHARER, J. M., BUTLER, M. & BUDMAN, H. M. 2014. Effects of nutrient levels and average culture pH on the glycosylation pattern of camelid-humanized monoclonal antibody. *Journal of Biotechnology,* 186**,** 98-109.

ALESSANDRI, L., OUELLETTE, D., ACQUAH, A., RIESER, M., LEBLOND, D., SALTARELLI, M., RADZIEJEWSKI, C., FUJIMORI, T. & CORREIA, I. 2012. Increased serum clearance of oligomannose species present on a human IgG1 molecule. *MAbs,* 4**,** 509-20.

ALPERT, A. J. 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *Journal of Chromatography,* 499**,** 177-96.

ALVAREZ, C., GARCIA-MATA, R., HAURI, H.-P. & SZTUL, E. 2001. The p115-interactive Proteins GM130 and Giantin Participate in Endoplasmic Reticulum-Golgi Traffic*. *Journal of Biological Chemistry,* 276**,** 2693-2700.

AMANN, T., SCHMIEDER, V., FAUSTRUP KILDEGAARD, H., BORTH, N. & ANDERSEN, M. R. 2019. Genetic engineering approaches to improve posttranslational modification of biopharmaceuticals in different production platforms. *Biotechnology and Bioengineering,* 116**,** 2778-2796.

ANDAG, U. & SCHMITT, H. D. 2003. Dsl1p, an Essential Component of the Golgi-Endoplasmic Reticulum Retrieval System in Yeast, Uses the Same Sequence Motif to Interact with Different Subunits of the COPI Vesicle Coat*. *Journal of Biological Chemistry,* 278**,** 51722-51734.

ASHWELL, G. & HARFORD, J. 1982. Carbohydrate-specific receptors of the liver. *Annual Review of Biochemistry,* 51**,** 531-54.

ASHWELL, G. & MORELL, A. G. 1974. The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. *Advances in Enzymology Related Areas of Molecular Biology,* 41**,** 99-128.

ATKINSON, P. H. & LEE, J. T. 1984. Co-translational excision of alpha-glucose and alpha-mannose in nascent vesicular stomatitis virus G protein. *Journal of Cell Biology,* 98**,** 2245-9.

BAILEY BLACKBURN, J., POKROVSKAYA, I., FISHER, P., UNGAR, D. & LUPASHIN, V. V. 2016. COG Complex Complexities: Detailed Characterization of a Complete Set of HEK293T Cells Lacking Individual COG Subunits. *Frontiers in Cell and Developmental Biology,* 4**,** 23.

BALLOU, L., COHEN, R. E. & BALLOU, C. E. 1980. Saccharomyces cerevisiae mutants that make mannoproteins with a truncated carbohydrate outer chain. *Journal of Biological Chemistry,* 255**,** 5986-5991.

BARNES, L. M., BENTLEY, C. M. & DICKSON, A. J. 2003. Stability of protein production from recombinant mammalian cells. *Biotechnology and Bioengineering,* 81**,** 631-639.

BATEMAN, A. C., KARAMANSKA, R., BUSCH, M. G., DELL, A., OLSEN, C. W. & HASLAM, S. M. 2010. Glycan analysis and influenza A virus infection of primary swine respiratory epithelial cells: the importance of NeuAc{alpha}2-6 glycans. *Journal of Biological Chemistry,* 285**,** 34016-26.

BECK, A. 2013. Glycosylation engineering of biopharmaceuticals. *Methods in Molecular Biology,* 988.

BEREK, C. & MILSTEIN, C. 1988. The Dynamic Nature of the Antibody Repertoire. *Immunological Reviews,* 105**,** 5-26.

BERGER, E. G., BUDDECKE, E., KAMERLING, J. P., KOBATA, A., PAULSON, J. C. & VLIEGENTHART, J. F. 1982. Structure, biosynthesis and functions of glycoprotein glycans. *Experientia,* 38**,** 1129-62.

BERGMANN, J. E. 1989. Chapter 4 Using Temperature-Sensitive Mutants of VSV to Study Membrane Protein Biogenesis. *In:* TARTAKOFF, A. M. (ed.) *Methods in Cell Biology.* Academic Press.

BERRIDGE, M. V., HERST, P. M. & TAN, A. S. 2005. Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. *Biotechnology Annual Review.* Elsevier.

BERTING, A., FARCET, M. R. & KREIL, T. R. 2010. Virus susceptibility of Chinese hamster ovary (CHO) cells and detection of viral contaminations by adventitious agent testing. *Biotechnology and Bioengineering,* 106**,** 598-607.

BHATTACHARYYA, L., KHAN, M. I., FANT, J. & BREWER, C. F. 1989. Formation of highly ordered cross-linked lattices between asparagine-linked oligosaccharides and lectins observed by electron microscopy. *Journal of Biological Chemistry,* 264**,** 11543-5.

BIGGE, J. C., PATEL, T. P., BRUCE, J. A., GOULDING, P. N., CHARLES, S. M. & PAREKH, R. B. 1995. Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Analytical Biochemistry,* 230**,** 229-238.

BLACKBURN, J. B., D'SOUZA, Z. & LUPASHIN, V. V. 2019. Maintaining order: COG complex controls Golgi trafficking, processing, and sorting. *FEBS letters,* 593**,** 2466-2487.

BLACKBURN, J. B., KUDLYK, T., POKROVSKAYA, I. & LUPASHIN, V. V. 2018. More than just sugars: Conserved oligomeric Golgi complex deficiency causes glycosylation-independent cellular defects. *Traffic (Copenhagen, Denmark),* 19**,** 463-480.

BLACKBURN, J. B. & LUPASHIN, V. V. 2016. Creating Knockouts of Conserved Oligomeric Golgi Complex Subunits Using CRISPR-Mediated Gene Editing Paired with a Selection Strategy Based on Glycosylation Defects Associated with Impaired COG Complex Function. *Methods in Molecular Biology,* 1496**,** 145-61.

BLOCK, M. R., GLICK, B. S., WILCOX, C. A., WIELAND, F. T. & ROTHMAN, J. E. 1988. Purification of an N-ethylmaleimide-sensitive protein catalyzing vesicular transport. *Proceedings of the National Academy of Sciences of the United States of America,* 85**,** 7852-6.

BONFANTI, L., MIRONOV, A. A., JR., MARTINEZ-MENARGUEZ, J. A., MARTELLA, O., FUSELLA, A., BALDASSARRE, M., BUCCIONE, R., GEUZE, H. J., MIRONOV, A. A. & LUINI, A. 1998. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell,* 95**,** 993-1003.

BORAIAH, S., PAUL, O., HAWKES, D., WICKHAM, M. & LORICH, D. G. 2009. Complications of recombinant human BMP-2 for treating complex tibial plateau fractures: a preliminary report. *Clinical Orthopedics and Related Research,* 467**,** 3257-62.

BOULENGUER, P., LEROY, Y., ALONSO, J. M., MONTREUIL, J., RICART, G., COLBERT, C., DUQUET, D., DEWAELE, C. & FOURNET, B. 1988. Continuous-flow fast atom bombardment-mass spectrometry of permethylated oligosaccharides: A comparative study of direct mixture analysis with packed capillary column liquid chromatography-fast atom bombardment-mass spectrometry. *Analytical Biochemistry,* 168**,** 164-170.

BOUSFIELD, G. R. & HARVEY, D. J. 2019. Follicle-Stimulating Hormone Glycobiology. *Endocrinology,* 160**,** 1515-1535.

BRÖCKER, C., ENGELBRECHT-VANDRÉ, S. & UNGERMANN, C. 2010. Multisubunit Tethering Complexes and Their Role in Membrane Fusion. *Current Biology,* 20**,** R943-R952.

BROOKS, S. A. 2004. Appropriate glycosylation of recombinant proteins for human use: implications of choice of expression system. *Molecular Biotechnology,* 28**,** 241-55.

BROWN, R. S. & LENNON, J. J. 1995. Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer. *Analytical Chemistry,* 67**,** 1998-2003.

CAMPBELL, M. P., ROYLE, L., RADCLIFFE, C. M., DWEK, R. A. & RUDD, P. M. 2008. GlycoBase and autoGU: tools for HPLC-based glycan analysis. *Bioinformatics,* 24**,** 1214-1216.

CAMPO, S., ANDREONE, L., AMBAO, V., URRUTIA, M., CALANDRA, R. S. & RULLI, S. B. 2019. Hormonal Regulation of Follicle-Stimulating Hormone Glycosylation in Males. *Frontiers in Endocrinology,* 10.

CARAMELO, J. J. & PARODI, A. J. 2015. A sweet code for glycoprotein folding. *FEBS Letters,* 589**,** 3379-3387.

CARR, C. M., GROTE, E., MUNSON, M., HUGHSON, F. M. & NOVICK, P. J. 1999. Sec1p binds to SNARE complexes and concentrates at sites of secretion. *Journal of Cell Biology,* 146**,** 333-44.

CHANG, D. K., CHENG, S. F., TRIVEDI, V. D. & LIN, K. L. 1999. Proline affects oligomerization of a coiled coil by inducing a kink in a long helix. *Journal of Structural Biology,* 128**,** 270-9.

CHANG, M. M., GAIDUKOV, L., JUNG, G., TSENG, W. A., SCARCELLI, J. J., CORNELL, R., MARSHALL, J. K., LYLES, J. L., SAKORAFAS, P., CHU, A.-H. A., COTE, K., TZVETKOVA, B., DOLATSHAHI, S., SUMIT, M., MULUKUTLA, B. C., LAUFFENBURGER, D. A., FIGUEROA, B., SUMMERS, N. M., LU, T. K. & WEISS, R. 2019. Small-molecule control of antibody N-glycosylation in engineered mammalian cells. *Nature Chemical Biology,* 15**,** 730-736.

CHATTERTON, J. E., HIRSCH, D., SCHWARTZ, J. J., BICKEL, P. E., ROSENBERG, R. D., LODISH, H. F. & KRIEGER, M. 1999. Expression cloning of LDLB, a gene essential

for normal Golgi function and assembly of the ldlCp complex. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 915-20.

CHUSAINOW, J., YANG, Y. S., YEO, J. H. M., TOH, P. C., ASVADI, P., WONG, N. S. C. & YAP, M. G. S. 2009. A study of monoclonal antibody-producing CHO cell lines: What makes a stable high producer? *Biotechnology and Bioengineering,* 102**,** 1182-1196.

CIUCANU, I. & COSTELLO, C. 2004. Elimination of Oxidative Degradation during the per-O-Methylation of Carbohydrates. *Journal of the American Chemical Society,* 125**,** 16213-9.

CIUCANU, I. & KEREK, F. 1984. A simple and rapid method for the permethylation of carbohydrates. *Carbohydrate Research,* 131**,** 209-217.

CLARY, D. O., GRIFF, I. C. & ROTHMAN, J. E. 1990. SNAPs, a family of NSF attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell,* 61**,** 709-21.

CLIMER, L. K., POKROVSKAYA, I. D., BLACKBURN, J. B. & LUPASHIN, V. V. 2018. Membrane detachment is not essential for COG complex function. *Molecular Biology of the Cell,* 29**,** 964-974.

COLOMBEL, J. F., SANDBORN, W. J., RUTGEERTS, P., ENNS, R., HANAUER, S. B., PANACCIONE, R., SCHREIBER, S., BYCZKOWSKI, D., LI, J., KENT, J. D. & POLLACK, P. F. 2007. Adalimumab for Maintenance of Clinical Response and Remission in Patients With Crohn's Disease: The CHARM Trial. *Gastroenterology,* 132**,** 52-65.

COMPTON, T. & COURTNEY, R. J. 1984. Evidence for post-translational glycosylation of a nonglycosylated precursor protein of herpes simplex virus type 1. *Journal of Virology,* 52**,** 630-637.

CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science,* 339**,** 819-23.

CONIBEAR, E. & STEVENS, T. H. 2000. Vps52p, Vps53p, and Vps54p Form a Novel Multisubunit Complex Required for Protein Sorting at the Yeast Late Golgi. *Molecular Biology of the Cell,* 11**,** 305-323.

CONTESSA, J. N., BHOJANI, M. S., FREEZE, H. H., REHEMTULLA, A. & LAWRENCE, T. S. 2008. Inhibition of N-Linked Glycosylation Disrupts Receptor Tyrosine Kinase Signaling in Tumor Cells. *Cancer Research,* 68**,** 3803-3809.

CORPER, A. L., SOHI, M. K., BONAGURA, V. R., STEINITZ, M., JEFFERIS, R., FEINSTEIN, A., BEALE, D., TAUSSIG, M. J. & SUTTON, B. J. 1997. Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. *Nature Structural and Molecular Biology,* 4**,** 374-81.

COSTA, A. R., WITHERS, J., RODRIGUES, M. E., MCLOUGHLIN, N., HENRIQUES, M., OLIVEIRA, R., RUDD, P. M. & AZEREDO, J. 2013. The impact of cell adaptation to serum-free conditions on the glycosylation profile of a monoclonal antibody produced by Chinese hamster ovary cells. *New biotechnology,* 30**,** 563-572.

