IMPERIAL COLLEGE LONDON

Assessment of the Interactions Between Bioprocess Conditions and Protein Glycosylation in Antibody-Producing Mammalian Cell Cultures

Jimenez Del Val, Ioscani

December 2012

Department of Chemical Engineering and Chemical Technology

Imperial College London

South Kensington Campus

London SW7 2AZ

A thesis submitted to Imperial College London in partial fulfilment of the requirements of the degree of Doctor of Philosophy

[…] “I have a new love for that glittering instrument, the human soul. It is a lovely and unique thing in the universe. It is always attacked and never destroyed – because *Thou mayest*.”

Abstract

The pharmaceutical industry is going through a rather turbulent period. Many blockbuster drugs have fallen off patent over the past two years and many more are expected to do so in the near future. In response, pharmaceutical companies have continued searching for products that will replace those that have lost patent protection. However, drug development and approval is extremely time-consuming and costly. So that this critical issue is addressed, industry experts and regulatory agencies have jointly proposed the implementation of Quality by Design (QbD) principles in the development and manufacture of all new drugs. Adoption of QbD is expected to reduce drug development cost and approval time. It is also expected to encourage innovation by developing drugs, and the processes used to manufacture them, around the mechanisms that relate process inputs with end product quality. Within this context, monoclonal antibodies (mAbs) are currently the highest-selling products of the biopharmaceutical industry and are projected to account for nearly half of the world's top-selling drugs by 2018. All currently commercialized mAbs contain N-linked glycans (complex carbohydrates) bound to their protein backbone. These carbohydrates, in turn, have been widely reported to impact the safety and efficacy of mAbs. Furthermore, it has widely been reported that bioprocess conditions heavily impact the composition and distribution of these glycans. For these reasons, mAb glycosylation is considered a critical quality attribute (CQA) of these therapeutic proteins under the QbD scope.

Based on QbD principles, the objective of this thesis was to generate a mathematical model that mechanistically relates the effect of nutrient availability throughout cell culture with the glycan profile of a mAb. The model was constructed from three individual ones. The first model describes the N-linked glycosylation process which occurs in the Golgi apparatus. The second model is unstructured and describes cell culture dynamics. The third and final model describes the biosynthetic pathway for nucleotide sugars. All three models were developed independently, but were adapted with features so that they could be interconnected. The glycosylation model approximates the Golgi apparatus to a single plug flow reactor where resident proteins (glycosylation enzymes and transport proteins) are recycled from distal portions of the Golgi space to proximal ones. Optimisation-based methods were developed to estimate unknown parameters of the model. The cell culture dynamics model was developed to represent cell growth, nutrient consumption and mAb synthesis. It was originally based on Monod kinetics, but was adapted to include experimentally-encountered complexity. The model for nucleotide metabolism was heuristically reduced from 35 constituting reactions to 7. Additional mechanistic features were adapted or included to ensure model fidelity.

Experimentally, batch cultures were performed with hybridoma (CRL-1606 from ATCC). Data for viable cell density, glucose, glutamine, lactate, ammonia and mAb titre were collected. Intracellular samples were produced by perchloric acid extraction. These samples were then analysed for nucleotide sugar content using a high performance anion exchange chromatographic method which was optimized to quantify eight nucleotide sugars and four nucleotides in 30min. mAb bound glycans were analysed by MALDI mass spectrometry. The experimental data was used to estimate the unknown parameters of the models.

The models – along with their associated parameters – were then combined to produce a coupled model that mechanistically relates nutrient availability with mAb glycosylation-associated quality. With further validation, such a model could be used for bioprocess design, control and optimization.

Acknowledgements

I would first like to thank my supervisor, Dr. Cleo Kontoravdi, for all her support, knowledge and patience throughout the past four years. I extend my full gratitude to the Mexican Council for Science and Technology (CONACYT), the Mario Molina Fund, the Mexican Ministry of Education and the BBSRC for their generous funding over the past five years. I would also like to thank Prof. Anne Dell and Dr. Stuart Haslam for their knowledge and support with the mass spectrometry work presented in this thesis. A special thank you to Dr. Antony Constantinou, Dr. Poh-Choo Pang and Dr. Kevin Canis for their practical help on the mass spectrometry. I would also like to show my gratitude to Dr. Karen Polizzi for her invaluable advice and support. I thank Philip Jędrzejewski, Carmen Huang and Tomer Lapidot for their work on the nucleotide sugar pathway modelling, Sarantos Kyriakopoulos for his patience and help with the HPLC, Kate Royle for proofreading sections of this thesis and all the people from the Bio-Systems Engineering group. A special thanks to Prof. Eduardo Bárzana for being a true mentor and friend. Finally, I would like to thank Prof. Mario J. Molina for being an inspiration.

I would like to dedicate this thesis to my family.

Declaration

I hereby certify that all material in this thesis which is not my own work has been properly acknowledged.

Ioscani Jiménez del Val

London, U.K.

Thursday 13 December, 2012

Notation

- OS_i : Oligosaccharide chain *i* (Acceptor)
- Pc_k : Nucleotide-sugar precursor *k* (Donor)
- B_k : By-product *k*. It is the nucleotide component of the NSDs (UDP, GDP or CMP)
- F_{in} : Flowrate of oligosaccharide 1 into the Golgi cisterna
- $F_{k,in}$: Flowrate of NSD *k* into the Golgi cisterna
- q : Volumetric flowrate into the Golgi cisterna
- D_i : Internal diameter of the Golgi space
- $\nu_{i,j}$: Stoichiometric coefficient of oligosaccharide *i* in reaction *j*
- $\nu_{k,j}$: Stoichiometric coefficient of NSD *k* in reaction *j*
- r_j : Rate of reaction *j*
- $E_j(z)$: Concentration of enzyme *j* along Golgi length, *z*
- $E_{j,MAX}$: Peak concentration of enzyme *j*
- $z_{j,MAX}$: Localization of the peak concentration of enzyme *j*
- ω_j : Width of the concentration profile of enzyme *j*
- $TP_k(z)$: Concentration of TP *k* along Golgi length, *z*
- $TP_{k,MAX}$: Peak concentration of TP *k*
- $z_{k,MAX}$: Localization of the peak concentration of TP *k*
- ω_k : Width of the concentration profile of TP *k*
- $k_{f,j}$: Rate-limiting turnover rate
- $K_{d,i}$: Dissociation constant of the Acceptor-Enzyme complex
- OS_z : Competing oligosaccharide substrates for the same enzyme (Competitor)
- $K_{d,z}$: Dissociation constant of the Competitor-Enzyme complex
- OS_{i-1} : Oligosaccharide product in the mannosidase reactions
- $K_{d,k}$: Dissociation constant of the Donor-Enzyme complex
- $K_{d,Bk}$: Dissociation constant of the BP-Enzyme complex
- OS_{i+1} : Oligosaccharide chain with additional sugar residue added (Product)
- $K_{d,k}$: Dissociation constant of the Donor-Enzyme complex

- $K_{d,i+1}$: Dissociation constant of the Enzyme-Product complex
- $k_{T,k}$: Transport turnover rate
- A_G : Surface area of the Golgi cisterna
- V_G : Volume of the Golgi cisterna
- Pc_k^{Cyt} : Nucleotide-sugar precursor *k* concentration on the cytosolic side of Golgi
- K_{Pc}^{Cyt} : Dissociation constant of the transport protein-NSD complex
- $BP_k^{G.L.}$: By-product *k* in the Golgi lumen (same as B_k)
- $K_{BP}^{G.L.}$: Dissociation constant of the transport protein-byproduct complex
- *ManI*: α-1,2 Mannosidase I
- *GnTI*: α-1,3 N-acetylglucosaminyl transferase I
- *ManII*: α-1,3/α-1,6 Mannosidase II
- *GnTII*: α-1,6 N-acetylglucosaminyl transferase II
- *FucT*: α-1,6 Fucosyltransferase (also known as FUT8)
- *GalT*: β-1,4 Galactosyltransferase
- *SiaT*: α-1,6 Sialyltransferase
- TP_1: transport protein for UDP-GlcNAc/UMP
- TP_2: transport protein for GDP-Fuc/GMP
- TP_3: Transport protein for UDP-Gal/UMP
- TP_4: Transport protein for CMP-Neu5Ac/CMP
- X_v: Viable cell density
- X_d: Dead cell density
- Glc_{ext}: Extracellular glucose concentration
- Gln_{ext}: Extracellular glutamine concentration
- *Lact*: Extracellular lactate concentration
- *Amm*: Extracellular ammonia concentration
- *Prod*: Extracellular product (mAb) concentration
- *IVC*: Integral viable cell concentration
- μ: Specific growth rate
- k_d: Specific death rate
- q_{Glc}: Specific glucose consumption rate
- q_{Gln}: Specific glucose consumption rate
- q_{Lact}: Specific lactate production rate
- q_{Amm}: Specific ammonia production rate
- q_{Prod}: Specific product (mAb) production rate
- $F_{Glc,in}$: Molar flowrate of glucose into the cell