COTTAM, N. P. & UNGAR, D. 2012. Retrograde vesicle transport in the Golgi. *Protoplasma,* 249**,** 943-55.

COTTER, R. J. 1992. Time-of-Flight Mass Spectrometry for the Structural Analysis of Biological Molecules. *Analytical Chemistry,* 64**,** 1027A-1039A.

CRICK, F. H. 1958. On protein synthesis. *Symposia of the Society for Experimental Biology,* 12**,** 138-63.

D'SOUZA, Z., BLACKBURN, J. B., KUDLYK, T., POKROVSKAYA, I. D. & LUPASHIN, V. V. 2019. Defects in COG-Mediated Golgi Trafficking Alter Endo-Lysosomal System in Human Cells. *Frontiers in Cell and Developmental Biology,* 7**,** 118.

D'SOUZA, Z., TAHER, F. S. & LUPASHIN, V. V. 2020. Golgi inCOGnito: From vesicle tethering to human disease. *Biochimica et Biophysica Acta (BBA) - General Subjects,* 1864**,** 129694.

D'ARCY, M. S. 2019. Cell death: a review of the major forms of apoptosis, necrosis and autophagy. *Cell Biology International,* 43**,** 582-592.

DAVIES, S. L., LOVELADY, C. S., GRAINGER, R. K., RACHER, A. J., YOUNG, R. J. & JAMES, D. C. 2013. Functional heterogeneity and heritability in CHO cell populations. *Biotechnology and Bioengineering,* 110**,** 260-274.

DAWSON, J. H. J. & GUILHAUS, M. 1989. Orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry,* 3**,** 155-159.

DAWSON, P. H. 2013. *Quadrupole mass spectrometry and its applications*, Elsevier.

DE GROOT, A. S. & SCOTT, D. W. 2007. Immunogenicity of protein therapeutics. *Trends in Immunology,* 28**,** 482-490.

DE HAAN, N., YANG, S., CIPOLLO, J. & WUHRER, M. 2020. Glycomics studies using sialic acid derivatization and mass spectrometry. *Nature Reviews Chemistry,* 4**,** 229-242.

DELANEY, J. & VOUROS, P. 2001. Liquid chromatography ion trap mass spectrometric analysis of oligosaccharides using permethylated derivatives. *Rapid Communications in Mass Spectrometry,* 15**,** 325-34.

DELL, A., ROGERS, M. E., THOMAS-OATES, J. E., HUCKERBY, T. N., SANDERSON, P. N. & NIEDUSZYNSKI, I. A. 1988. Fast-atom-bombardment mass-spectrometric strategies for sequencing sulphated oligosaccharides. *Carbohydrate Research,* 179**,** 7-19.

DENZEL, A., MOLINARI, M., TRIGUEROS, C., MARTIN, J. E., VELMURGAN, S., BROWN, S., STAMP, G. & OWEN, M. J. 2002. Early postnatal death and motor disorders in mice congenitally deficient in calnexin expression. *Molecular and Cellular Biology,* 22**,** 7398-404.

DERIŠ, H., CINDRIĆ, A., LAUBER, M., PETROVIĆ, T., BIELIK, A., TARON, C. H., VAN WINGERDEN, M., LAUC, G. & TRBOJEVIĆ-AKMAČIĆ, I. 2021. Robustness and repeatability of GlycoWorks RapiFluor-MS IgG N-glycan profiling in a long-term high-throughput glycomic study. *Glycobiology,* 31**,** 1062-1067.

DOLE, M., HINES, R. L., MACK, L. L., MOBLEY, R. C., FERGUSON, L. D. & ALICE, M. B. 1968a. Gas Phase Macroions. *Macromolecules,* 1**,** 96-97.

DOLE, M., MACK, L. L., HINES, R. L., MOBLEY, R. C., FERGUSON, L. D. & ALICE, M. B. 1968b. Molecular Beams of Macroions. *The Journal of Chemical Physics,* 49**,** 2240-2249.

DOMANN, P. J., PARDOS-PARDOS, A. C., FERNANDES, D. L., SPENCER, D. I., RADCLIFFE, C. M., ROYLE, L., DWEK, R. A. & RUDD, P. M. 2007. Separation-based glycoprofiling approaches using fluorescent labels. *Proteomics,* 7 Suppl 1**,** 70-6.

DONDAPATI, S. K., STECH, M., ZEMELLA, A. & KUBICK, S. 2020. Cell-Free Protein Synthesis: A Promising Option for Future Drug Development. *BioDrugs,* 34**,** 327-348.

DONINI, R., HASLAM, S. M. & KONTORAVDI, C. 2021. Glycoengineering Chinese hamster ovary cells: a short history. *Biochemical Society Transactions,* 49**,** 915-931.

DORDAL, M. S., WANG, F. F. & GOLDWASSER, E. 1985. The role of carbohydrate in erythropoietin action. *Endocrinology,* 116**,** 2293-9.

DOUDNA, J. A. & CHARPENTIER, E. 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science,* 346.

DRICU, A., CARLBERG, M., WANG, M. & LARSSON, O. 1997. Inhibition of *N*-linked Glycosylation Using Tunicamycin Causes Cell Death in Malignant Cells: Role of Down-Regulation of the Insulin-like Growth Factor 1 Receptor in Induction of Apoptosis. *Cancer Research,* 57**,** 543-548.

DRÖSCHER, A. 1998. Camillo Golgi and the discovery of the Golgi apparatus. *Histochemistry and Cell Biology,* 109**,** 425-430.

DUKSIN, D. & MAHONEY, W. C. 1982. Relationship of the structure and biological activity of the natural homologues of tunicamycin. *Journal of Biological Chemistry,* 257**,** 3105-9.

DUMONT, J., EUWART, D., MEI, B., ESTES, S. & KSHIRSAGAR, R. 2016. Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Critical Reviews in Biotechnology,* 36**,** 1110-1122.

DUNPHY, W. G., FRIES, E., URBANI, L. J. & ROTHMAN, J. E. 1981. Early and late functions associated with the Golgi apparatus reside in distinct compartments. *Proceedings of the National Academy of Sciences of the United States of America,* 78**,** 7453-7.

DUNPHY, W. G. & ROTHMAN, J. E. 1983. Compartmentation of asparagine-linked oligosaccharide processing in the Golgi apparatus. *Journal of Cell Biology,* 97**,** 270-5.

EDWARDS, E., LIVANOS, M., KRUEGER, A., DELL, A., HASLAM, S. M., MARK SMALES, C. & BRACEWELL, D. G. 2022. Strategies to control therapeutic antibody glycosylation during bioprocessing: Synthesis and separation. *Biotechnology and Bioengineering,* 119**,** 1343-1358.

EHRET, J., ZIMMERMANN, M., EICHHORN, T. & ZIMMER, A. 2019. Impact of cell culture media additives on IgG glycosylation produced in Chinese hamster ovary cells. *Biotechnology and Bioengineering,* 116**,** 816-830.

EHRING, H., KARAS, M. & HILLENKAMP, F. 1992. Role of photoionization and photochemistry in ionization processes of organic molecules and relevance for matrix-assisted laser desorption Ionization mass spectrometry. *Organic Mass Spectrometry,* 27**,** 472-480.

ELBEIN, A. D., SOLF, R., DORLING, P. R. & VOSBECK, K. 1981. Swainsonine: an inhibitor of glycoprotein processing. *Proceedings of the National Academy of Sciences of the United States of America,* 78**,** 7393-7.

ESCHBACH, J. W., EGRIE, J. C., DOWNING, M. R., BROWNE, J. K. & ADAMSON, J. W. 1987. Correction of the anemia of end-stage renal disease with recombinant human erythropoietin. Results of a combined phase I and II clinical trial. *The New England Journal of Medicine,* 316**,** 73-8.

FAIRBANKS, G., STECK, T. L. & WALLACH, D. F. 1971. Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry,* 10**,** 2606-17.

FANG, F., OREND, G., WATANABE, N., HUNTER, T. & RUOSLAHTI, E. 1996. Dependence of Cyclin E-CDK2 Kinase Activity on Cell Anchorage. *Science,* 271**,** 499-502.

FARQUHAR, M. G. & PALADE, G. E. 1981. The Golgi apparatus (complex)-(1954-1981)-from artifact to center stage. *Journal of Cell Biology,* 91**,** 77s-103s.

FENN, J. B., MANN, M., MENG, C. K., WONG, S. F. & WHITEHOUSE, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science,* 246**,** 64-71.

FERRARA, C., GRAU, S., JÄGER, C., SONDERMANN, P., BRÜNKER, P., WALDHAUER, I., HENNIG, M., RUF, A., RUFER, A. C., STIHLE, M., UMAÑA, P. & BENZ, J. 2011. Unique carbohydrate–carbohydrate interactions are required for high affinity binding between FcγRIII and antibodies lacking core fucose. *Proceedings of the National Academy of Sciences,* 108**,** 12669-12674.

FERREIRA, C. R., XIA, Z. J., CLEMENT, A., PARRY, D. A., DAVIDS, M., TAYLAN, F., SHARMA, P., TURGEON, C. T., BLANCO-SANCHEZ, B., NG, B. G., LOGAN, C. V., WOLFE, L. A., SOLOMON, B. D., CHO, M. T., DOUGLAS, G., CARVALHO, D. R., BRATKE, H., HAUG, M. G., PHILLIPS, J. B., WEGNER, J., TIEMEYER, M., AOKI, K., UNDIAGNOSED DISEASES, N., SCOTTISH GENOME, P., NORDGREN, A., HAMMARSJO, A., DUKER, A. L., ROHENA, L., HOVE, H. B., EK, J., ADAMS, D., TIFFT, C. J., ONYEKWELI, T., WEIXEL, T., MACNAMARA, E., RADTKE, K., POWIS, Z., EARL, D., GABRIEL, M., RUSSI, A. H. S., BRICK, L., KOZENKO, M., THAM, E., RAYMOND, K. M., PHILLIPS, J. A., 3RD, TILLER, G. E., WILSON, W. G., HAMID, R., MALICDAN, M. C. V., NISHIMURA, G., GRIGELIONIENE, G., JACKSON, A., WESTERFIELD, M., BOBER, M. B., GAHL, W. A. & FREEZE, H. H. 2018. A Recurrent De Novo Heterozygous COG4 Substitution Leads to Saul-Wilson Syndrome, Disrupted Vesicular Trafficking, and Altered Proteoglycan Glycosylation. *American Journal of Human Genetics,* 103**,** 553-567.

FISHER, P., SPENCER, H., THOMAS-OATES, J., WOOD, A. J. & UNGAR, D. 2019. Modeling Glycan Processing Reveals Golgi-Enzyme Homeostasis upon Trafficking Defects and Cellular Differentiation. *Cell Reports,* 27**,** 1231-1243.e6.

FISHER, P. & UNGAR, D. 2016. Bridging the Gap between Glycosylation and Vesicle Traffic. *Frontiers in Cell and Developmental Biology,* 4**,** 15.

FLYNN, G. C., CHEN, X., LIU, Y. D., SHAH, B. & ZHANG, Z. 2010. Naturally occurring glycan forms of human immunoglobulins G1 and G2. *Molecular Immunology,* 47**,** 2074-2082.

FORTHAL, D. N. 2014. Functions of Antibodies. *Microbiology spectrum,* 2**,** 1-17.

FOULQUIER, F. 2009. COG defects, birth and rise! *Biochimica et Biophysica Acta* 1792**,** 896-902.

FOULQUIER, F., VASILE, E., SCHOLLEN, E., CALLEWAERT, N., RAEMAEKERS, T., QUELHAS, D., JAEKEN, J., MILLS, P., WINCHESTER, B., KRIEGER, M., ANNAERT, W. & MATTHIJS, G. 2006. Conserved oligomeric Golgi complex subunit 1 deficiency reveals a previously uncharacterized congenital disorder of glycosylation type II. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 3764-9.

FRANCO, C. R., ROCHA, H. A., TRINDADE, E. S., SANTOS, I. A., LEITE, E. L., VEIGA, S. S., NADER, H. B. & DIETRICH, C. P. 2001. Heparan sulfate and control of cell division: adhesion and proliferation of mutant CHO-745 cells lacking xylosyl transferase. *Brazilian Journal of Medical and Biological Research,* 34**,** 971-5.

FREEZE, H. H., CHONG, J. X., BAMSHAD, M. J. & NG, B. G. 2014. Solving glycosylation disorders: fundamental approaches reveal complicated pathways. *American Journal of Human Genetics,* 94**,** 161-75.

FRIES, E. & ROTHMAN, J. E. 1980. Transport of vesicular stomatitis virus glycoprotein in a cell-free extract. *Proceedings of the National Academy of Sciences of the United States of America,* 77**,** 3870-4.

FRYE, C., DESHPANDE, R., ESTES, S., FRANCISSEN, K., JOLY, J., LUBINIECKI, A., MUNRO, T., RUSSELL, R., WANG, T. & ANDERSON, K. 2016. Industry view on the relative importance of "clonality" of biopharmaceutical-producing cell lines. *Biologicals,* 44**,** 117-122.

FUKUDA, M., KANNO, E., ISHIBASHI, K. & ITOH, T. 2008. Large Scale Screening for Novel Rab Effectors Reveals Unexpected Broad Rab Binding Specificity. *Molecular & Cellular Proteomics,* 7**,** 1031.

GAUDRY, J.-P., AROD, C., SAUVAGE, C., BUSSO, S., DUPRAZ, P., PANKIEWICZ, R. & ANTONSSON, B. 2008. Purification of the extracellular domain of the membrane protein GlialCAM expressed in HEK and CHO cells and comparison of the glycosylation. *Protein Expression and Purification,* 58**,** 94-102.

GHASEMI, M., TURNBULL, T., SEBASTIAN, S. & KEMPSON, I. 2021. The MTT Assay: Utility, Limitations, Pitfalls, and Interpretation in Bulk and Single-Cell Analysis. *International Journal of Molecular Sciences,* 22.

GIBSON, R., SCHLESINGER, S. & KORNFELD, S. 1979. The nonglycosylated glycoprotein of vesicular stomatitis virus is temperature-sensitive and undergoes intracellular aggregation at elevated temperatures. *Journal of Biological Chemistry,* 254**,** 3600-3607.

GILLESPIE, L., GERSTENBERG, K., ANA-SOSA-BATIZ, F., PARSONS, M., FARRUKEE, R., KRABBE, M., SPANN, K., BROOKS, A., LONDRIGAN, S. & READING, P. 2016. DC-SIGN and L-SIGN Are Attachment Factors That Promote Infection of Target Cells by Human Metapneumovirus in the Presence or Absence of Cellular Glycosaminoglycans. *Journal of Virology,* 90**,** JVI.00537-16.

GILLINGHAM, A. K. & MUNRO, S. 2016. Finding the Golgi: Golgin Coiled-Coil Proteins Show the Way. *Trends in Cell Biology,* 26**,** 399-408.

GIROLDI, L. A. & SCHALKEN, J. A. 1993. Decreased expression of the intercellular adhesion molecule E-cadherin in prostate cancer: biological significance and clinical implications. *Cancer Metastasis Rev,* 12**,** 29-37.

GLICK, B. S., ELSTON, T. & OSTER, G. 1997. A cisternal maturation mechanism can explain the asymmetry of the Golgi stack. *FEBS Letters,* 414**,** 177-81.

GLICK, B. S. & LUINI, A. 2011. Models for Golgi traffic: a critical assessment. *Cold Spring Harbour Perspectives in Biology,* 3**,** a005215.

GOETZE, A. M., LIU, Y. D., ZHANG, Z., SHAH, B., LEE, E., BONDARENKO, P. V. & FLYNN, G. C. 2011. High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology,* 21**,** 949-959.

GOLDWASSER, E. 1976. Erythropoietin. *Blut,* 33**,** 135-140.

GOLDWASSER, E. 1996. Erythropoietin: a somewhat personal history. *Perspectives in Biology and Medicine,* 40**,** 18-32.

GOUDSMIT, S. A. 1948. A Time-of-Flight Mass Spectrometer. *Physical Review,* 74**,** 622-623.

GOVENDER, S., CSIMMA, C., GENANT, H. K., VALENTIN-OPRAN, A., AMIT, Y., ARBEL, R., ARO, H., ATAR, D., BISHAY, M., BÖRNER, M. G., CHIRON, P., CHOONG, P., CINATS, J., COURTENAY, B., FEIBEL, R., GEULETTE, B., GRAVEL, C., HAAS, N., RASCHKE, M., HAMMACHER, E., VAN DER VELDE, D., HARDY, P., HOLT, M., JOSTEN, C., KETTERL, R. L., LINDEQUE, B., LOB, G., MATHEVON, H., MCCOY, G.,

MARSH, D., MILLER, R., MUNTING, E., OEVRE, S., NORDSLETTEN, L., PATEL, A., POHL, A., RENNIE, W., REYNDERS, P., ROMMENS, P. M., RONDIA, J., ROSSOUW, W. C., DANEEL, P. J., RUFF, S., RÜTER, A., SANTAVIRTA, S., SCHILDHAUER, T. A., GEKLE, C., SCHNETTLER, R., SEGAL, D., SEILER, H., SNOWDOWNE, R. B., STAPERT, J., TAGLANG, G., VERDONK, R., VOGELS, L., WECKBACH, A., WENTZENSEN, A. & WISNIEWSKI, T. 2002. Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: a prospective, controlled, randomized study of four hundred and fifty patients. *Journal of Bone and Joint Surgery American Volume,* 84**,** 2123-34.

GRAMER, M. J., ECKBLAD, J. J., DONAHUE, R., BROWN, J., SHULTZ, C., VICKERMAN, K., PRIEM, P., VAN DEN BREMER, E. T., GERRITSEN, J. & VAN BERKEL, P. H. 2011. Modulation of antibody galactosylation through feeding of uridine, manganese chloride, and galactose. *Biotechnology and Bioengineering,* 108**,** 1591-602.

GRANOVSKY, M., FATA, J., PAWLING, J., MULLER, W. J., KHOKHA, R. & DENNIS, J. W. 2000. Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nature Medicine,* 6**,** 306-312.

GRILO, A. L. & MANTALARIS, A. 2019. The Increasingly Human and Profitable Monoclonal Antibody Market. *Trends in Biotechnology,* 37**,** 9-16.

GRUNWALD-GRUBER, C., THADER, A., MARESCH, D., DALIK, T. & ALTMANN, F. 2017. Determination of true ratios of different N-glycan structures in electrospray ionization mass spectrometry. *Analytical and Bioanalytical Chemistry,* 409**,** 2519-2530.

GU, J., ISAJI, T., XU, Q., KARIYA, Y., GU, W., FUKUDA, T. & DU, Y. 2012. Potential roles of N-glycosylation in cell adhesion. *Glycoconjugate Journal,* 29**,** 599-607.

GUILFORD, P., HOPKINS, J., HARRAWAY, J., MCLEOD, M., MCLEOD, N., HARAWIRA, P., TAITE, H., SCOULAR, R., MILLER, A. & REEVE, A. E. 1998. E-cadherin germline mutations in familial gastric cancer. *Nature,* 392**,** 402-5.

GUILHAUS, M., SELBY, D. & MLYNSKI, V. 2000. Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrometry Reviews,* 19**,** 65-107.

GUPTA, S. K. & SHUKLA, P. 2018. Glycosylation control technologies for recombinant therapeutic proteins. *Applied Microbiology and Biotechnology,* 102**,** 10457-10468.

HA, J. Y., CHOU, H. T., UNGAR, D., YIP, C. K., WALZ, T. & HUGHSON, F. M. 2016. Molecular architecture of the complete COG tethering complex. *Nature Structural & Molecular Biology,* 23**,** 758-60.

HA, T. K., KIM, D., KIM, C. L., GRAV, L. M. & LEE, G. M. 2022. Factors affecting the quality of therapeutic proteins in recombinant Chinese hamster ovary cell culture. *Biotechnology Advances,* 54**,** 107831.

HAIJES, H. A., JAEKEN, J., FOULQUIER, F. & VAN HASSELT, P. M. 2018. Hypothesis: lobe A (COG1-4)-CDG causes a more severe phenotype than lobe B (COG5-8)-CDG. *Journal of Medical Genetics,* 55**,** 137-142.

HALL, M. K., WEIDNER, D. A., DAYAL, S. & SCHWALBE, R. A. 2014. Cell surface N-glycans influence the level of functional E-cadherin at the cell-cell border. *FEBS Open Bio,* 4**,** 892-7.

HAM, R. G. 1965. Clonal growth of mammalian cells in a chemically defined, synthetic medium. *Proceedings of the National Academy of Sciences of the United States of America,* 53**,** 288-93.

HAMMER, J. A., 3RD & WU, X. S. 2002. Rabs grab motors: defining the connections between Rab GTPases and motor proteins. *Current Opinion in Cell Biology,* 14**,** 69-75.

HAMMOND, C., BRAAKMAN, I. & HELENIUS, A. 1994. Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 913-7.

HAN, L. & COSTELLO, C. E. 2013. Mass spectrometry of glycans. *Biochemistry. Biokhimiia,* 78**,** 710-720.

HANDLOGTEN, M. W., LEE-O'BRIEN, A., ROY, G., LEVITSKAYA, S. V., VENKAT, R., SINGH, S. & AHUJA, S. 2018. Intracellular response to process optimization and impact on productivity and product aggregates for a high-titer CHO cell process. *Biotechnology and Bioengineering,* 115**,** 126-138.

HANG, Q., ZHOU, Y., HOU, S., ZHANG, D., YANG, X., CHEN, J., BEN, Z., CHENG, C. & SHEN, A. 2013. Asparagine-linked glycosylation of bone morphogenetic protein-2 is required for secretion and osteoblast differentiation. *Glycobiology,* 24**,** 292-304.

HANSON, S. R., CULYBA, E. K., HSU, T. L., WONG, C. H., KELLY, J. W. & POWERS, E. T. 2009. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 3131-6.

HARVEY, D. J. 2011. Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry. *Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences,* 879**,** 1196-225.

HARVEY, D. J. 2018. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update for 2013-2014. *Mass Spectrometry Reviews*.

HASE, S., IKENAKA, T. & MATSUSHIMA, Y. 1981. A highly sensitive method for analyses of sugar moieties of glycoproteins by fluorescence labeling. *Journal of Biochemistry,* 90**,** 407-14.

HAWORTH, W. N. 1925. A Revision of the Structural Formula of Glucose. *Nature,* 116**,** 430-430.

HAYTER, P. M., CURLING, E. M. A., BAINES, A. J., JENKINS, N., SALMON, I., STRANGE, P. G. & BULL, A. T. 1991. Chinese hamster ovary cell growth and interferon production kinetics in stirred batch culture. *Applied Microbiology and Biotechnology,* 34**,** 559-564.

HEFFNER, K. M., WANG, Q., HIZAL, D. B., CAN, O. & BETENBAUGH, M. J. 2018. Glycoengineering of Mammalian Expression Systems on a Cellular Level. *Advanced Biochemical Engineering Biotechnology*.

HIGEL, F., SEIDL, A., SÖRGEL, F. & FRIESS, W. 2016. N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *European Journal of Pharmaceutics and Biopharmaceutics,* 100**,** 94-100.

HILLENKAMP, F. & KARAS, M. 1990. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods in Enzymology,* 193**,** 280-95.

HIPPLE, J. A. & THOMAS, H. A. 1949. A Time-of-Flight Mass Spectrometer with Varying Field. *Physical Review,* 75**,** 1616-1616.

HOSSLER, P., KHATTAK, S. F. & LI, Z. J. 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology,* 19**,** 936-49.

IRIBARNE, J. V. & THOMSON, B. A. 1976. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics,* 64**,** 2287-2294.

ISAJI, T., GU, J., NISHIUCHI, R., ZHAO, Y., TAKAHASHI, M., MIYOSHI, E., HONKE, K., SEKIGUCHI, K. & TANIGUCHI, N. 2004. Introduction of bisecting GlcNAc into integrin alpha5beta1 reduces ligand binding and down-regulates cell adhesion and cell migration. *Journal of Biological Chemistry,* 279**,** 19747-54.

ISAJI, T., SATO, Y., FUKUDA, T. & GU, J. 2009. N-glycosylation of the I-like domain of beta1 integrin is essential for beta1 integrin expression and biological function: identification of the minimal N-glycosylation requirement for alpha5beta1. *Journal of Biological Chemistry,* 284**,** 12207-16.

ISAJI, T., SATO, Y., ZHAO, Y., MIYOSHI, E., WADA, Y., TANIGUCHI, N. & GU, J. 2006. N-glycosylation of the beta-propeller domain of the integrin alpha5 subunit is essential for alpha5beta1 heterodimerization, expression on the cell surface, and its biological function. *Journal of Biological Chemistry,* 281**,** 33258-67.

JACOBS, K., SHOEMAKER, C., RUDERSDORF, R., NEILL, S. D., KAUFMAN, R. J., MUFSON, A., SEEHRA, J., JONES, S. S., HEWICK, R., FRITSCH, E. F. & ET AL. 1985. Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature,* 313**,** 806-10.

JAEKEN, J. 2011. Congenital disorders of glycosylation (CDG): it's (nearly) all in it! *Journal of Inherited Metabolic Disease,* 34**,** 853-8.

JAHN, R. & SCHELLER, R. H. 2006. SNAREs--engines for membrane fusion. *Nature Reviews Molecular Cell Biology,* 7**,** 631-43.

JASKOLLA, T. W. & KARAS, M. 2011. Compelling evidence for Lucky Survivor and gas phase protonation: the unified MALDI analyte protonation mechanism. *Journal of the American Society for Mass Spectrometry,* 22**,** 976-88.

JELKMANN, W. 2007. Erythropoietin after a century of research: younger than ever. *European Journal of Haematology,* 78**,** 183-205.

JELKMANN, W. & LUNDBY, C. 2011. Blood doping and its detection. *Blood,* 118**,** 2395-2404.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science,* 337**,** 816-821.