- $F_{Gln,in}$: Molar flowrate of glutamine into the cell
- $r_{met,Glc}$: Rate of intracellular glucose consumption for metabolism
- $r_{met,Gln}$: Rate of intracellular glutamine consumption for metabolism
- Glc_{int}: Intracellular glucose concentration
- Gln_{int}: Intracellular glutamine concentration
- *UDP-GlcNAc*: Intracellular uridine diphosphate N-acetylglucosamine concentration
- *UDP-Glc*: Intracellular concentration of uridine diphosphate glucose concentration
- *GDP-Man*: Intracellular concentration of guanosine diphosphate mannose
- *UDP-GalNAc*: Intracellular uridine diphosphate N-acetylgalactosamine concentration
- *CMP-Neu5Ac*: Intracellular cytosine monophosphate N-acetylneuraminic acid concentration
- *UDP-Gal*: Intracellular uridine diphosphate galactose concentration
- *GDP-Fuc*: Intracellular guanosine diphosphate fucose concentration
- $r_{UDPGlcNAc}$: UDP-GlcNAc synthesis rate from glucose and glutamine
- r_{UDPGlc}: UDP-Glc synthesis rate from glucose
- r_{GDPMan}: GDP-Man synthesis rate from glucose
- $r_{UDPGalNAc}$: UDP-GalNAc synthesis rate from UDP-GlcNAc
- $r_{CMPNeu5Ac}$: CMP-Neu5Ac synthesis from UDP-GlcNAc
- r_{UDPGal}: UDP-Gal synthesis rate from UDP-Glc
- r_{GDPFuc}: GDP-Fuc synthesis rate from GDP-Man
- μ_{max}: Maximum specific growth rate
- $k_{d,max,glc}$: Maximum specific death rate for glucose depletion
- $k_{d,max,gln}$: Maximum specific death rate for glutamine depletion
- $K_{M,glc}$; Monod constant for glucose
- $K_{M,gln}$; Monod constant for glutamine
- $K_{d,glc}$; Inverse saturation constant for glucose depletion
- $K_{d,gln}$; Inverse saturation constant for glutamine depletion
- $Y_{X/Glc}$: Yield coefficient of viable cells from glucose
- $Y_{X/Gln}$: Yield coefficient of viable cells from glutamine
- $Y_{Lact/Glc}$: Yield coefficient of lactate from glucose
- $Y_{Amm/Gln}$: Yield coefficient of ammonia from glutamine
- $Y_{mAb/Xv}$: Yield coefficient of mAb from viable cells
- k_T^{Glc}: Glucose transport rate coefficient (from extracellular to intracellular)
- k_T^{Gln}: Glutamine transport rate coefficient (from extracellular to intracellular)
- $k_{met,max}^{Glc}$: Maximum glucose consumption rate for metabolism
- $k_{met,max}^{Gln}$: Maximum glutamine consumption rate for metabolism

- $K_{M,met}^{Glc}$: Menten coefficient for metabolic glucose consumption
- $K_{M,met}^{Gln}$: Menten coefficient for metabolic glutamine consumption
- $k_f^{UDPGlcNAc}$: Turnover rate for UDP-GlcNAc synthesis
- k_f^{UDPGlc}: Turnover rate for UDP-Glc synthesis
- k_f^{GDPMan}: Turnover rate for GDP-Man synthesis
- $k_f^{UDPGalNAc}$: Turnover rate for UDP-GalNAc synthesis
- $k_f^{CMPNeu5Ac}$: Turnover rate for CMP-Neu5Ac synthesis
- k_f^{UDPGal}: Turnover rate for UDP-Gal synthesis
- k_f^{GDPFuc}: Turnover rate for GDP-Fuc synthesis
- $K_{M,glc}^{UDPGlcNAc}$: Glucose Menten coefficient for UDP-GlcNAc synthesis
- $K_{M,gln}^{UDPGlcNAc}$: Glutamine Menten coefficient for UDP-GlcNAc synthesis
- K_M^{UDPGlc}: Menten coefficient for UDP-Glc synthesis
- K_M^{GDPMan}: Menten coefficient for GDP-Man synthesis
- $K_M^{UDPGalNAc}$: Menten coefficient for UDP-GalNAc synthesis
- $K_M^{CMPNeu5Ac}$: Menten coefficient for CMP-Neu5Ac synthesis
- K_M^{UDPGal}: Menten coefficient for UDP-Gal synthesis
- K_M^{GDPFuc}: Menten coefficient for GDP-Fuc synthesis
- $K_i^{CMPNeu5Ac}$: Inhibition coefficient for CMP-Neu5Ac synthesis from UDP-GlcNAc
- K_i^{GDPFuc}: Inhibition coefficient for GDP-Fuc synthesis from GDP-Man
- $K_{i,T}^{CMPNeu5Ac}$: Inhibition coefficient for CMP-Neu5Ac transport from UDP-GlcNAc
- $k_{Cons,max}^{UDPGalNAc}$: Maximum consumption rate of UDP-GalNAc
- $K_{M,Cons}^{UDPGalNAc}$: Menten coefficient for UDP-GalNAc consumption
- $K_{UDPGlcNAc,glyc}$: Saturation coefficient for UDP-GlcNAc consumption in glycosylation
- $K_{UDPGlc,glyc}$: Saturation coefficient for UDP-Glc consumption in glycosylation
- $K_{GDPMan,glyc}$: Saturation coefficient for GDP-Man consumption in glycosylation
- $K_{CMPNeu5Ac,glyc}$: Saturation coefficient for CMP-Neu5Ac consumption in glycosylation
- $K_{UDPGal,glyc}$: Saturation coefficient for UDP-Gal consumption in glycosylation
- $K_{GDPFuc,glyc}$: Saturation coefficient for GDP-Fuc consumption in glycosylation

Contents

Abstract ... iii

Acknowledgements ... iv

Declaration ... vi

Notation ... vii

Contents ... xi

List of figures ... xiv

List of tables ... xvi

Chapter 1: Introduction ... 1

Chapter 2: Literature Review ... 6

2.1. Pharmaceutical Quality by Design ... 6

2.2. Monoclonal antibodies: structure and function ... 11

2.3. Protein Glycosylation ... 16

2.4. Effects of glycosylation on mAb structure and function ... 20

2.5. Effects of bioprocess conditions on protein glycosylation ... 23

2.6. Previous mathematical models for protein glycosylation ... 35

2.7. Techniques for mAb glycosylation analysis ... 39

Chapter 3: Golgi N-linked glycosylation model ... 57

3.1. Definition of the mathematical model 57

3.2. Cisternal maturation and material balances 58

3.3. Enzyme kinetics 60

3.4. Nucleotide sugar donor transport 62

3.5. Golgi glycosylation parameter values 63

3.6. Parameter estimation and model simulation 68

3.7. Comparison with previous models 70

3.8. Case study: glycotransferase silencing 74

Chapter 4: Culture dynamics and NSD model 77

4.1. Cell culture dynamics model 77

4.2. NSD metabolism 81

4.3. NSD reaction network and model reduction 82

Chapter 5: Experimental materials and methods 91

5.1. Cell culture 91

5.2. NSD quantification 94

5.3. Intracellular glucose analysis 107

5.4. Mass Spectroscopy for glycan analysis 108

Chapter 6: Results and discussion 114

6.1. Parameter estimation strategies 114

6.2. Batch culture of HFN 7.1 cells 115

6.3. Case study: glutamine-limited culture 124

Chapter 7: Conclusions and future work 128

7.1. Summary of results 128

7.2. Main conclusions 129

7.3. Future work 130

List of publications 133

References 134

Appendix I: Rate expression derivation 153

Appendix II: Transport equations derivation 162

Appendix III: Glycosylation model parameter derivation 165

Appendix IV: Rate of NSD consumption vs. rate of NSD transport into Golgi 166

Appendix V: HPAEC Supporting information 167

List of figures

Figure 1. Comparison of an IgG-mAb with small molecule therapeutics 3

Figure 2. Overall strategy for linking nutrient availability with mAb Fc N-glycosylation 5

Figure 3. Generalised domain organisation in antibody molecules 12

Figure 4. Structure of IgG 14

Figure 5. Cytotoxic effector functions of mAbs[23] 15

Figure 6. Protein N-linked Glycosylation Process[23] 17

Figure 7. Types of N-linked oligosaccharide structures 18

Figure 8. Oligosaccharide processing reactions of the Golgi apparatus[29] 19

Figure 9. Nucleotide sugar metabolism in mammalian cells[66, 67] 20

Figure 10. Common oligosaccharide structures on the Fc of mAbs 21

Figure 11. Cleavage of the peptide backbone of mAbs[29] 40

Figure 12. mAb Fc N-linked glycosylation reaction scheme[225] 58

Figure 13. Cisternal maturation of the Golgi apparatus viewed as a PFR[225] 59

Figure 14. Definition of Golgi resident protein concentration profiles 60

Figure 15. Resident protein profiles along the Golgi Apparatus[225] 66

Figure 16. Calculated spatial and temporal oligosaccharide profiles within Golgi. 70

Figure 17. Comparison of present model with previous ones[225] 71

Figure 18. Calculated vs. experimental oligosaccharide profiles[225] 72

Figure 19. FucT silencing[225] 75

Figure 20. Reduced NSD biosynthetic pathway[23] 84

Figure 21. Chromatogram replicating the method by Tomiya et al.[288] 96

Figure 22. Chromatogram with 3mM NaOH in E1 and E2 96

Figure 23. Representative chromatogram of the optimised method 98

Figure 24. Comparison of PCA and ACN for nucleotide and NSD extraction 103

Figure 25. Comparison of quenching and quenching and washing with 0.9% NaCl 105

Figure 26. Typical HPAEC chromatogram of intracellular nucleotide and NSD pools 106

Figure 27. Reducing SDS-PAGE gel of purified mAbs 109

Figure 28. Typical mass spectrum of mAb-bound N-glycans 112

Figure 29. X_v, X_d, IVCC, viability and [mAb] profiles in batch culture 116

Figure 30. Extracellular glucose, lactate, glutamine and ammonia in batch overgrowth 117

Figure 31. NSD profiles throughout batch culture 119

Figure 32. Intracellular nucleotide profiles throughout batch culture 121

Figure 33. Estimated Golgi glycosylation enzyme profiles 122

Figure 34. Comparison of measured mAb glycan profiles vs. simulation results 123

Figure 35. Cell culture dynamics under glutamine deprivation 125

Figure 36. NSDs under glutamine deprivation 126

Figure 37. mAb glycan profiles under glutamine deprivation 127

List of tables

Table 1. Global market for mAbs in 2011[1, 2, 48] 13

Table 2. Effect of dissolved oxygen tension (DOT) on protein N-glycosylation[29] 26

Table 3. Differences between mAb glycosylation profiles between the CHO and NS0[143] 32

Table 4. Impact of bioprocess inputs on protein N-glycosylation[23] 34

Table 5. Commonly used UV and fluorescent labels for quantification of glycans[29] 47

Table 6. Analytical techniques for mAb glycosylation[29] 50

Table 7. Kinetics of N-linked glycosylation enzymes 61

Table 8. Enzyme turnover and dissociation constants 64

Table 9. Transport protein parameters 65

Table 10. Intracellular, cytosolic and Golgi NSD concentrations 65

Table 11. Estimated dissociation constants with present model[225] 69

Table 12. Estimated parameters with previous models[225] 73

Table 13. Tomiya et al.[288] HPAEC conditions for NSD quantification 95

Table 14. Optimised HPAEC conditions for NSD quantification 97

Table 15. Retention time variation over 50 injections 98

Table 16. Calibration curve slope and linearity 99

Table 17. Optimised HPAEC sensitivity parameters 100

Table 18. Estimated parameters for culture dynamics model 117

Table 19. Estimated parameters for NSD model 118

Table 20. Estimated dissociation constants 123

Chapter 1

Introduction

The market for monoclonal antibodies (mAbs) has expanded remarkably over the past fifteen years. Between 2009 and 2010, it grew approximately 10% and achieved worldwide sales of over $43 billion.[1, 2] As of June 2011, there were 28 mAbs approved in the EU and US, of which 42% are prescribed for the treatment of several types of cancer – including colorectal and breast cancer – and another 32% for autoimmune disorders, such as rheumatoid arthritis, psoriasis, lupus erythematosus and Crohn's disease. The remaining 26% are indicated for other conditions such as osteoporosis, transplant rejection, bacterial infections, asthma and wet age-related macular degeneration.[1, 2]

Commercial success of mAbs is expected to continue, with their sales projected to account for nearly half of the world's top 100 drugs by 2018.[2] Considering this very promising outlook, pharmaceutical companies have continued driving new mAb candidates through the development pipeline, with approximately 350 currently undergoing clinical studies, of which 25 were in phase III trials as of November 2011.[3] Among these are so-called next-generation mAbs such as bi-specific antibodies, antibody-drug conjugates, glycoengineered antibodies and antibody fragments, all of which intend to replicate the success achieved by their first-generation counterparts.[4-6]

Despite the promising trends for mAbs, "big" – or innovative – pharmaceutical companies have been facing an increasingly challenging environment. From 2004 to 2011, large pharmaceutical companies lost approximately $112 billion in sales due to patent expiration of several high-profile drugs, and this trend is expected to continue with an estimated $148 billion to be lost in the next six years.[2] While these losses and a struggling global economy have resulted in a relatively bleak outlook for overall growth of innovative pharmaceutical

companies, the generics market has been growing at a rapid pace, precisely due to patent expirations.[2] In response, innovative pharmaceutical companies continue searching for new products which will replace the blockbuster drugs that have lost patent protection.[2, 7] However, industry analysts argue that the cost and relatively low throughput of drug development and approval are the major obstacles that are preventing this replacement from occurring.[8, 9] Recent estimates put the cost of taking a new drug to market at approximately $1.3 billion,[8, 10, 11] and although figures in this range have been questioned,[12, 13] most reports suggest that approximately 60% of the cost for taking a new drug to market arises from the clinical stages of drug development.[8, 10, 11, 13-15]

Within this context, it has been proposed that two areas are crucial for improving R&D productivity: firstly, reduction of attrition rates, i.e. reducing the probability of products failing at late clinical phases; and secondly, reduction of the time it takes for drug candidates to go through different clinical stages.[8, 9, 16] A fundamental target for reducing attrition rates and development cycle time is product quality assurance. More specifically, identification of the key quality attributes that impact drug safety and efficacy could assist in identifying which drug candidates should not be progressed to further clinical trials, thus aiding in attrition predictability. In addition, understanding the mechanisms by which manufacturing conditions impact product quality would potentially reduce drug development cycle time by allowing process optimization to occur simultaneously with scale-up and clinical trials. Considering this, regulatory agencies and industry experts proposed implementation of the Quality by Design (QbD) paradigm for development of all new drug products in the pipeline. Fundamentally, QbD aims to build quality into the product at every stage of development by making use of quantitative mechanistic relationships between process inputs and the different features that impact product quality.

QbD has been successfully implemented in small molecule therapeutics (SMT) manufacture on several occasions.[17-21] Its implementation, however, has been considerably more challenging in the field of biological pharmaceuticals. This has largely been due to the size and complexity of biological molecules. Figure 1 compares the scale of a monoclonal antibody with three SMTs which are, on average, 500 times smaller than the mAb. Compounding the effect of size is three-dimensional structure. In all cases, the structure of each molecule defines its safety and efficacy as therapeutics, although it is clear that variation in structure is more likely to occur within a large molecule than in smaller ones. In addition, SMTs are generally produced by well-defined and characterised chemical processes, whereas mAbs are produced by mammalian cells in culture, a process that is inherently more complex and is not yet fully understood.

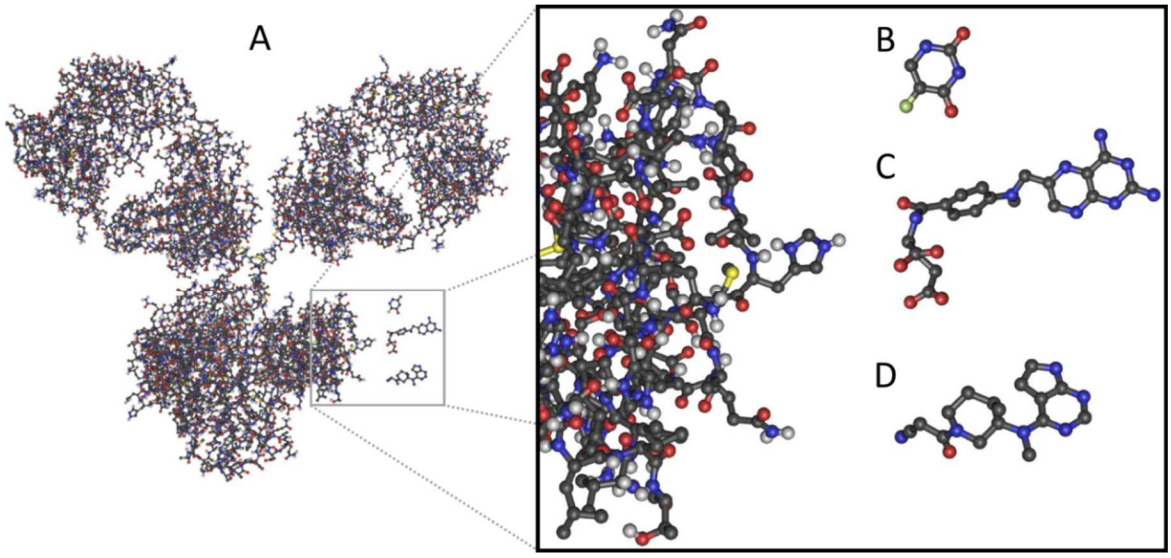


Figure 1. Comparison of an IgG-mAb with small molecule therapeutics
Crystal structure of an IgG1 mAb[22] (A), molecular structure of 3 SMTs which are indicated for autoimmune disorders and cancer: Fluorouracil (B), Methotrexate (C) and Tofacitinib (D).