JOAO, H. C. & DWEK, R. A. 1993. Effects of glycosylation on protein structure and dynamics in ribonuclease B and some of its individual glycoforms. *European Journal of Biochemistry,* 218**,** 239-44.

JONES, P. T., DEAR, P. H., FOOTE, J., NEUBERGER, M. S. & WINTER, G. 1986. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature,* 321**,** 522.

KAILEMIA, M. J., RUHAAK, L. R., LEBRILLA, C. B. & AMSTER, I. J. 2014. Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Analytical Chemistry,* 86**,** 196-212.

KALAY, H., AMBROSINI, M., VAN BERKEL, P. H., PARREN, P. W., VAN KOOYK, Y. & GARCÍA VALLEJO, J. J. 2012. Online nanoliquid chromatography-mass spectrometry and nanofluorescence detection for high-resolution quantitative N-glycan analysis. *Analytical Biochemistry,* 423**,** 153-62.

KANEKO, Y., NIMMERJAHN, F. & RAVETCH, J. V. 2006. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science,* 313**,** 670-3.

KAO, F. T. & PUCK, T. T. 1968. Genetics of somatic mammalian cells, VII. Induction and isolation of nutritional mutants in Chinese hamster cells. *Proceedings of the National Academy of Sciences of the United States of America,* 60**,** 1275-81.

KARAS, M., BACHMANN, D., BAHR, U. & HILLENKAMP, F. 1987. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes,* 78**,** 53-68.

KARAS, M., BACHMANN, D. & HILLENKAMP, F. 1985. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry,* 57**,** 2935-2939.

KARAS, M., GLÜCKMANN, M. & SCHÄFER, J. 2000. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *Journal of Mass Spectrometry,* 35**,** 1-12.

KARAS, M. & HILLENKAMP, F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry,* 60**,** 2299-301.

KAUFMANN, R., HILLENKAMP, F. & WECHSUNG, R. 1979. The laser microprobe mass analyzer (LAMMA): a new instrument for biomedical microprobe analysis. *Medicinal Progress Through Technology,* 6**,** 109-21.

KAWASAKI, N., ITOH, S., HASHII, N., TAKAKURA, D., QIN, Y., HUANG, X. & YAMAGUCHI, T. 2009. The significance of glycosylation analysis in development of biopharmaceuticals. *Biological and Pharmaceutical Bulletin,* 32**,** 796-800.

KAWASAKI, N., OHTA, M., HYUGA, S., HYUGA, M. & HAYAKAWA, T. 2000. Application of Liquid Chromatography/Mass Spectrometry and Liquid Chromatography with Tandem Mass Spectrometry to the Analysis of the Site-Specific Carbohydrate Heterogeneity in Erythropoietin. *Analytical Biochemistry,* 285**,** 82-91.

KELLEHER, D. J., KARAOGLU, D., MANDON, E. C. & GILMORE, R. 2003. Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Molecular Cell,* 12**,** 101-11.

KELLY, E. E., GIORDANO, F., HORGAN, C. P., JOLLIVET, F., RAPOSO, G. & MCCAFFREY, M. W. 2012. Rab30 is required for the morphological integrity of the Golgi apparatus. *Biology of the Cell,* 104**,** 84-101.

KESER, T., PAVIĆ, T., LAUC, G. & GORNIK, O. 2018. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-glycan Analysis. *Frontiers in Chemistry,* 6.

KHAN, S. N. & LANE, J. M. 2004. The use of recombinant human bone morphogenetic protein-2 (rhBMP-2) in orthopaedic applications. *Expert Opinion on Biological Therapy,* 4**,** 741-8.

KILDEGAARD, H. F., FAN, Y., SEN, J. W., LARSEN, B. & ANDERSEN, M. R. 2016. Glycoprofiling effects of media additives on IgG produced by CHO cells in fed-batch bioreactors. *Biotechnology and Bioengineering,* 113**,** 359-66.

KIM, Y. J., BORSIG, L., VARKI, N. M. & VARKI, A. 1998. P-selectin deficiency attenuates tumor growth and metastasis. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 9325-30.

KINGSLEY, D. M., KOZARSKY, K. F., SEGAL, M. & KRIEGER, M. 1986. Three types of low density lipoprotein receptor-deficient mutant have pleiotropic defects in the

synthesis of N-linked, O-linked, and lipid-linked carbohydrate chains. *Journal of Cell Biology,* 102**,** 1576-85.

KITADA, T., MIYOSHI, E., NODA, K., HIGASHIYAMA, S., IHARA, H., MATSUURA, N., HAYASHI, N., KAWATA, S., MATSUZAWA, Y. & TANIGUCHI, N. 2001. The addition of bisecting N-acetylglucosamine residues to E-cadherin down-regulates the tyrosine phosphorylation of beta-catenin. *Journal of Biological Chemistry,* 276**,** 475-80.

KLINGLER, F., MATHIAS, S., SCHNEIDER, H., BUCK, T., RAAB, N., ZEH, N., SHIEH, Y.-W., PFANNSTIEL, J. & OTTE, K. 2021. Unveiling the CHO surfaceome: Identification of cell surface proteins reveals cell aggregation-relevant mechanisms. *Biotechnology and Bioengineering,* 118**,** 3015-3028.

KNOSSOW, M., GAUDIER, M., DOUGLAS, A., BARRÈRE, B., BIZEBARD, T., BARBEY, C., GIGANT, B. & SKEHEL, J. J. 2002. Mechanism of neutralization of influenza virus infectivity by antibodies. *Virology,* 302**,** 294-8.

KO, P., MISAGHI, S., HU, Z., ZHAN, D., TSUKUDA, J., YIM, M., SANFORD, M., SHAW, D., SHIRATORI, M., SNEDECOR, B., LAIRD, M. & SHEN, A. 2018. Probing the importance of clonality: Single cell subcloning of clonally derived CHO cell lines yields widely diverse clones differing in growth, productivity, and product quality. *Biotechnology Progress,* 34**,** 624-634.

KOHLER, G. & MILSTEIN, C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature,* 256**,** 495-7.

KONRAD, M. W., STORRIE, B., GLASER, D. A. & THOMPSON, L. H. 1977. Clonal variation in colony morphology and growth of CHO cells cultured on agar. *Cell,* 10**,** 305-12.

KORNFELD, R. & KORNFELD, S. 1985. Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry,* 54**,** 631-64.

KRANZ, C., NG, B. G., SUN, L., SHARMA, V., EKLUND, E. A., MIURA, Y., UNGAR, D., LUPASHIN, V., WINKEL, R. D., CIPOLLO, J. F., COSTELLO, C. E., LOH, E., HONG, W. & FREEZE, H. H. 2007. COG8 deficiency causes new congenital disorder of glycosylation type IIh. *Human Molecular Genetics,* 16**,** 731-41.

KRAPP, S., MIMURA, Y., JEFFERIS, R., HUBER, R. & SONDERMANN, P. 2003. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *Journal of Molecular Biology,* 325**,** 979-89.

KRIEGER, M., BROWN, M. S. & GOLDSTEIN, J. L. 1981. Isolation of Chinese hamster cell mutants defective in the receptor-mediated endocytosis of low density lipoprotein. *Journal of Molecular Biology,* 150**,** 167-84.

KRUSIUS, T., FINNE, J. & RAUVALA, H. 1978. The Poly(glycosyl) Chains of Glycoproteins. *European Journal of Biochemistry,* 92**,** 289-300.

KUKURUZINSKA, M. A., APEKIN, V., LAMKIN, M. S., HILTZ, A., RODRIGUEZ, A., LIN, C. C., PAZ, M. A. & OPPENHEIM, F. G. 1994. Antisense RNA to the First N-Glycosylation Gene, ALG7, Inhibits Protein N-Glycosylation and Secretion by Xenopus Oocytes. *Biochemical and Biophysical Research Communications,* 198**,** 1248-1254.

KUROGOCHI, M., MORI, M., OSUMI, K., TOJINO, M., SUGAWARA, S.-I., TAKASHIMA, S., HIROSE, Y., TSUKIMURA, W., MIZUNO, M., AMANO, J., MATSUDA, A., TOMITA, M., TAKAYANAGI, A., SHODA, S.-I. & SHIRAI, T. 2015. Glycoengineered Monoclonal Antibodies with Homogeneous Glycan (M3, G0, G2, and A2) Using

a Chemoenzymatic Approach Have Different Affinities for FcγRIIIa and Variable Antibody-Dependent Cellular Cytotoxicity Activities. *PloS one,* 10**,** e0132848-e0132848.

LALONDE, M.-E. & DUROCHER, Y. 2017. Therapeutic glycoprotein production in mammalian cells. *Journal of Biotechnology,* 251**,** 128-140.

LAUBER, M. A., YU, Y.-Q., BROUSMICHE, D. W., HUA, Z., KOZA, S. M., MAGNELLI, P., GUTHRIE, E., TARON, C. H. & FOUNTAIN, K. J. 2015. Rapid Preparation of Released N-Glycans for HILIC Analysis Using a Labeling Reagent that Facilitates Sensitive Fluorescence and ESI-MS Detection. *Analytical Chemistry,* 87**,** 5401-5409.

LAUFMAN, O., FREEZE, H. H., HONG, W. & LEV, S. 2013a. Deficiency of the Cog8 subunit in normal and CDG-derived cells impairs the assembly of the COG and Golgi SNARE complexes. *Traffic,* 14**,** 1065-77.

LAUFMAN, O., HONG, W. & LEV, S. 2013b. The COG complex interacts with multiple Golgi SNAREs and enhances fusogenic assembly of SNARE complexes. *Journal of Cell Science,* 126**,** 1506-16.

LAUFMAN, O., KEDAN, A., HONG, W. & LEV, S. 2009. Direct interaction between the COG complex and the SM protein, Sly1, is required for Golgi SNARE pairing. *EMBO Journal,* 28**,** 2006-17.

LEE, H. S., QI, Y. & IM, W. 2015a. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Scientific Reports,* 5**,** 8926.

LEE, J. S., GRAV, L. M., LEWIS, N. E. & FAUSTRUP KILDEGAARD, H. 2015b. CRISPR/Cas9-mediated genome engineering of CHO cell factories: Application and perspectives. *Biotechnology Journal,* 10**,** 979-994.

LEE, S. J., EVERS, S., ROEDER, D., PARLOW, A. F., RISTELI, J., RISTELI, L., LEE, Y. C., FEIZI, T., LANGEN, H. & NUSSENZWEIG, M. C. 2002. Mannose receptor-mediated regulation of serum glycoprotein homeostasis. *Science,* 295**,** 1898-901.

LEMMON, M. A. & SCHLESSINGER, J. 2010. Cell signaling by receptor tyrosine kinases. *Cell,* 141**,** 1117-34.

LEROY, J. G. 2006. Congenital disorders of N-glycosylation including diseases associated with O- as well as N-glycosylation defects. *Pediatric Research,* 60**,** 643-56.

LI, F., VIJAYASANKARAN, N., SHEN, A. Y., KISS, R. & AMANULLAH, A. 2010. Cell culture processes for monoclonal antibody production. *mAbs,* 2**,** 466-479.

LIAO, M. M. & SUNSTROM, N.-A. 2006. A transient expression vector for recombinant protein production in Chinese hamster ovary cells. *Journal of Chemical Technology & Biotechnology,* 81**,** 82-88.

LIEWEN, H., MEINHOLD-HEERLEIN, I., OLIVEIRA, V., SCHWARZENBACHER, R., LUO, G., WADLE, A., JUNG, M., PFREUNDSCHUH, M. & STENNER-LIEWEN, F. 2005. Characterization of the human GARP (Golgi associated retrograde protein) complex. *Experimental Cell Research,* 306**,** 24-34.

LIN, C.-W., TSAI, M.-H., LI, S.-T., TSAI, T.-I., CHU, K.-C., LIU, Y.-C., LAI, M.-Y., WU, C.-Y., TSENG, Y.-C., SHIVATARE, S. S., WANG, C.-H., CHAO, P., WANG, S.-Y., SHIH, H.-W., ZENG, Y.-F., YOU, T.-H., LIAO, J.-Y., TU, Y.-C., LIN, Y.-S., CHUANG, H.-Y., CHEN, C.-L., TSAI, C.-S., HUANG, C.-C., LIN, N.-H., MA, C., WU, C.-Y. & WONG, C.-H. 2015. A common glycan structure on immunoglobulin G for enhancement of effector functions. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 10611-10616.

LIN, F. K., SUGGS, S., LIN, C. H., BROWNE, J. K., SMALLING, R., EGRIE, J. C., CHEN, K. K., FOX, G. M., MARTIN, F., STABINSKY, Z. & ET AL. 1985. Cloning and expression of the human erythropoietin gene. *Proceedings of the National Academy of Sciences of the United States of America,* 82**,** 7580-4.

LINK, T., BÄCKSTRÖM, M., GRAHAM, R., ESSERS, R., ZÖRNER, K., GÄTGENS, J., BURCHELL, J., TAYLOR-PAPADIMITRIOU, J., HANSSON, G. C. & NOLL, T. 2004. Bioprocess development for the production of a recombinant MUC1 fusion protein expressed by CHO-K1 cells in protein-free medium. *Journal of Biotechnology,* 110**,** 51-62.