Molecular complexity of the mAb and incomplete understanding of the bioprocesses employed for its manufacture highlights the importance of elucidating the mechanisms that relate bioprocess conditions with the quality features of mAbs, and thus also underlines the difficulties associated with implementing QbD for the development of these drugs. Knowledge gained on the interplay between bioprocess conditions and mAb quality is currently playing a critical role for the development of new mAbs in the pipeline and is considered of paramount importance in the context of the currently struggling pharmaceutical market.[23]

When considering QbD for mAb development and manufacture, it is necessary to understand which of their molecular features impact their safety and therapeutic efficacy. Broadly, a fully functional mAb consists of a properly folded protein backbone which has been modified by the addition of functional groups and other side-chains. These additions to the peptide backbone are commonly referred to as post-translational modifications (PTMs), and although fragmentation, carboxylation, amidation and sulfation are PTMs commonly associated with currently marketed biologics, glycosylation is considered to be the most prevalent and complex one.[24] More specifically, all commercially available mAbs are based on the gamma immunoglobulin (IgG) isotype, which contains a consensus asparagine-linked (N-linked) glycosylation site on the Cγ2 domains of its crystallizable fragment (Fc). In turn, the nature and distribution of glycan structures bound to this fragment have been reported to greatly influence the immunogenicity, serum half-life and therapeutic mechanism of mAbs,[25-27] thus defining these glycan distributions as critical for mAb quality in the QbD context. Furthermore, it has been widely reported that bioprocess conditions employed for mAb manufacture (e.g. nutrient

availability, dissolved oxygen, pH, temperature, stirring speed and culture medium supplementation) influence the composition and distribution of the glycans bound to mAbs.[28, 29] Despite the wealth of publications describing the mechanisms underlying the influence of certain bioprocess conditions on protein glycosylation, few have been fully quantitative. Finally, authors have suggested that quantitative, mechanistic mathematical models will be central to implementing QbD to biopharmaceutical process development,[29, 30] and the case of mAb glycosylation should be no different.

Given the commercial importance of mAbs, the current climate in the pharmaceutical sector and the ensuing need for quantitative mechanistic mathematical models describing drug quality as a function of manufacturing processes and conditions, with this work we propose the definition of a mathematical model that attempts to mechanistically describe the glycosylation profiles of mAbs as a function of nutrient availability during cell culture.

The first objective, therefore, was to develop a mathematical model that represents the glycosylation process which occurs in the Golgi apparatus. Because none of the three previously published models for Golgi glycosylation have the ability to represent the effect of extracellular culture conditions on the glycosylation process,[31-33] our model includes transport of nucleotide sugars (NSDs) from the cytosol to the Golgi apparatus. NSDs are metabolic products which are synthesized from extracellular nutrients and consumed in the glycosylation reactions in the Golgi apparatus. Accordingly, these metabolites were considered as crucial for drawing mechanistic links between nutrient availability, cell metabolism and mAb glycosylation.

A second mathematical model was developed to describe cell culture dynamics. This model was based on Monod kinetics, but was expanded to include intracellular accumulation of glucose, glutamine and seven nucleotide sugars. The purpose of this model is to mechanistically link extracellular nutrient availability, cell growth, mAb productivity, intracellular NSD availability and the Golgi glycosylation process. A conceptual description of the objectives of this work is presented in Figure 2.

Experimentally, mAb-producing mammalian cells were grown under batch conditions and typical data (viable cell density, extracellular glucose, glutamine, lactate and ammonia; as well as mAb titre) were collected. In addition, methods for extracting and quantifying the intracellular pools of nucleotide sugars were developed and optimised with the aim of obtaining time-profiles for their availability throughout cell culture. Finally, the glycan composition of the produced mAb was monitored under batch operation with MALDI-TOF mass spectrometry. The collected experimental data was subsequently used to estimate the unknown parameters of the mathematical models.

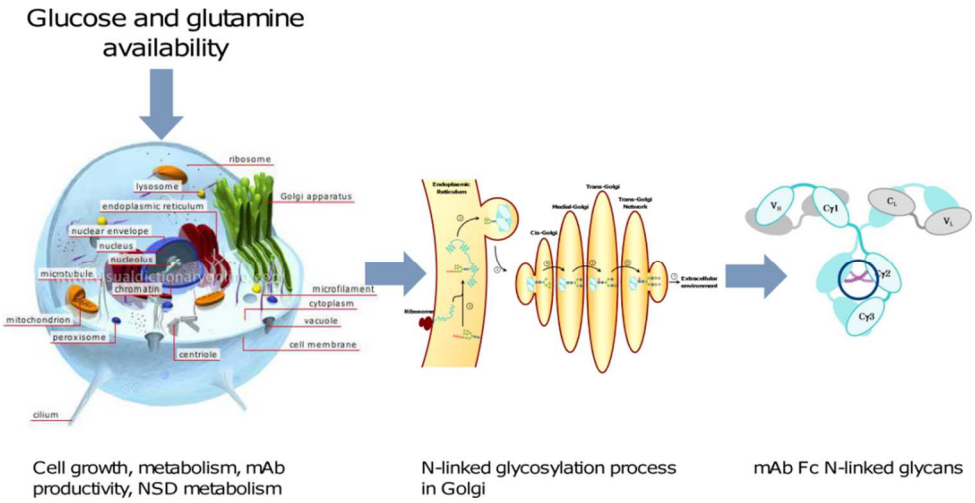


Figure 2. Overall strategy for linking nutrient availability with mAb Fc N-glycosylation
A model that describes cell growth, mAb productivity and NSD metabolism as a function of glucose and glutamine availability will be coupled to a model for N-linked glycosylation in the Golgi apparatus. The coupled model will be able to represent mAb Fc N-glycosylation as a function of glucose and glutamine availability throughout cell culture.

The culmination of this work is a single, combined mathematical model that quantitatively and mechanistically relates cell culture dynamics, intracellular metabolite accumulation and mAb glycosylation. The aim of this coupled model is to assist in predicting mAb glycosylation-associated quality, and thus aid in the development, manufacture and approval of next-generation mAbs under the QbD scope.

This thesis will first present an overview of the current market for mAbs, and the increasingly challenging environment the companies that are producing them are facing. This will be followed by a review of the relevant literature that has described protein N-linked glycosylation, the impact it has on mAb structure and function, and the interplay between bioprocess conditions and protein glycosylation. Subsequently, a description of Quality by Design and the application of its principles in the context of mAb glycosylation-associated quality and bioprocessing will be presented. The thesis will continue with a description of mathematical models, their construction, their underlying assumptions, as well as the optimisation strategies developed to estimate their unknown parameters. A description of the experimental methods employed throughout the project will be presented with special emphasis on the development and optimization of the NSD extraction and quantification protocols that were established through this work. Both modelling and experimental results will then be presented and compared in order to assess and establish the validity of assumptions that were made and experimental procedures that were established or followed. Finally, the thesis will conclude by presenting the main findings of the work along with recommendations for future paths in this line of research.

Chapter 2

Literature Review

2.1. Pharmaceutical Quality by Design

As was discussed in the introduction to this thesis, Quality by Design (QbD) was proposed along with other strategies as a means for pharmaceutical companies to improve their drug development throughput by building quality into the product at every stage of development and manufacture. This section will review the underlying concepts of QbD applied to drug development and manufacture, contrast these concepts with the current state of affairs in the industry and highlight the advantages associated with adopting the QbD paradigm for drug development and manufacture.

2.1.1. Current state of pharmaceutical process development

Drug development begins by identifying molecules that have the biochemical potential to treat illnesses. Pharmaceutical companies then select drug candidates based on their manufacturability and potential profitability for optimization and, eventually, are progressed to preclinical and clinical studies so that their safety and therapeutic efficacy is assessed. Concurrently, manufacturing is scaled up so that sufficient material is available throughout all stages of product development. Also throughout these phases, data concerning drug safety, efficacy and manufacture is reviewed for approval by the corresponding regulatory agencies. Approval requires that the drug product is shown to be produced consistently, that it is safe and finally, that it is efficacious for its indication. Despite decades of advances in drug manufacturing, pharmaceutical process development and approval is still extremely lengthy, highly expensive and uncertain.[34]

Pharmaceutical companies have largely relied on the Quality by Testing (QbT) approach.[34] Under QbT, the ranges for drug substance properties that yield acceptable safety and efficacy, also known as product quality attributes, are associated with a set of process inputs in terms of process parameters and raw materials during the preclinical and clinical stages of development. In turn, process inputs that have empirically shown to yield acceptable product quality are defined and often maintained unchanged after phase II clinical trials so that manufacturers avoid costs associated with additional testing for regulatory compliance.[35] From this point on, the product is tested for quality compliance *after* manufacturing runs with these pre-defined process inputs.[34-37] This approach relies solely on empirical knowledge, excluding mechanistic knowledge of how process inputs are linked with product quality, and thus, uncouples the end-quality of the product from the manufacturing process.

There are several drawbacks of the QbT approach.[29, 35, 38] The first is that development and approval of pharmaceutical processes has historically been extremely time-consuming and costly. It has been estimated that drug approval takes between 7 to 10 years[9, 39] and costs $1.2 to $1.8 bn per approved drug when the risk of failure is included.[8, 40] These issues are correlated and may be attributed to the limited understanding between product quality characteristics, therapeutic mechanisms, and the effect process inputs have on product quality.

Secondly, process control is not established by mechanistic links between process inputs and product quality, rendering the process susceptible to generating off-spec products. In such cases, identification of the source(s) of failure is difficult.

Additionally, the range of process conditions approved under QbT is tightly defined and input disturbances outside this space require validation, which equates to further delays and expense. This discourages pharmaceutical companies from optimising processes or implementing innovative technologies. Finally, manufacturing processes developed and approved under QbT generate a limited amount of knowledge because mechanistic links between process inputs and product quality are largely disregarded. This restricts knowledge transfer between processes and products. Furthermore, this lack of generated mechanistic knowledge increases the likelihood of suboptimal process performance.

Small molecules and protein therapeutics currently undergo the same QbT development and approval process and suffer the same caveats. In the QbT context, bioprocess inputs are empirically defined so that the protein quality characteristics, such as e.g. aggregation, folding, methylation and glycosylation, lie within acceptable ranges for safety and therapeutic efficacy. As occurs with small molecules, QbT disjoins therapeutic protein quality from the bioprocess,

leading to inadequate understanding of the relationship between process inputs and product quality and thus limiting the potential for bioprocess optimization.

The inefficiencies of QbT have led to significant investments in drug discovery rather than process understanding and optimization. This, in turn, has translated into reduced cost-effectiveness, lower number of novel manufacturing processes and an overall drop in number and quality of innovative drugs.[8, 35] The issues discussed above, as well as the need for increasing drug development throughput at lower cost, have led the regulatory bodies and pharmaceutical companies to propose the introduction of Quality by Design principles to the manufacture of all new drugs – including therapeutic proteins – along the development pipeline.[34, 38, 41]

2.1.2. Implementation of QbD to pharmaceutical process development

Pharmaceutical QbD is a conceptual framework for the development and approval of pharmaceutical manufacturing processes that aims to build quality – particularly with respect to safety and therapeutic efficacy – into the product at every stage of process development.[29, 37, 38] Application of QbD principles to process development is outlined in the Process Analytical Technology (PAT) guideline "PAT – A Framework for Innovative Pharmaceutical Manufacturing and Quality Assurance[42]" by the US Federal Drug Administration (USFDA) and in the guidance documents "ICH Q8 Pharmaceutical Development,[37]" "ICH Q9 Quality Risk Management[43]" and "ICH Q10 Pharmaceutical Quality Systems[44]" from the International Conference on Harmonization (ICH), which is an association constituted by the USFDA, the European Medicines Agency (EMA), the Pharmaceuticals and Medical Devices Agency of Japan (PMDA) and several experts from the pharmaceutical industry.

QbD implementation is an information-driven process where all available knowledge on the drug product including, but not limited to, its therapeutic mechanisms, its manufacturing process and potential sources of variability is used to define a range of manufacturing conditions that will ultimately ensure *in vivo* product safety and efficacy. More specifically, the ICH guidelines [37, 43, 45] require the definition of the quality target product profile, identification of the critical quality attributes (CQAs) of the drug product, identification of process inputs that affect product CQAs, Selection of the appropriate manufacturing process, and definition of a control strategy. These steps are discussed in detail below.

Definition of the quality target product profile (QTPP)

The QTPP is the set of quality characteristics that ensure that the drug product is safe and efficacious. Considerations to define the QTPP include the route of administration, dosage form, delivery systems, dosage strength, sterility, purity and stability.[37]

Identification of the critical quality attributes (CQAs) of the drug product

A CQA is defined in ICH Q8 as "a physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality" [37]. As the definition implies, identification of CQAs requires thorough physicochemical and biological characterization of the drug product and in-depth knowledge of the properties that have a higher influence on its safety and efficacy.

Identification of process inputs that affect product CQAs

Once CQAs are identified, it is necessary to determine not only which process inputs (raw materials and process conditions) impact CQAs, but also how these inputs interact to affect the CQAs. It is possible that the required knowledge may not be available for certain processes (particularly novel ones) and should be established through the combination of prior knowledge, mechanistic modelling, experimentation and, finally, a risk assessment so that the influence of material attributes and process conditions on CQAs is ranked according to likelihood and extent of impact. For this purpose, under QbD, there is special emphasis on design of experiments (DoE) so that the interaction between individual process inputs and their impact on product CQAs is taken into consideration. Crucially, all elements involved in this section directly couple manufacturing process conditions with the CQAs in a robust, systematic and information-driven manner.

Selection of the appropriate manufacturing process

The aforementioned risk assessment will dictate a multi-dimensional design space of allowable process input values. ICH Q8 defines the design space as "an established multidimensional combination and interaction of material attributes and/or process parameters demonstrated to provide assurance of quality."[37] Guided by the design space, the manufacturing process conditions that ensure product quality can be identified. The selected process must be robust such that the likelihood of the CQAs being within the range that ensures safety and efficacy is maximised.

Definition of a control strategy

With the appropriate manufacturing process in place, a risk mitigation strategy can be established. This involves monitoring and controlling raw material specifications such that no impact is observed on the drug product's CQAs. In addition, the process parameters that influence the CQAs are controlled, ideally through online measurements and robust control systems, at every stage of the manufacturing process so that the desired product quality is met.

An enabling tool for QbD implementation is process analytical technology (PAT). PAT is defined by the USFDA and the ICH as "a system for designing, analyzing and controlling manufacturing through timely measurements of CQAs and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality.[42]" The elements of PAT concerning process design and control have been presented in the description of QbD above. However, PAT also concerns the "timely measurement of CQAs and performance attributes"[42] throughout the implementation of QbD to manufacturing process development. Experimental methods for accurately identifying and measuring the physical, chemical and biological properties of drug products are essential for defining QTPPs and CQAs early in process development. Further along in the implementation of QbD, it is also necessary to accurately measure the material attributes and process parameters that affect product CQAs. Finally, once the manufacturing process is established, appropriate analytical technologies are necessary to monitor process parameters ideally online (or at line) so that process metrics can be used for process control. Analytical methods therefore constitute a core element throughout process development under the QbD scope.[46]

The three major regulatory bodies (USFDA, EMA and PMDA) are encouraging implementation of the QbD approach for the development of all new drugs in the pipeline.[19, 34, 36, 38] QbD is expected to reduce process approval time and costs, limit regulatory intervention and encourage optimization and innovation by building processes around the mechanistic relationships between inputs and product quality. Because these relationships should be based on mechanistic knowledge, process outputs are expected to be more predictable and to require less regulation. This, in turn, would considerably reduce approval time and development costs. In addition, the more ample design space created through QbD would allow inputs to vary more without the need for additional approval. Predictability would also translate into more robust control systems that would decrease the likelihood of failed batches. Finally, the knowledge generated along with the broader design space and more flexible regulatory approval procedure would encourage process optimization and could potentially contribute to the development of novel processes as well as the discovery and design of next generation drugs.

The QbD framework has already been implemented in the manufacture of small-molecule therapeutics (SMTs) with relative ease. However, the implementation of QbD to protein therapeutics (PTs) has met more resistance. This can be attributed to the fact that the physical and chemical processes underlying the manufacture of SMTs is better understood and the mechanisms relating process inputs with SMT quality are easier to define. In contrast, the cellular mechanisms by which PTs are produced are less well understood, and the structural complexity of PTs makes their isolation, separation, purification and overall quality control significantly more challenging. Despite this, the regulatory agencies and several authors believe that sufficient knowledge is available or can be gained through current experimental and modelling methodologies to elucidate mechanistic relations between bioprocesses and PT quality, thus allowing for implementation of QbD principles to the development of therapeutic protein manufacturing processes in the near future.[34, 38]

Implementation of the QbD paradigm in biologics should increase knowledge on the therapeutic mechanisms of biotherapeutics considerably. This could lead to the improvement of existing products and may contribute to the discovery and development of new biologics. It will also enable more robust process control, process optimization and, potentially, development of novel and efficient manufacturing platforms Implied in this is the considerable reduction in approval times which would heavily reduce costs of product and process development and eventually translate into lower costs for healthcare providers and patients.