LIU, J., VOLK, K. J., KERNS, E. H., KLOHR, S. E., LEE, M. S. & ROSENBERG, I. E. 1993. Structural characterization of glycoprotein digests by microcolumn liquid chromatography-ionspray tandem mass spectrometry. *Journal of Chromatography A,* 632**,** 45-56.

LIU, L. 2015. Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc-Fusion Proteins. *Journal of Pharmaceutical Sciences,* 104**,** 1866-1884.

LIWOSZ, A., LEI, T. & KUKURUZINSKA, M. A. 2006. N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *Journal of Biological Chemistry,* 281**,** 23138-49.

LOH, E. & HONG, W. 2004. The binary interacting network of the conserved oligomeric Golgi tethering complex. *Journal of Biological Chemistry,* 279**,** 24640-8.

LOUIE, S., HALEY, B., MARSHALL, B., HEIDERSBACH, A., YIM, M., BROZYNSKI, M., TANG, D., LAM, C., PETRYNIAK, B., SHAW, D., SHIM, J., MILLER, A., LOWE, J. B., SNEDECOR, B. & MISAGHI, S. 2017. FX knockout CHO hosts can express desired ratios of fucosylated or afucosylated antibodies with high titers and comparable product quality. *Biotechnology and Bioengineering,* 114**,** 632-644.

LU, R.-M., HWANG, Y.-C., LIU, I. J., LEE, C.-C., TSAI, H.-Z., LI, H.-J. & WU, H.-C. 2020. Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science,* 27**,** 1.

LUCOCQ, J. M. & WARREN, G. 1987. Fragmentation and partitioning of the Golgi apparatus during mitosis in HeLa cells. *EMBO Journal,* 6**,** 3239-46.

LUNENFELD, B., BILGER, W., LONGOBARDI, S., ALAM, V., D'HOOGHE, T. & SUNKARA, S. K. 2019. The Development of Gonadotropins for Clinical Use in the Treatment of Infertility. *Frontiers in endocrinology,* 10**,** 429-429.

MALPHETTES, L., FREYVERT, Y., CHANG, J., LIU, P. Q., CHAN, E., MILLER, J. C., ZHOU, Z., NGUYEN, T., TSAI, C., SNOWDEN, A. W., COLLINGWOOD, T. N., GREGORY, P. D. & COST, G. J. 2010. Highly efficient deletion of FUT8 in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. *Biotechnology and Bioengineering,* 106**,** 774-83.

MAMYRIN, B. A., KARATAEV, V. I., SHMIKK, D. V. & ZAGULIN, V. A. 1973. Mass-Reflectron a new nonmagnetic time-of-flight high-resolution mass-spectrometer. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki,* 64**,** 82-89.

MANZA, L. L., STAMER, S. L., HAM, A. J., CODREANU, S. G. & LIEBLER, D. C. 2005. Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics,* 5**,** 1742-5.

MARSHALL, R. D. 1972. Glycoproteins. *Annual Review of Biochemistry,* 41**,** 673-702.

MARTH, J. D. & GREWAL, P. K. 2008. Mammalian glycosylation in immunity. *Nature Reviews Immunology,* 8**,** 874-887.

MCEVER, R. P. 1997. Selectin-carbohydrate interactions during inflammation and metastasis. *Glycoconjucate Journal,* 14**,** 585-91.

MCNEW, J. A., PARLATI, F., FUKUDA, R., JOHNSTON, R. J., PAZ, K., PAUMET, F., SOLLNER, T. H. & ROTHMAN, J. E. 2000. Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature,* 407**,** 153-9.

MECHREF, Y., HU, Y., DESANTOS-GARCIA, J. L., HUSSEIN, A. & TANG, H. 2013. Quantitative glycomics strategies. *Molecular and Cellular Proteomics,* 12**,** 874-84.

MESAELI, N., NAKAMURA, K., ZVARITCH, E., DICKIE, P., DZIAK, E., KRAUSE, K. H., OPAS, M., MACLENNAN, D. H. & MICHALAK, M. 1999. Calreticulin is essential for cardiac development. *Journal of Cell Biology,* 144**,** 857-68.

MESSINA, A., PALMIGIANO, A., ESPOSITO, F., FIUMARA, A., BORDUGO, A., BARONE, R., STURIALE, L., JAEKEN, J. & GAROZZO, D. 2021. HILIC-UPLC-MS for high throughput and isomeric N-glycan separation and characterization in Congenital Disorders Glycosylation and human diseases. *Glycoconjugate Journal,* 38**,** 201-211.

MILLER, V. J., SHARMA, P., KUDLYK, T. A., FROST, L., ROFE, A. P., WATSON, I. J., DUDEN, R., LOWE, M., LUPASHIN, V. V. & UNGAR, D. 2013. Molecular insights into vesicle tethering at the Golgi by the conserved oligomeric Golgi (COG) complex and the golgin TATA element modulatory factor (TMF). *Journal of Biological Chemistry,* 288**,** 4229-40.

MILLER, V. J. & UNGAR, D. 2012. Re'COG'nition at the Golgi. *Traffic,* 13**,** 891-7.

MIMURA, Y., CHURCH, S., GHIRLANDO, R., ASHTON, P. R., DONG, S., GOODALL, M., LUND, J. & JEFFERIS, R. 2000. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Molecular Immunology,* 37**,** 697-706.

MIMURA, Y., KATOH, T., SALDOVA, R., O'FLAHERTY, R., IZUMI, T., MIMURA-KIMURA, Y., UTSUNOMIYA, T., MIZUKAMI, Y., YAMAMOTO, K., MATSUMOTO, T. & RUDD, P. M. 2018. Glycosylation engineering of therapeutic IgG antibodies: challenges for the safety, functionality and efficacy. *Protein Cell,* 9**,** 47-62.

MININ, A. A. 1997. Dispersal of Golgi apparatus in nocodazole-treated fibroblasts is a kinesin-driven process. *Journal of Cell Science,* 110 (Pt 19)**,** 2495-505.

MISURA, K. M., SCHELLER, R. H. & WEIS, W. I. 2000. Three-dimensional structure of the neuronal-Sec1-syntaxin 1a complex. *Nature,* 404**,** 355-62.

MIYAKE, T., KUNG, C. K. & GOLDWASSER, E. 1977. Purification of human erythropoietin. *Journal of Biological Chemistry,* 252**,** 5558-5564.

MIZUSHIMA, T., YAGI, H., TAKEMOTO, E., SHIBATA-KOYAMA, M., ISODA, Y., IIDA, S., MASUDA, K., SATOH, M. & KATO, K. 2011. Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycans. *Genes to Cells,* 16**,** 1071-80.

MOLINARI, M., GALLI, C., VANONI, O., ARNOLD, S. M. & KAUFMAN, R. J. 2005. Persistent Glycoprotein Misfolding Activates the Glucosidase II/UGT1-Driven Calnexin Cycle to Delay Aggregation and Loss of Folding Competence. *Molecular Cell,* 20**,** 503-512.

MORI, K., KUNI-KAMOCHI, R., YAMANE-OHNUKI, N., WAKITANI, M., YAMANO, K., IMAI, H., KANDA, Y., NIWA, R., IIDA, S., UCHIDA, K., SHITARA, K. & SATOH, M. 2004. Engineering Chinese hamster ovary cells to maximize effector function of

produced antibodies using FUT8 siRNA. *Biotechnology and Bioengineering,* 88**,** 901-8.

MUKHOPADHYAY, S., BACHERT, C., SMITH, D. R. & LINSTEDT, A. D. 2010. Manganese-induced trafficking and turnover of the cis-Golgi glycoprotein GPP130. *Molecular Biology of the Cell,* 21**,** 1282-1292.

MULLOY, B., DELL, A., STANLEY, P. & J, H. P. 2015. Structural Analysis of Glycans. *In:* VARKI, A., CUMMINGS, R. D., ESKO, J. D., STANLEY, P., HART, G. W., AEBI, M., DARVILL, A. G., KINOSHITA, T., PACKER, N. H., PRESTEGARD, J. H., SCHNAAR, R. L. & SEEBERGER, P. H. (eds.) *Essentials of Glycobiology.* 3rd ed. Cold Spring Harbor (NY).

NEEDS, P. W. & SELVENDRAN, R. R. 1993. Avoiding oxidative degradation during sodium hydroxide/methyl iodide-mediated carbohydrate methylation in dimethyl sulfoxide. *Carbohydrate Research,* 245**,** 1-10.

NG, B. G., SHARMA, V., SUN, L., LOH, E., HONG, W., TAY, S. K. & FREEZE, H. H. 2011. Identification of the first COG-CDG patient of Indian origin. *Molecular Genetics and Metabolism,* 102**,** 364-7.

NIMMERJAHN, F., ANTHONY, R. M. & RAVETCH, J. V. 2007. Agalactosylated IgG antibodies depend on cellular Fc receptors for in vivo activity. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 8433-8437.

NOBRE, L. T. D. B., VIDAL, A. A. J., ALMEIDA-LIMA, J., OLIVEIRA, R. M., PAREDES-GAMERO, E. J., MEDEIROS, V. P., TRINDADE, E. S., FRANCO, C. R. C., NADER, H. B. & ROCHA, H. A. O. 2013. Fucan effect on CHO cell proliferation and migration. *Carbohydrate Polymers,* 98**,** 224-232.

NORTH, S. J., HUANG, H. H., SUNDARAM, S., JANG-LEE, J., ETIENNE, A. T., TROLLOPE, A., CHALABI, S., DELL, A., STANLEY, P. & HASLAM, S. M. 2010. Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity. *Journal of Biological Chemistry,* 285**,** 5759-75.

NOVO, J. B., TARGINO VALOTA, R. C., MORO, A. M., RAW, I. & HO, P. L. 2014. Adaptation of glucocerebrosidase-producing CHO cells to serum-free suspension culture. *BMC Proceedings,* 8**,** P46.

O'FLAHERTY, R., TRBOJEVIC-AKMACIC, I., GREVILLE, G., RUDD, P. M. & LAUC, G. 2017. The Sweet Spot for Biologics: Recent Advances in Characterization of Biotherapeutic Glycoproteins. *Expert Review of Proteomics*.

O'SHEA, E. K., KLEMM, J. D., KIM, P. S. & ALBER, T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science,* 254**,** 539-44.

OKA, T., UNGAR, D., HUGHSON, F. M. & KRIEGER, M. 2004. The COG and COPI complexes interact to control the abundance of GEARs, a subset of Golgi integral membrane proteins. *Molecular Biology of the Cell,* 15**,** 2423-35.

OKA, T., VASILE, E., PENMAN, M., NOVINA, C. D., DYKXHOORN, D. M., UNGAR, D., HUGHSON, F. M. & KRIEGER, M. 2005. Genetic analysis of the subunit organization and function of the conserved oligomeric golgi (COG) complex: studies of COG5- and COG7-deficient mammalian cells. *Journal of Biological Chemistry,* 280**,** 32736-45.

OURTH, D. D. & MACDONALD, A. B. 1977. Neutralization of tetanus toxin by human and rabbit immunoglobulin classes and subunits. *Immunology,* 33**,** 807-815.

PADMANABHAN, V., LANG, L. L., SONSTEIN, J., KELCH, R. P. & BEITINS, I. Z. 1988. Modulation of serum follicle-stimulating hormone bioactivity and isoform

distribution by estrogenic steroids in normal women and in gonadal dysgenesis. *The Journal of Clinical Endocrinology and Metabolism,* 67**,** 465-73.

PANCHALINGAM, K. M., JUNG, S., ROSENBERG, L. & BEHIE, L. A. 2015. Bioprocessing strategies for the large-scale production of human mesenchymal stem cells: a review. *Stem Cell Research & Therapy,* 6**,** 225-225.

PAREKH, R. B., DWEK, R. A., SUTTON, B. J., FERNANDES, D. L., LEUNG, A., STANWORTH, D., RADEMACHER, T. W., MIZUOCHI, T., TANIGUCHI, T., MATSUTA, K. & ET AL. 1985. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature,* 316**,** 452-7.

PARODI, A. J., MENDELZON, D. H., LEDERKREMER, G. Z. & MARTIN-BARRIENTOS, J. 1984. Evidence that transient glucosylation of protein-linked Man9GlcNAc2, Man8GlcNAc2, and Man7GlcNAc2 occurs in rat liver and Phaseolus vulgaris cells. *Journal of Biological Chemistry,* 259**,** 6351-7.

PARSONS, J. T., HORWITZ, A. R. & SCHWARTZ, M. A. 2010. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. *Nature Reviews Molecular Cell Biology,* 11**,** 633-643.

PASQUALINI, R. & HEMLER, M. E. 1994. Contrasting roles for integrin beta 1 and beta 5 cytoplasmic domains in subcellular localization, cell proliferation, and cell migration. *The Journal of Cell Biology,* 125**,** 447-460.

PAUL, W. & STEINWEDEL, H. 1953. A new mass spectrometer without a magnetic field. *Zeitschrift fuer Naturforschung (West Germany) Divided into Z. Nautrforsch., A, and Z. Naturforsch., B: Anorg. Chem., Org. Chem., Biochem., Biophys.,* 8.