2.2. Monoclonal antibodies: structure and function

Antibodies, which are more formally named immunoglobulins (Igs), are the biomolecules that mediate humoral adaptative immune response in vertebrates. Immunoglobulins are glycoproteins that are produced by a type of white blood cells named B-lymphocytes. Being glycoproteins, Igs consist of a polypeptide backbone (protein backbone) that has oligosaccharide structures (glycans) covalently bound to it. Most mammalian Igs are formed by four different polypeptide chains covalently bound together by disulphide bridges. The four constituting chains are two identical light chains and two identical heavy chains. Each chain is composed of separate structural domains. The light chains consist of two domains –a constant domain (C_L), which defines either the kappa (κ) or lambda (λ) isotype, and a variable domain (V_L), which dictates the antibody's affinity for its antigen. The heavy chains also contain a set of constant structural domains (C_H) and a single variable domain (V_H), which serves the same purpose as their V_L counterpart. The constant domains of the heavy chains define five different isotypes, which are IgA, IgD, IgE, IgG and IgM. IgA, IgD, IgE and IgG heavy chains consist of three

constant structural domains, while IgM heavy chains contain four constant domains. Each of these domains, regardless of it being constant or variable, has a three dimensional structure that is defined by intermolecular forces and disulphide bonds. Figure 3 shows the domain structure and intra and inter chain disulphide bonding within immunoglobulins.

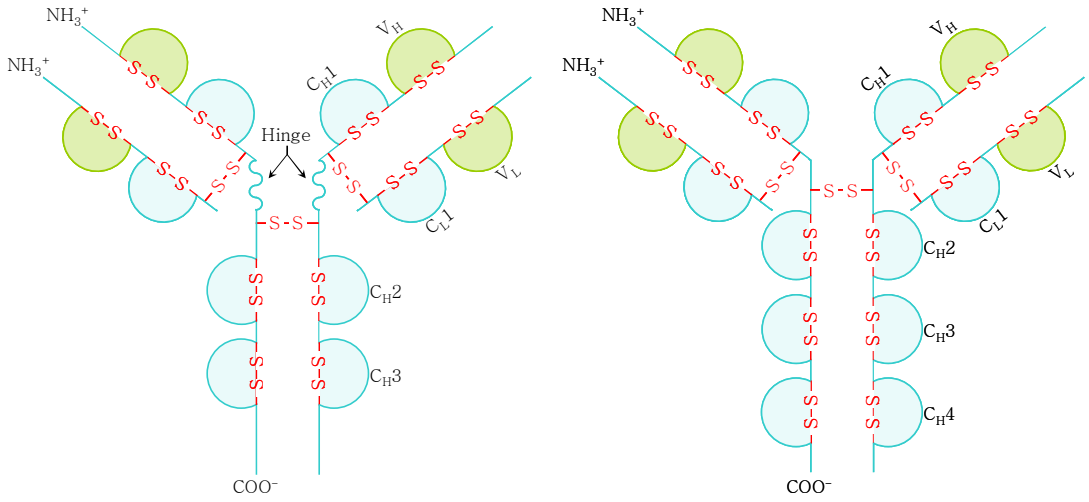


Figure 3. Generalised domain organisation in antibody molecules
On the left we have the generalised domain organisation of antibodies which contain only four heavy chain domains (IgA, IgD and IgG). On the right we have the organisation of IgM, which contains five heavy chain domains and has no hinge region. The sense of the polypeptide chains is represented shown by the NH_4^+ and COO^- annotations. The disulphide bonds which generate each separate domain and which also bind the separate chains to each other are presented in red. The constant domains can be seen in blue and the variable domains are presented in green.

Monoclonal antibodies are immunoglobulins that are produced by the clones of a single cell that has been immortalised (it has the ability to reproduce indefinitely *in vitro*). The original method by which B-lymphocytes were immortalised was hybridisation. The hybridoma technology consists in fusing a B-lymphocyte that produces a desired antibody with a tumorous B-lymphocyte.[47] Since the advent of this technology, focus was set on generating a new class of drugs that could bind, with very high specificity, to pathogenic targets within patients. From the first mAbs that were generated to avoid transplant rejection, a vast number of them are currently approved by regulatory bodies for the treatment of immunological disorders, infectious diseases and cancer. As of June, 2011, 28 mAbs have been approved by the U.S. Food and Drug Administration (USFDA) or the European Medicines Agency (EMEA).[1, 2] Table 1 shows the currently approved mAbs, their antibody isotype, the antigens they bind to and the category of illnesses for which they are currently prescribed.

Table 1 shows that all approved mAbs are based on the IgG isotype. For this reason, this review will focus on the structure of IgG and the mechanisms by which it functions. IgG has the same basic structure shown in Figure 3, but can be seen in more detail in Figure 4. It has two

heavy chains, and two light chains. Its gamma (γ) heavy chain is composed of four structural domains, which are named Cγ1, Cγ2, Cγ3 and V_H. The IgG light chains contain two domains which are named C_L and V_H, and are responsible for antigen binding.

Table 1. Global market for mAbs in 2011[1, 2, 48]

Generic name	US trade name	Company	Antibody Format	Therapeutical Category	2011 Global Sales ($ billion)
Infliximab	Remicade	Centocor	Chimeric, IgG1κ, anti-TNF	Immunological	$ 9.00
Adalimumab	Humira	Abbott	Humanized, IgG1κ, anti-TNFa	Immunological	$ 8.20
Rituximab	Rituxan	Biogen IDEC and Genentech/Roche	Chimeric, IgG1κ, anti-CD20	Cancer	$ 6.80
Bevacizumab	Avastin	Genentech/Roche	Humanized, IgG1κ, anti-VEGF	Cancer	$ 6.00
Trastuzumab	Herceptin	Genentech/Roche	Humanized, IgG1κ, anti-HER2	Cancer	$ 5.90
Ranibizumab	Lucentis	Genentech/Roche	Humanized, IgG1κ Fab, anti-VEGF	Ophthalmic	$ 3.80
Cetuximab	Erbitux	ImClone Systems	Chimeric, IgG1κ, anti-EGF receptor	Cancer	$ 1.90
Natalizumab	Tysabri	Biogen Idec	Humanized, IgG4κ, anti-a4-integrin	Immunological	$ 1.50
Palivizumab	Synagis	MedImmune	Humanized, IgG1κ, anti-RSV	Anti-infective	$ 1.00
Abciximab	ReoPro	Centocor	Chimeric, IgG1κ, anti-GPIIb/IIIa	Hemostasis	
Alemtuzumab	Campath	Genzyme	Humanized, IgG1κ, anti-CD52	Cancer	
Basiliximab	Simulect	Novartis/Cerimon	Chimeric, IgG1κ, anti-CD25	Immunological	
Basiliximab	Simulect	Novartis	Chimeric, IgG1κ, anti-CD25	Immunological	
Belimumab	Benlysta	Human Genome Sciences/GlaxoSmithKline	Human, IgG1λ, anti-BLyS	Immunological	
Canakinumab	Ilaris	Novartis	Human, IgG1κ, anti-IL1β	Immunological	
Catumaxomab	Removab	Fresnius/Trion	Trispecific, mouse IgG2a/rat IgG2b, anti-EpCAM and CD3	Cancer	
Certolizumab pegol	Cimzia	UCB	Humanized, pegylated IgG1κ-derived Fab, anti-TNF-α	Immunological	
Denosumab	Prolia/XGEVA	Amgen/GlaxoSmithKline/Daiichi-Sankyo/AstraZeneca	Human, IgG2κ, anti-RANKL	Cancer	
Eculizumab	Soliris	Alexion	Humanized, IgG2/4κ, anti-complement-5	Immunological	$ 3.60
Golimumab	Simponi	Johnson & Johnson/Merck & Co./Mitsubishi Tanabe Pharma	Human, IgG1κ, anti-TNF-α	Immunological	
Ibritumomab tiuxetan	Zevalin	Biogen Idec	Murine, IgG1κ, anti-CD20 radiolabeled (Y-90)	Cancer	
Ipilimumab	Yervoy	Bristo-Myers Squibb/Medarex	Human IgG1κ, anti-CTLA4	Cancer	
Muromonab-CD3	Orthoclone OKT3	Johnson &Johnson	Murine, IgG2α, anti-CD3	Immunological	
Ofatumumab	Arzerra	GlaxoSmithKline/Genmab	Human, IgG2κ, anti-CD20	Cancer	
Omalizumab	Xolair	Genentech/Roche	Humanized, IgG1κ, anti-IgE	Immunological	
Panitumumab	Vectibix	Amgen	Human, IgG2κ, anti-EGF receptor	Cancer	
Tocilizumab	Actemra	Roche/Chugai	Humanized, IgG1κ, anti-IL-6R	Immunological	
Tositumomab-I131	Bexxar	Corixa	Murine, IgG2αλ, anti-CD20; radiolabeled (I-131)	Cancer	
Ustekinumab	Stelara	Centocor/Johnson & Johnson	Human, IgG1κ, anti-IL12/23	Immunological	

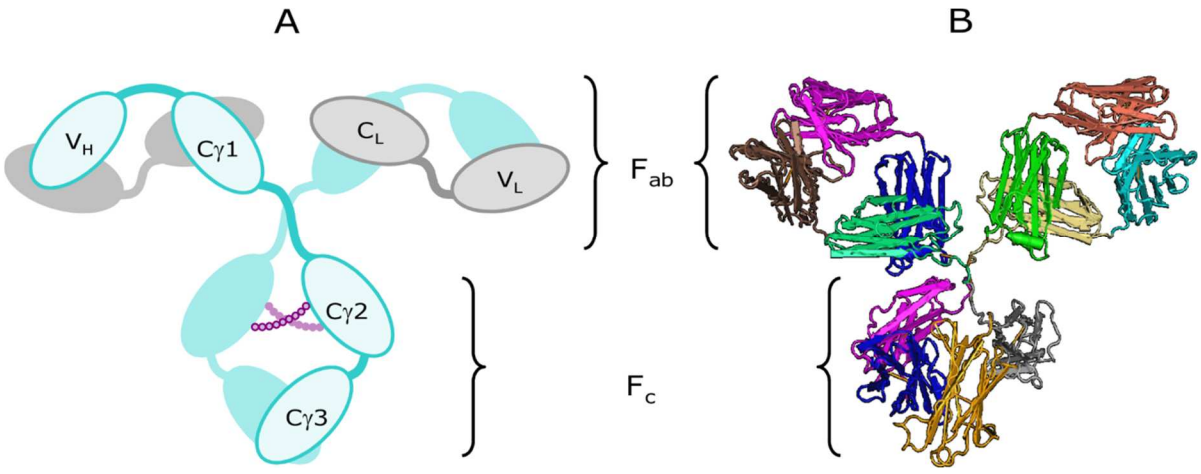


Figure 4. Structure of IgG
Figure 4A shows the spatial arrangement of IgG. The chains are identical; hence, they are arranged such that they are rotated on their *z* axis and situated one in front of the other. The chains that can be seen between the Cγ2 domains (○ and ●) represent the consensus glycosylation site of IgG. Figure 4B presents the ribbon structure of IgG, obtained from crystallographic data.[22] The separate domains are presented in different colours. Additionally, the secondary structure (α-helices and β-sheets) can be seen in panel B. The antigen binding fragment (Fab) and the constant or crystallisable fragment (Fc) are also shown

The mechanism of action of monoclonal antibodies is directly a function of their structure. The variable regions within their heavy and light chains bind with very high specificity to the antibody's respective antigen, while the constant domains determine the manner in which the antigen-bound immunoglobulin will either neutralize the antigen or activate an immune response within the host organism to remove it. The IgG isotype has been selected as the base for therapeutic mAbs because of its ability to generate immune effector functions such as antibody-dependent cellular cytotoxicity (ADCC)[49-51] and complement-dependent cellular cytotoxicity (CDC).[52, 53] These mechanisms are schematically shown in Figure 5.

Additionally, IgG-based mAbs have been shown to modulate immune response in patients.[54-56] Immune response modulation is mediated by FcγRII-B2 receptors which are expressed on the membrane of Macrophages, neutrophils and eosinophils. The purpose of these receptors is to adjust the concentrations at which immune complexes (antigen/antibody complexes) activate other cells of the immune system.[57] Anthony and collaborators[55] proposed a model in which IgGs with specific oligosaccharide structures trigger up-regulation of the synthesis of FcγRII-B2 receptors. This implies that IgGs with certain oligosaccharide structures have the ability to regulate the concentrations at which they will trigger immune responses. Further research on the mechanisms by which this occurs has been reviewed recently by Nimmerjahn and Ravetch.[56]

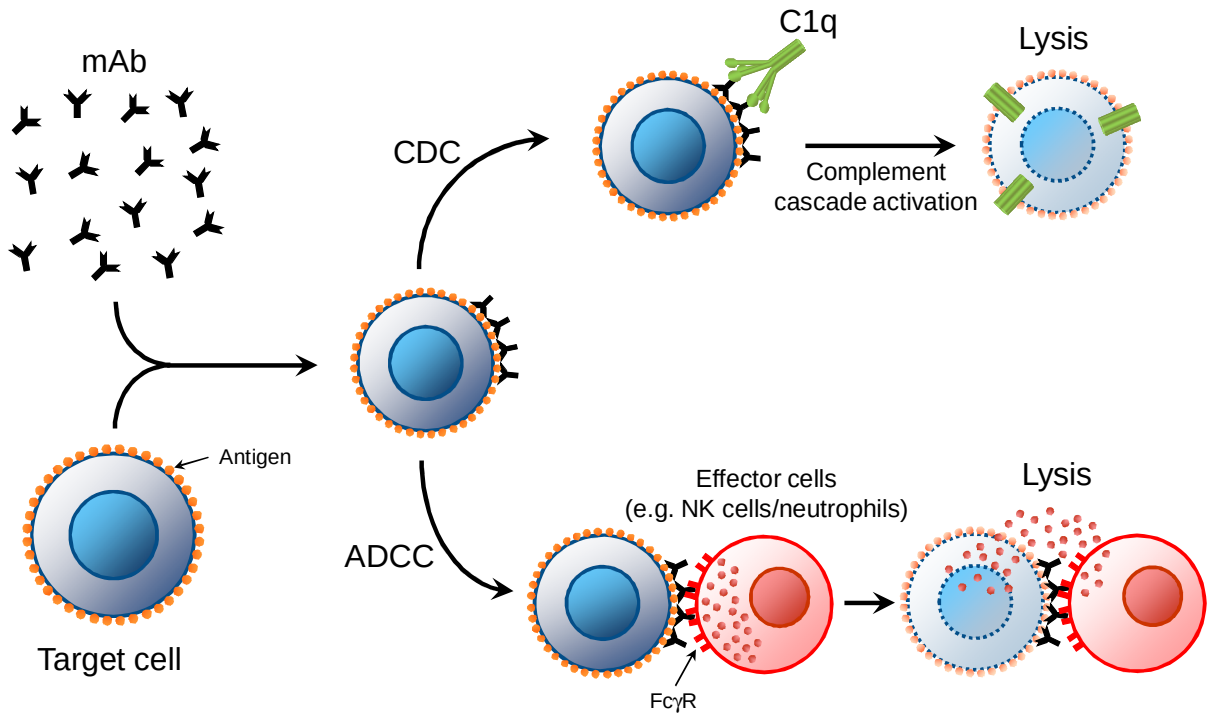


Figure 5. Cytotoxic effector functions of mAbs[23]
Antibody-dependent cellular cytotoxicity (ADCC): **The constant fragment of the mAb binds to the FcγRIII receptors which are present on the surface of natural killer (NK) cells and to the FcγRII-A receptors on the surface of neutrophils. These immune system cells recognise clusters of mAbs bound to the antigenic epitopes of whole-cell immunological threats. NK and neutrophil cells react by secreting granulocytes, which consist of a cocktail of cell lytic agents that eventually lead the immunological threat towards apoptosis, also known as cell death.[57]**
Complement dependent cellular cytotoxicity (CDC): **This mechanism is also used by the immune system to eliminate whole-cell pathogens. The Fc of antigen-bound mAbs binds to the C1q component of the complement complex. This interaction initiates the complement cascade which eventually generates the membrane attack complex, and this ultimately destroys the pathogen's membrane, thus removing it.[57]**

Considering the QbD guidelines described in previous sections, knowledge of the effects therapeutic products are meant to elicit in patients is crucial. The therapeutic indication of all currently commercialized mAbs is presented in Table 1. For treatment of cancer, mAbs are expected to generate either ADCC and/or CDC to lyse or neutralise cancer cells. The same effect is desired when mAbs are prescribed for anti-infective purposes. In the case of immunological disorders, monoclonal antibodies are expected to modulate immune response. Structural integrity of the antigen-binding (Fab) and constant (Fc) fragments of all mAbs is intrinsically necessary for their therapeutic efficacy. If the structure of their Fab is compromised, they will not bind with sufficient specificity or affinity to their antigen. If their Fc is structurally flawed, the mAbs will be unable to elicit effector functions, and thus, the host organism will be unable to remove the immunological threat.

Under QbD, the structural integrity of mAbs is defined as a critical quality attribute because it *defines* the therapeutic mechanism these products are meant to elicit in patients. Moreover, the considerations mentioned in this section provide an outline of which product properties are to be monitored and controlled to ensure that it retains the appropriate safety and therapeutic properties for the treatment it is prescribed for.