PEANNE, R., LEGRAND, D., DUVET, S., MIR, A.-M., MATTHIJS, G., ROHRER, J. & FOULQUIER, F. 2010. Differential effects of lobe A and lobe B of the Conserved Oligomeric Golgi complex on the stability of β1,4-galactosyltransferase 1 and α2,6-sialyltransferase 1. *Glycobiology,* 21**,** 864-876.

PESCHKE, B., KELLER, C. W., WEBER, P., QUAST, I. & LÜNEMANN, J. D. 2017. Fc-Galactosylation of Human Immunoglobulin Gamma Isotypes Improves C1q Binding and Enhances Complement-Dependent Cytotoxicity. *Frontiers in Immunology,* 8.

PETITJEAN, O., GIRARDI, E., NGONDO, R. P., LUPASHIN, V. & PFEFFER, S. 2020. Genome-Wide CRISPR-Cas9 Screen Reveals the Importance of the Heparan Sulfate Pathway and the Conserved Oligomeric Golgi Complex for Synthetic Double-Stranded RNA Uptake and Sindbis Virus Infection. *mSphere,* 5**,** e00914-20.

PHILIPPIDIS, A. 2018. *The Top 15 Best-Selling Drugs of 2017* [Online]. Genetic Engineering & Biotechnology News. Available: https://www.genengnews.com/lists/the-top-15-best-selling-drugs-of-2017/ [Accessed 08/12 2018].

PINHO, S. S., FIGUEIREDO, J., CABRAL, J., CARVALHO, S., DOURADO, J., MAGALHÃES, A., GÄRTNER, F., MENDONFA, A. M., ISAJI, T., GU, J., CARNEIRO, F., SERUCA, R., TANIGUCHI, N. & REIS, C. A. 2013. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochimica et Biophysica Acta,* 1830**,** 2690-700.

PLUMMER, T. H., JR., ELDER, J. H., ALEXANDER, S., PHELAN, A. W. & TARENTINO, A. L. 1984. Demonstration of peptide:N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. *Journal of Biological Chemistry,* 259**,** 10700-4.

PLUMMER, T. H., JR. & TARENTINO, A. L. 1991. Purification of the oligosaccharide-cleaving enzymes of Flavobacterium meningosepticum. *Glycobiology,* 1**,** 257-63.

PODOS, S. D., REDDY, P., ASHKENAS, J. & KRIEGER, M. 1994. LDLC encodes a brefeldin A-sensitive, peripheral Golgi protein required for normal Golgi function. *Journal of Cell Biology,* 127**,** 679-91.

POKROVSKAYA, I. D., WILLETT, R., SMITH, R. D., MORELLE, W., KUDLYK, T. & LUPASHIN, V. V. 2011. Conserved oligomeric Golgi complex specifically regulates the maintenance of Golgi glycosylation machinery. *Glycobiology,* 21**,** 1554-1569.

POON, B., KHA, T., TRAN, S. & DASS, C. R. 2016. Bone morphogenetic protein-2 and bone therapy: successes and pitfalls. *Journal of Pharmacy and Pharmacology,* 68**,** 139-147.

PRABHU, A., GADRE, R. & GADGIL, M. 2018. Zinc supplementation decreases galactosylation of recombinant IgG in CHO cells. *Applied Microbiology and Biotechnology,* 102**,** 5989-5999.

PRESLEY, J. F., COLE, N. B., SCHROER, T. A., HIRSCHBERG, K., ZAAL, K. J. M. & LIPPINCOTT-SCHWARTZ, J. 1997. ER-to-Golgi transport visualized in living cells. *Nature,* 389**,** 81-85.

PUCK, T. T. 1985. Development of the Chinese hamster ovary (CHO) cell for use in somatic cell genetics. *Molecular Cell Genetics***,** 37-64.

PUCK, T. T., CIECIURA, S. J. & ROBINSON, A. 1958. Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects. *The Journal of Experimental Medicine,* 108**,** 945-956.

RABOUILLE, C., HUI, N., HUNTE, F., KIECKBUSCH, R., BERGER, E. G., WARREN, G. & NILSSON, T. 1995. Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *Journal of Cell Science,* 108 (Pt 4)**,** 1617-27.

RADIONOVA, A., FILIPPOV, I. & DERRICK, P. J. 2016. In pursuit of resolution in time-of-flight mass spectrometry: A historical perspective. *Mass Spectrometry Reviews,* 35**,** 738-757.

RAM, R. J., LI, B. & KAISER, C. A. 2002. Identification of Sec36p, Sec37p, and Sec38p: Components of Yeast Complex That Contains Sec34p and Sec35p. *Molecular Biology of the Cell,* 13**,** 1484-1500.

RAN, F. A., HSU, P. D., LIN, C.-Y., GOOTENBERG, J. S., KONERMANN, S., TREVINO, A., SCOTT, D. A., INOUE, A., MATOBA, S., ZHANG, Y. & ZHANG, F. 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-1389.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols,* 8**,** 2281.

RAYLEIGH, L. 1882. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 14**,** 184-186.

RAYMOND, C., ROBOTHAM, A., SPEARMAN, M., BUTLER, M., KELLY, J. & DUROCHER, Y. 2015. Production of α2,6-sialylated IgG1 in CHO cells. *MAbs,* 7**,** 571-83.

REPORTLINKER. 2018. *Biologics Global Market Opportunities And Strategies To 2021* [Online]. Available: https://www.reportlinker.com/p05482361/Biologics-Global-Market-Opportunities-And-Strategies-To.html [Accessed 08/12 2018].

REUSCH, D., HABERGER, M., FALCK, D., PETER, B., MAIER, B., GASSNER, J., HOOK, M., WAGNER, K., BONNINGTON, L., BULAU, P. & WUHRER, M. 2015. Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles—Part 2: Mass spectrometric methods. *mAbs,* 7**,** 732-742.

REYNDERS, E., FOULQUIER, F., LEAO TELES, E., QUELHAS, D., MORELLE, W., RABOUILLE, C., ANNAERT, W. & MATTHIJS, G. 2009. Golgi function and dysfunction in the first COG4-deficient CDG type II patient. *Human Molecular Genetics,* 18**,** 3244-56.

RIBLETT, A. M., BLOMEN, V. A., JAE, L. T., ALTAMURA, L. A., DOMS, R. W., BRUMMELKAMP, T. R. & WOJCECHOWSKYJ, J. A. 2015. A Haploid Genetic Screen Identifies Heparan Sulfate Proteoglycans Supporting Rift Valley Fever Virus Infection. *Journal of Virology,* 90**,** 1414-1423.

RICHARDS, F. M., MCKEE, S. A., RAJPAR, M. H., COLE, T. R., EVANS, D. G., JANKOWSKI, J. A., MCKEOWN, C., SANDERS, D. S. & MAHER, E. R. 1999. Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. *Human Molecular Genetics,* 8**,** 607-10.

RICHARDSON, B. C., SMITH, R. D., UNGAR, D., NAKAMURA, A., JEFFREY, P. D., LUPASHIN, V. V. & HUGHSON, F. M. 2009. Structural basis for a human glycosylation disorder caused by mutation of the COG4 gene. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 13329-34.

RITA COSTA, A., ELISA RODRIGUES, M., HENRIQUES, M., AZEREDO, J. & OLIVEIRA, R. 2010. Guidelines to cell engineering for monoclonal antibody production. *European Journal of Pharmaceutics and Biopharmaceutics,* 74**,** 127-138.

RITTERSHAUS, E. S. C., REHMANN, M. S., XU, J., HE, Q., HILL, C., SWANBERG, J., BORYS, M. C., LI, Z.-J. & KHETAN, A. 2022. N-1 Perfusion Platform Development Using a Capacitance Probe for Biomanufacturing. *Bioengineering (Basel, Switzerland),* 9**,** 128.

ROCHA, H. A., FRANCO, C. R., TRINDADE, E. S., VEIGA, S. S., LEITE, E. L., NADER, H. B. & DIETRICH, C. P. 2005. Fucan inhibits Chinese hamster ovary cell (CHO) adhesion to fibronectin by binding to the extracellular matrix. *Planta medica,* 71**,** 628-633.

RODRIGUES, M. E., COSTA, A. R., HENRIQUES, M., CUNNAH, P., MELTON, D. W., AZEREDO, J. & OLIVEIRA, R. 2013. Advances and Drawbacks of the Adaptation to Serum-Free Culture of CHO-K1 Cells for Monoclonal Antibody Production. *Applied Biochemistry and Biotechnology,* 169**,** 1279-1291.

ROMANO, S. J. 2005. Selectin antagonists : therapeutic potential in asthma and COPD. *Treatments in Respiratory Medicine,* 4**,** 85-94.

ROSEN, S. D. 2004. Ligands for L-Selectin: Homing, Inflammation, and Beyond. *Annual Review of Immunology,* 22**,** 129-156.

ROSSER, M. P., XIA, W., HARTSELL, S., MCCAMAN, M., ZHU, Y., WANG, S., HARVEY, S., BRINGMANN, P. & COBB, R. R. 2005. Transient transfection of CHO-K1-S using serum-free medium in suspension: A rapid mammalian protein expression system. *Protein Expression and Purification,* 40**,** 237-243.

ROTHMAN, J. E. & ORCI, L. 1990. Movement of proteins through the Golgi stack: a molecular dissection of vesicular transport. *The FASEB Journal,* 4**,** 1460-8.

ROYLE, L., CAMPBELL, M. P., RADCLIFFE, C. M., WHITE, D. M., HARVEY, D. J., ABRAHAMS, J. L., KIM, Y.-G., HENRY, G. W., SHADICK, N. A., WEINBLATT, M. E.,

LEE, D. M., RUDD, P. M. & DWEK, R. A. 2008. HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. *Analytical Biochemistry,* 376**,** 1-12.

RUGGIANO, A., FORESTI, O. & CARVALHO, P. 2014. Quality control: ER-associated degradation: protein quality control and beyond. *The Journal of Cell Biology,* 204**,** 869-879.

RUHAAK, L. R., ZAUNER, G., HUHN, C., BRUGGINK, C., DEELDER, A. M. & WUHRER, M. 2010. Glycan labeling strategies and their use in identification and quantification. *Analytical and Bioanalytical Chemistry,* 397**,** 3457-81.

RUIZ-CANADA, C., KELLEHER, D. J. & GILMORE, R. 2009. Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell,* 136**,** 272-83.

RUOSLAHTI, E. & REED, J. C. 1994. Anchorage dependence, integrins, and apoptosis. *Cell,* 77**,** 477-8.

RUSH, J. S. 2015. Role of Flippases in Protein Glycosylation in the Endoplasmic Reticulum. *Lipid Insights,* 8**,** 45-53.

SANTIAGO, Y., CHAN, E., LIU, P.-Q., ORLANDO, S., ZHANG, L., URNOV, F. D., HOLMES, M. C., GUSCHIN, D., WAITE, A., MILLER, J. C., REBAR, E. J., GREGORY, P. D., KLUG, A. & COLLINGWOOD, T. N. 2008. Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proceedings of the National Academy of Sciences,* 105**,** 5809-5814.

SASAKI, H., BOTHNER, B., DELL, A. & FUKUDA, M. 1987. Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *Journal of Biological Chemistry,* 262**,** 12059-76.

SASAKI, H., OCHI, N., DELL, A. & FUKUDA, M. 1988. Site-specific glycosylation of human recombinant erythropoietin: analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry. *Biochemistry,* 27**,** 8618-26.

SCHWARTZ, M. A. & ASSOIAN, R. K. 2001. Integrins and cell proliferation: regulation of cyclin-dependent kinases via cytoplasmic signaling pathways. *Journal of Cell Science,* 114**,** 2553-2560.

SEALOVER, N. R., DAVIS, A. M., BROOKS, J. K., GEORGE, H. J., KAYSER, K. J. & LIN, N. 2013. Engineering Chinese Hamster Ovary (CHO) cells for producing recombinant proteins with simple glycoforms by zinc-finger nuclease (ZFN)—mediated gene knockout of mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase (Mgat1). *Journal of Biotechnology,* 167**,** 24-32.

SHESTAKOVA, A., ZOLOV, S. & LUPASHIN, V. 2006. COG Complex-Mediated Recycling of Golgi Glycosyltransferases is Essential for Normal Protein Glycosylation. *Traffic,* 7**,** 191-204.

SHIELDS, R. L., LAI, J., KECK, R., O'CONNELL, L. Y., HONG, K., MENG, Y. G., WEIKERT, S. H. & PRESTA, L. G. 2002. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *Journal of Biological Chemistry,* 277**,** 26733-40.

SHINKAWA, T., NAKAMURA, K., YAMANE, N., SHOJI-HOSAKA, E., KANDA, Y., SAKURADA, M., UCHIDA, K., ANAZAWA, H., SATOH, M., YAMASAKI, M., HANAI, N. & SHITARA, K. 2003. The Absence of Fucose but Not the Presence of Galactose or Bisecting N-Acetylglucosamine of Human IgG1 Complex-type Oligosaccharides

Shows the Critical Role of Enhancing Antibody-dependent Cellular Cytotoxicity. *Journal of Biological Chemistry,* 278**,** 3466-3473.

SHIOZAKI, H., OKA, H., INOUE, M., TAMURA, S. & MONDEN, M. 1996. E-cadherin mediated adhesion system in cancer cells. *Cancer,* 77**,** 1605-1613.

SINACORE, M. S., DRAPEAU, D. & ADAMSON, S. R. 2000. Adaptation of mammalian cells to growth in serum-free media. *Molecular Biotechnology,* 15**,** 249-257.

SINCLAIR, A. M. & ELLIOTT, S. 2005. Glycoengineering: The effect of glycosylation on the properties of therapeutic proteins. *Journal of Pharmaceutical Sciences,* 94**,** 1626-1635.

SINHA, S., PIPES, G., TOPP, E. M., BONDARENKO, P. V., TREUHEIT, M. J. & GADGIL, H. S. 2008. Comparison of LC and LC/MS Methods for Quantifying N-Glycosylation in Recombinant IgGs. *Journal of the American Society for Mass Spectrometry,* 19**,** 1643-1654.

SINKA, R., GILLINGHAM, A. K., KONDYLIS, V. & MUNRO, S. 2008. Golgi coiled-coil proteins contain multiple binding sites for Rab family G proteins. *The Journal of Cell Biology,* 183**,** 607-615.

SKEENE, K., WALKER, M., CLARKE, G., BERGSTROM, E., GENEVER, P., UNGAR, D. & THOMAS-OATES, J. 2017. One Filter, One Sample, and the N- and O-Glyco(proteo)me: Toward a System to Study Disorders of Protein Glycosylation. *Analytical Chemistry,* 89**,** 5840-5849.

SLAMON, D. J., CLARK, G. M., WONG, S. G., LEVIN, W. J., ULLRICH, A. & MCGUIRE, W. L. 1987. Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2/neu Oncogene. *Science,* 235**,** 177-182.

SMITH, K. G. C. & CLATWORTHY, M. R. 2010. FcgammaRIIB in autoimmunity and infection: evolutionary and therapeutic implications. *Nature reviews. Immunology,* 10**,** 328-343.

SMITH, R. D. & LUPASHIN, V. V. 2008. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydrate Research,* 343**,** 2024-31.

SMITH, R. D., WILLETT, R., KUDLYK, T., POKROVSKAYA, I., PATON, A. W., PATON, J. C. & LUPASHIN, V. V. 2009. The COG complex, Rab6 and COPI define a novel Golgi retrograde trafficking pathway that is exploited by SubAB toxin. *Traffic (Copenhagen, Denmark),* 10**,** 1502-1517.

SÖLLNER, T., WHITEHEART, S. W., BRUNNER, M., ERDJUMENT-BROMAGE, H., GEROMANOS, S., TEMPST, P. & ROTHMAN, J. E. 1993. SNAP receptors implicated in vesicle targeting and fusion. *Nature,* 362**,** 318-24.

SPIVAK, J. L. & HOGANS, B. B. 1989. The in vivo metabolism of recombinant human erythropoietin in the rat. *Blood,* 73**,** 90-9.

STADLMANN, J., PABST, M., KOLARICH, D., KUNERT, R. & ALTMANN, F. 2008. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics,* 8**,** 2858-2871.

STAHL, P. D. 1990. The macrophage mannose receptor: current status. *American Journal of Respiratory Cell and Molecular Biology,* 2**,** 317-8.

STANLEY, P. 2011. Golgi glycosylation. *Cold Spring Harbor Perspectives in Biology,* 3.

STANLEY, P., TANIGUCHI, N. & AEBI, M. 2015. N-Glycans. *In:* VARKI, A., CUMMINGS, R. D., ESKO, J. D., STANLEY, P., HART, G. W., AEBI, M., DARVILL, A. G., KINOSHITA, T., PACKER, N. H., PRESTEGARD, J. H., SCHNAAR, R. L. & SEEBERGER, P. H. (eds.) *Essentials of Glycobiology.* 3rd ed. Cold Spring Harbor (NY).

STRUPAT, K., KARAS, M. & HILLENKAMP, F. 1991. 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption—ionization mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes,* 111**,** 89-102.

STURIALE, L., BIANCA, S., GAROZZO, D., TERRACCIANO, A., AGOLINI, E., MESSINA, A., PALMIGIANO, A., ESPOSITO, F., BARONE, C., NOVELLI, A., FIUMARA, A., JAEKEN, J. & BARONE, R. 2019. ALG12-CDG: novel glycophenotype insights endorse the molecular defect. *Glycoconjugate Journal,* 36**,** 461-472.

SUMYA, F. T., POKROVSKAYA, I. D. & LUPASHIN, V. 2021. Development and Initial Characterization of Cellular Models for COG Complex-Related CDG-II Diseases. *Frontiers in Genetics,* 12**,** 733048.

SUN, T., LI, C., HAN, L., JIANG, H., XIE, Y., ZHANG, B., QIAN, X., LU, H. & ZHU, J. 2015. Functional knockout of FUT8 in Chinese hamster ovary cells using CRISPR/Cas9 to produce a defucosylated antibody. *Engineering in Life Sciences,* 15**,** 660-666.

SUN, X., TAO, L., YI, L., OUYANG, Y., XU, N., LI, D., LINHARDT, R. J. & ZHANG, Z. 2017. N-glycans released from glycoproteins using a commercial kit and comprehensively analyzed with a hypothetical database. *Journal of Pharmaceutical Analysis,* 7**,** 87-94.

SUTTAPITUGSAKUL, S., XIAO, H., SMEEKENS, J. & WU, R. 2017. Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. *Molecular Biosystems,* 13**,** 2574-2582.

SUTTON, C. W., ONEILL, J. A. & COTTRELL, J. S. 1994. Site-Specific Characterization of Glycoprotein Carbohydrates by Exoglycosidase Digestion and Laser-Desorption Mass Spectrometry. *Analytical Biochemistry,* 218**,** 34-46.

SUVOROVA, E. S., DUDEN, R. & LUPASHIN, V. V. 2002. The Sec34/Sec35p complex, a Ypt1p effector required for retrograde intra-Golgi trafficking, interacts with Golgi SNAREs and COPI vesicle coat proteins. *Journal of Cell Biology,* 157**,** 631-43.

SUVOROVA, E. S., KURTEN, R. C. & LUPASHIN, V. V. 2001. Identification of a human orthologue of Sec34p as a component of the cis-Golgi vesicle tethering machinery. *Journal of Biological Chemistry,* 276**,** 22810-8.

SZTUL, E. & LUPASHIN, V. 2009. Role of vesicle tethering factors in the ER-Golgi membrane traffic. *FEBS Letters,* 583**,** 3770-83.

TANAKA, K., WAKI, H., IDO, Y., AKITA, S., YOSHIDA, Y., YOSHIDA, T. & MATSUO, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry,* 2**,** 151-153.

TARENTINO, A. L. & PLUMMER, T. H. 1994. [4] Enzymatic deglycosylation of asparagine-linked glycans: Purification, properties, and specificity of oligosaccharide-cleaving enzymes from Flavobacterium meningosepticum. *Methods in Enzymology.* Academic Press.

TAYLOR, G. I. 1964. Disintegration of water drops in an electric field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 280**,** 383-397.

TAYLOR, M. E. & DRICKAMER, K. 1993. Structural requirements for high affinity binding of complex ligands by the macrophage mannose receptor. *Journal of Biological Chemistry,* 268**,** 399-404.

TEJWANI, V., ANDERSEN, M. R., NAM, J. H. & SHARFSTEIN, S. T. 2018. Glycoengineering in CHO Cells: Advances in Systems Biology. *Biotechnology Journal,* 13**,** 1700234.

TEMPLETON, N., DEAN, J., REDDY, P. & YOUNG, J. D. 2013. Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture. *Biotechnology and Bioengineering,* 110**,** 2013-24.

TERBUSH, D. R., MAURICE, T., ROTH, D. & NOVICK, P. 1996. The Exocyst is a multiprotein complex required for exocytosis in Saccharomyces cerevisiae. *The EMBO Journal,* 15**,** 6483-94.

THOMAS-OATES, J. E. & DELL, A. 1989. Fast atom bombardment-mass spectrometry strategies for analysing glycoprotein glycans. *Biochemical Society Transactions,* 17**,** 243-245.

THOMSON, B. A. & IRIBARNE, J. V. 1979. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *The Journal of Chemical Physics,* 71**,** 4451-4463.

TOMIYA, N., KURONO, M., ISHIHARA, H., TEJIMA, S., ENDO, S., ARATA, Y. & TAKAHASHI, N. 1987. Structural analysis of N-linked oligosaccharides by a combination of glycopeptidase, exoglycosidases, and high-performance liquid chromatography. *Analytical Biochemistry,* 163**,** 489-499.

TRETTER, V., ALTMANN, F. & MÄRZ, L. 1991. Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached alpha 1-3 to the asparagine-linked N-acetylglucosamine residue. *European Journal of Biochemistry,* 199**,** 647-52.

UCHIMURA, K. & ROSEN, S. D. 2006. Sulfated L-selectin ligands as a therapeutic target in chronic inflammation. *Trends in Immunology,* 27**,** 559-65.

ULLOA-AGUIRRE, A., TIMOSSI, C., DAMIÁN-MATSUMURA, P. & DIAS, J. A. 1999. Role of glycosylation in function of follicle-stimulating hormone. *Endocrine,* 11**,** 205-15.

UMBAS, R., ISAACS, W. B., BRINGUIER, P. P., SCHAAFSMA, H. E., KARTHAUS, H. F., OOSTERHOF, G. O., DEBRUYNE, F. M. & SCHALKEN, J. A. 1994. Decreased E-cadherin expression is associated with poor prognosis in patients with prostate cancer. *Cancer Research,* 54**,** 3929-33.

UMBAS, R., SCHALKEN, J. A., AALDERS, T. W., CARTER, B. S., KARTHAUS, H. F., SCHAAFSMA, H. E., DEBRUYNE, F. M. & ISAACS, W. B. 1992. Expression of the cellular adhesion molecule E-cadherin is reduced or absent in high-grade prostate cancer. *Cancer Research,* 52**,** 5104-9.

UNGAR, D., OKA, T., BRITTLE, E. E., VASILE, E., LUPASHIN, V. V., CHATTERTON, J. E., HEUSER, J. E., KRIEGER, M. & WATERS, M. G. 2002. Characterization of a mammalian Golgi-localized protein complex, COG, that is required for normal Golgi morphology and function. *Journal of Cell Biology,* 157**,** 405-15.

UNGAR, D., OKA, T., KRIEGER, M. & HUGHSON, F. M. 2006. Retrograde transport on the COG railway. *Trends in Cell Biology,* 16**,** 113-20.

UNGAR, D., OKA, T., VASILE, E., KRIEGER, M. & HUGHSON, F. M. 2005. Subunit architecture of the conserved oligomeric Golgi complex. *Journal of Biological Chemistry,* 280**,** 32729-35.

URNOV, F. D., MILLER, J. C., LEE, Y.-L., BEAUSEJOUR, C. M., ROCK, J. M., AUGUSTUS, S., JAMIESON, A. C., PORTEUS, M. H., GREGORY, P. D. & HOLMES, M. C. 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature,* 435**,** 646-651.

VALABREGA, G., MONTEMURRO, F. & AGLIETTA, M. 2007. Trastuzumab: mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer. *Annals of Oncology,* 18**,** 977-984.

VAN DEVENTER, S. J. 1997. Tumour necrosis factor and Crohn's disease. *Gut,* 40**,** 443-448.

VANRHEENEN, S. M., CAO, X., LUPASHIN, V. V., BARLOWE, C. & WATERS, M. G. 1998. Sec35p, a novel peripheral membrane protein, is required for ER to Golgi vesicle docking. *The Journal of Cell Biology,* 141**,** 1107-1119.

VARKI, A. 2016. Biological roles of glycans. *Glycobiology,* 27**,** 3-49.

VASILE, E., OKA, T., ERICSSON, M., NAKAMURA, N. & KRIEGER, M. 2006. IntraGolgi distribution of the Conserved Oligomeric Golgi (COG) complex. *Experimental Cell Research,* 312**,** 3132-41.

VESTAL, M. L., JUHASZ, P. & MARTIN, S. A. 1995. Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry,* 9**,** 1044-1050.

VUORILUOTO, K., JOKINEN, J., KALLIO, K., SALMIVIRTA, M., HEINO, J. & IVASKA, J. 2008. Syndecan-1 supports integrin α2β1-mediated adhesion to collagen. *Experimental Cell Research,* 314**,** 3369-3381.

WADA, Y., AZADI, P., COSTELLO, C. E., DELL, A., DWEK, R. A., GEYER, H., GEYER, R., KAKEHI, K., KARLSSON, N. G., KATO, K., KAWASAKI, N., KHOO, K. H., KIM, S., KONDO, A., LATTOVA, E., MECHREF, Y., MIYOSHI, E., NAKAMURA, K., NARIMATSU, H., NOVOTNY, M. V., PACKER, N. H., PERREAULT, H., PETER-KATALINIC, J., POHLENTZ, G., REINHOLD, V. N., RUDD, P. M., SUZUKI, A. & TANIGUCHI, N. 2007. Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology,* 17**,** 411-22.

WALSH, G. 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology,* 32**,** 992-1000.

WALTER, D. M., PAUL, K. S. & WATERS, M. G. 1998. Purification and characterization of a novel 13 S hetero-oligomeric protein complex that stimulates in vitro Golgi transport. *Journal of Biological Chemistry,* 273**,** 29565-76.

WALTHER, C. G., WHITFIELD, R. & JAMES, D. C. 2016. Importance of Interaction between Integrin and Actin Cytoskeleton in Suspension Adaptation of CHO cells. *Applied Biochemistry and Biotechnology,* 178**,** 1286-1302.

WANG, Q., CHUNG, C. Y., ROSENBERG, J. N., YU, G. & BETENBAUGH, M. J. 2018. Application of the CRISPR/Cas9 Gene Editing Method for Modulating Antibody Fucosylation in CHO Cells. *Methods in Molecular Biology,* 1850**,** 237-257.

WANG, Y. & SEEMANN, J. 2011. Golgi biogenesis. *Cold Spring Harbor Perspectives in Biology,* 3**,** a005330.

WATERS. 2017. *GlycoWorks RapiFluor-MS N-Glycan Kit-96 Sample Care and Use Manual* [Online]. Waters Corporation. Available: https://www.waters.com/webassets/cms/support/docs/715004793en.pdf [Accessed 12 Jun 2022].

WATERS, M. G., CLARY, D. O. & ROTHMAN, J. E. 1992. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. *Journal of Cell Biology,* 118**,** 1015-26.

WAWRZYNCZAK, E. J., CUMBER, A. J., PARNELL, G. D., JONES, P. T. & WINTER, G. 1992. Blood clearance in the rat of a recombinant mouse monoclonal antibody lacking

the N-linked oligosaccharide side chains of the CH2 domains. *Molecular Immunology,* 29**,** 213-20.

WEBER, T., ZEMELMAN, B. V., MCNEW, J. A., WESTERMANN, B., GMACHL, M., PARLATI, F., SÖLLNER, T. H. & ROTHMAN, J. E. 1998. SNAREpins: minimal machinery for membrane fusion. *Cell,* 92**,** 759-72.

WEI, J. H. & SEEMANN, J. 2017. Golgi ribbon disassembly during mitosis, differentiation and disease progression. *Current Opinions in Cell Biology,* 47**,** 43-51.

WEICKHARDT, C., MORITZ, F. & GROTEMEYER, J. 1996. Time-of-flight mass spectrometry: State-of the-art in chemical analysis and molecular science. *Mass Spectrometry Reviews,* 15**,** 139-162.

WEINBLATT, M. E., KEYSTONE, E. C., FURST, D. E., MORELAND, L. W., WEISMAN, M. H., BIRBARA, C. A., TEOH, L. A., FISCHKOFF, S. A. & CHARTASH, E. K. 2003. Adalimumab, a fully human anti–tumor necrosis factor α monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: The ARMADA trial. *Arthritis and Rheumatism,* 48**,** 35-45.

WHITEHOUSE, C. M., DREYER, R. N., YAMASHITA, M. & FENN, J. B. 1985. Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical Chemistry,* 57**,** 675-9.

WHYTE, J. R. C. & MUNRO, S. 2001. The Sec34/35 Golgi Transport Complex Is Related to the Exocyst, Defining a Family of Complexes Involved in Multiple Steps of Membrane Traffic. *Developmental Cell,* 1**,** 527-537.

WILEY, W. C. & MCLAREN, I. H. 1955. Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments,* 26**,** 1150-1157.

WILLETT, R., KUDLYK, T., POKROVSKAYA, I., SCHÖNHERR, R., UNGAR, D., DUDEN, R. & LUPASHIN, V. 2013a. COG complexes form spatial landmarks for distinct SNARE complexes. *Nature Communications,* 4**,** 1553-1553.

WILLETT, R., POKROVSKAYA, I., KUDLYK, T. & LUPASHIN, V. 2014. Multipronged interaction of the COG complex with intracellular membranes. *Cell Logistics,* 4**,** e27888.

WILLETT, R., UNGAR, D. & LUPASHIN, V. 2013b. The Golgi puppet master: COG complex at center stage of membrane trafficking interactions. *Histochemistry and cell biology,* 140**,** 271-283.

WILLIAMS, R. L., GREENE, S. M. & MCPHERSON, A. 1987. The crystal structure of ribonuclease B at 2.5-A resolution. *Journal of Biological Chemistry,* 262**,** 16020-31.

WILM, M. 2011. Principles of Electrospray Ionization. *Molecular & Cellular Proteomics,* 10**,** M111.009407.

WILSON, K. M., JAGGER, A. M., WALKER, M., SEINKMANE, E., FOX, J. M., KRÖGER, R., GENEVER, P. & UNGAR, D. 2018. Glycans modify mesenchymal stem cell differentiation to impact on the function of resulting osteoblasts. *Journal of Cell Science,* 131**,** jcs209452.

WILSON, K. M., THOMAS-OATES, J. E., GENEVER, P. G. & UNGAR, D. 2016. Glycan Profiling Shows Unvaried N-Glycomes in MSC Clones with Distinct Differentiation Potentials. *Frontiers in Cell and Developmental Biology,* 4**,** 52.

WIŚNIEWSKI, J. R., ZIELINSKA, D. F. & MANN, M. 2011. Comparison of ultrafiltration units for proteomic and N-glycoproteomic analysis by the filter-aided sample preparation method. *Analytical Biochemistry,* 410**,** 307-9.

WITKOS, T. M. & LOWE, M. 2015. The Golgin Family of Coiled-Coil Tethering Proteins. *Frontiers in Cell and Developmental Biology,* 3**,** 86.

WITKOS, T. M. & LOWE, M. 2017. Recognition and tethering of transport vesicles at the Golgi apparatus. *Current Opinions in Cell Biology,* 47**,** 16-23.

WONG, M. & MUNRO, S. 2014. Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. *Science,* 346**,** 1256898.

WOOLLARD, K. J. & CHIN-DUSTING, J. 2007. Therapeutic targeting of p-selectin in atherosclerosis. *Inflammation and Allergy Drug Targets,* 6**,** 69-74.

WORMALD, M. R. & DWEK, R. A. 1999. Glycoproteins: glycan presentation and protein-fold stability. *Structure,* 7**,** R155-R160.

WOZNEY, J. M. 1992. The bone morphogenetic protein family and osteogenesis. *Molecular Reproduction and Development,* 32**,** 160-167.

WOZNEY, J. M., ROSEN, V., CELESTE, A. J., MITSOCK, L. M., WHITTERS, M. J., KRIZ, R. W., HEWICK, R. M. & WANG, E. A. 1988. Novel regulators of bone formation: molecular clones and activities. *Science,* 242**,** 1528-34.

WRIGHT, A. & MORRISON, S. L. 1998. Effect of C2-associated carbohydrate structure on Ig effector function: studies with chimeric mouse-human IgG1 antibodies in glycosylation mutants of Chinese hamster ovary cells. *Journal of Immunology,* 160**,** 3393-402.

WU, X., STEET, R. A., BOHOROV, O., BAKKER, J., NEWELL, J., KRIEGER, M., SPAAPEN, L., KORNFELD, S. & FREEZE, H. H. 2004. Mutation of the COG complex subunit gene COG7 causes a lethal congenital disorder. *Nature Medicine,* 10**,** 518-23.

WURM, F. M. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology,* 22**,** 1393-1398.

XIA, Z.-J., ZENG, X.-X. I., TAMBE, M., NG, B. G., DONG, P. D. S. & FREEZE, H. H. 2021. A Dominant Heterozygous Mutation in COG4 Causes Saul-Wilson Syndrome, a Primordial Dwarfism, and Disrupts Zebrafish Development via Wnt Signaling. *Frontiers in Cell and Developmental Biology,* 9**,** 720688-720688.

XU, J., REHMANN, M. S., XU, M., ZHENG, S., HILL, C., HE, Q., BORYS, M. C. & LI, Z. J. 2020. Development of an intensified fed-batch production platform with doubled titers using N-1 perfusion seed for cell culture manufacturing. *Bioresources and Bioprocessing,* 7**,** 17.

XU, X., NAGARAJAN, H., LEWIS, N. E., PAN, S., CAI, Z., LIU, X., CHEN, W., XIE, M., WANG, W., HAMMOND, S., ANDERSEN, M. R., NEFF, N., PASSARELLI, B., KOH, W., FAN, H. C., WANG, J., GUI, Y., LEE, K. H., BETENBAUGH, M. J., QUAKE, S. R., FAMILI, I., PALSSON, B. O. & WANG, J. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nature Biotechnology,* 29**,** 735-41.

YAMAMOTO, M., TAKAHASHI, Y. & TABATA, Y. 2006. Enhanced bone regeneration at a segmental bone defect by controlled release of bone morphogenetic protein-2 from a biodegradable hydrogel. *Tissue Engineering,* 12**,** 1305-11.

YAMANE-OHNUKI, N., KINOSHITA, S., INOUE-URAKUBO, M., KUSUNOKI, M., IIDA, S., NAKANO, R., WAKITANI, M., NIWA, R., SAKURADA, M., UCHIDA, K., SHITARA, K. & SATOH, M. 2004. Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnology and Bioengineering,* 87**,** 614-22.

YORK, W. S., KIEFER, L. L., ALBERSHEIM, P. & DARVILL, A. G. 1990. Oxidation of oligoglycosyl alditols during methylation catalyzed by sodium hydroxide and iodomethane in methyl sulfoxide. *Carbohydrate Research,* 208**,** 175-182.

YOSHIMURA, M., IHARA, Y., MATSUZAWA, Y. & TANIGUCHI, N. 1996. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *Journal of Biological Chemistry,* 271**,** 13811-5.

YOSHIMURA, M., NISHIKAWA, A., IHARA, Y., TANIGUCHI, S. & TANIGUCHI, N. 1995. Suppression of lung metastasis of B16 mouse melanoma by N-acetylglucosaminyltransferase III gene transfection. *Proceedings of the National Academy of Sciences,* 92**,** 8754-8758.

ZAIA, J. 2004. Mass spectrometry of oligosaccharides. *Mass Spectrometry Reviews,* 23**,** 161-227.

ZAMBRANO, E., ZARIÑÁN, T., OLIVARES, A., BARRIOS-DE-TOMASI, J. & ULLOA-AGUIRRE, A. 1999. Receptor binding activity and in vitro biological activity of the human FSH charge isoforms as disclosed by heterologous and homologous assay systems: implications for the structure-function relationship of the FSH variants. *Endocrine,* 10**,** 113-21.

ZELENY, J. 1914. The Electrical Discharge from Liquid Points, and a Hydrostatic Method of Measuring the Electric Intensity at Their Surfaces. *Physical Review,* 3**,** 69-91.

ZELENY, J. 1917. Instability of Electrified Liquid Surfaces. *Physical Review,* 10**,** 1-6.

ZHANG, L., CASTAN, A., STEVENSON, J., CHATZISSAVIDOU, N., VILAPLANA, F. & CHOTTEAU, V. 2019. Combined effects of glycosylation precursors and lactate on the glycoprofile of IgG produced by CHO cells. *Journal of Biotechnology,* 289**,** 71-79.

ZHANG, L., WANG, M., CASTAN, A., HJALMARSSON, H. & CHOTTEAU, V. 2021. Probabilistic model by Bayesian network for the prediction of antibody glycosylation in perfusion and fed-batch cell cultures. *Biotechnology and Bioengineering,* 118**,** 3447-3459.

ZHANG, M., KOSKIE, K., ROSS, J. S., KAYSER, K. J. & CAPLE, M. V. 2010. Enhancing glycoprotein sialylation by targeted gene silencing in mammalian cells. *Biotechnology and Bioengineering,* 105**,** 1094-1105.

ZHENG, K., YARMARKOVICH, M., BANTOG, C., BAYER, R. & PATAPOFF, T. W. 2014. Influence of glycosylation pattern on the molecular properties of monoclonal antibodies. *mAbs,* 6**,** 649-658.

ZOLOV, S. N. & LUPASHIN, V. V. 2005. Cog3p depletion blocks vesicle-mediated Golgi retrograde trafficking in HeLa cells. *The Journal of Cell Biology,* 168**,** 747-759.

ZUPKE, C., BRADY, L. J., SLADE, P. G., CLARK, P., CASPARY, R. G., LIVINGSTON, B., TAYLOR, L., BIGHAM, K., MORRIS, A. E. & BAILEY, R. W. 2015. Real-time product attribute control to manufacture antibodies with defined N-linked glycan levels. *Biotechnology Progress,* 31**,** 1433-1441.

Appendix 1

Herceptin heavy chain amino acid sequence

EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADSVKGRF
TISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSS
KSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNV
NHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE
VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKL
TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Appendix 1 | Amino acid sequence for Herceptin heavy chain, expressed by the Cobra biologics plasmid used in chapter 4.