2.3. Protein Glycosylation

Protein glycosylation is the post-translational modification by which oligosaccharide structures covalently bind to the polypeptide backbone of a protein. There are two types of protein glycosylation: serine/threonine-linked glycosylation (O-linked) and asparagine-linked glycosylation (N-linked). O-linked glycosylation is seldom seen in antibodies,[58, 59] and therefore, focus will only be set upon N-linked glycosylation in this review. Asparagine-linked glycosylation is a process that begins in the endoplasmic reticulum and concludes in the Golgi apparatus. The overall process has been well described by several authors.[60, 61] A schematic description of N-linked glycosylation is presented in Figure 6.

The end product of the N-glycosylation process is a properly folded protein with an oligosaccharide chain bound to it. N-linked oligosaccharides can be of the high mannose, hybrid or complex types. Despite their classification, N-linked oligosaccharides contain a common core structure which consists of three mannose and two GlcNAc residues ($Man_3GlcNAc_2$). The two GlcNAc residues form the base that is bound to the asparagine residue of the protein backbone. The first of the mannose residues is attached to the second GlcNAc of the base. Bound to this initial mannose residue are the two other mannose residues, one joined through an α-1,3 linkage and the other one attached through an α-1,6 bond. These mannose residues give rise to the two arms that characterise all N-linked glycans. A detailed representation of the core oligosaccharide structure can be seen in Figure 7.

As their name implies, high mannose oligosaccharide structures only contain mannose residues bound to the core. For the purposes of this review, we shall consider that high mannose oligosaccharides contain up to six mannose structures apart from the core mannoses, and their specific configuration, including individual bond structure, can be seen in Figure 7. Complex-type oligosaccharides are characterised by having branches (antennae) stemming from the core structure. Depending on the degree of processing, available enzymes and protein-related stearic constraints, complex oligosaccharides can have two, three or four antennae. In turn, each antenna can contain GlcNAc, galactose and sialic acid residues.

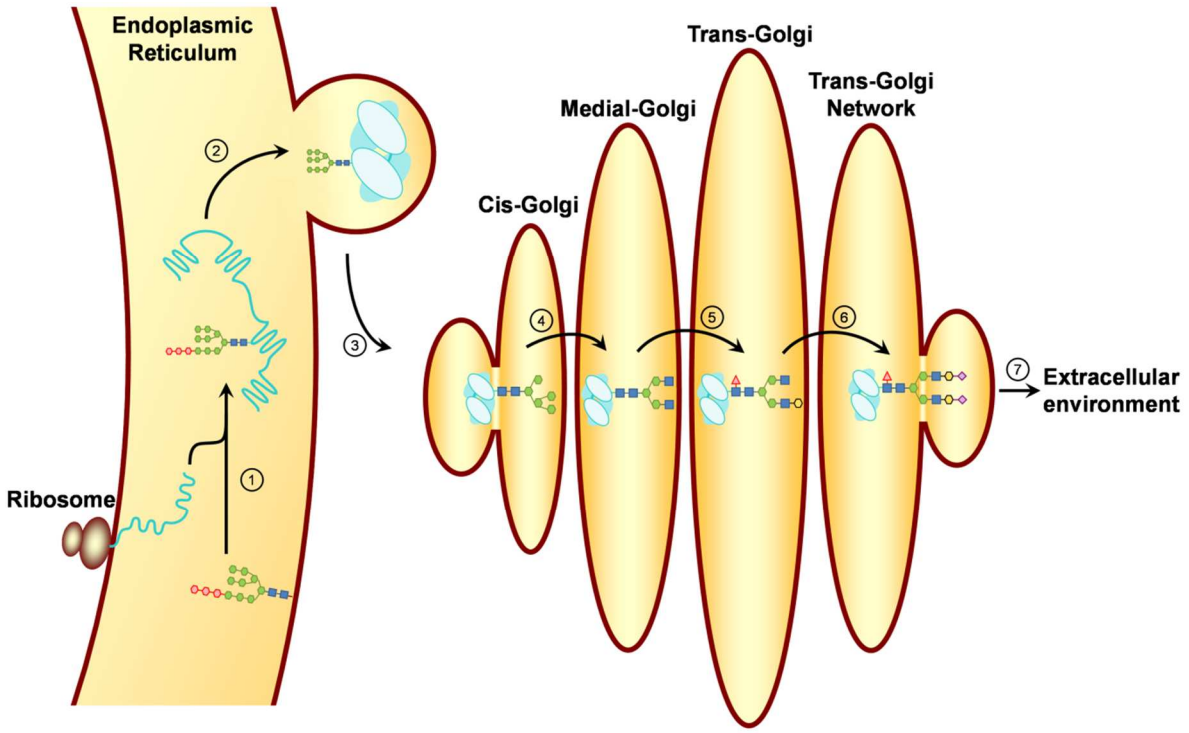


Figure 6. Protein N-linked Glycosylation Process[23]

Step 1: **The precursor oligosaccharide with a simplified formula of $GlcNAc_2Man_8Gluc_3$ is transferred onto the asparagine residue within the consensus amino acid sequon Asn-X-Ser or Asn-X-Thr (where X is any amino acid except proline) of the nascent polypeptide chain.** ***Step 2:*** **The polypeptide chain acquires its secondary, tertiary and quaternary structure by means of intra and intermolecular interactions and disulphide bonding. It is worth mentioning that the glucose (⬢) residues participate in the proper folding of the amino acid backbone as quality control markers. All three glucose residues must be removed so that the properly folded protein can be secreted to the Golgi apparatus.** ***Step 3:*** **The glycoprotein is transported by means of vesicles to the Golgi apparatus.** ***Steps 4 - 6:*** **The oligosaccharide structure bound to the polypeptide backbone is modified by a series of sugar trimming and addition reactions. The trimming reactions remove mannose (⬢) residues from the oligosaccharide. The sugar-transfer reactions add N-acetylglucosamine (GlcNAc) (■), fucose (▲), galactose (⬢) and N-acetylneuraminic acid (sialic acid or Neu5Ac) (◆) residues to the oligosaccharide chain. All of these reactions are catalysed by enzymes whose concentration varies from one compartment of the Golgi apparatus to the other.[62, 63]** ***Step 7:*** **Once the oligosaccharide is processed within the Golgi apparatus, the cell secretes the glycoprotein to the extracellular medium.**

Additionally, complex-type oligosaccharides can contain a fucose residue bound to the GlcNAc that is attached to the asparagine of the protein backbone. Since it is bound directly onto the core of the oligosaccharide, this monosaccharide residue is commonly referred to as core fucose. The structure of complex oligosaccharides can also be seen in Figure 7. Hybrid-type oligosaccharides are a combination of high mannose and complex type oligosaccharides. The α-1,6 arm of the core has mannose residues attached to it as occurs in high-mannose-type oligosaccharides. The α-1,3 arm of hybrid-type oligosaccharides contains monosaccharides that correspond to the complex-type. Again, the structure for hybrid-type glycans can be seen Figure 7.

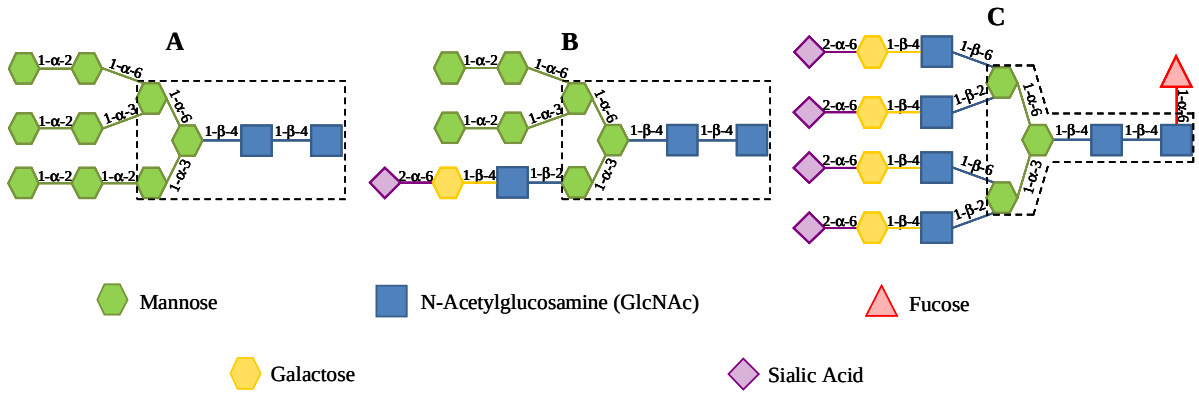


Figure 7. Types of N-linked oligosaccharide structures
Panel A shows a high mannose oligosaccharide, panel B shows a hybrid oligosaccharide and panel C shows a complex tetraantennary oligosaccharide with core fucosylation. The configuration of the bonds is also presented. The greek letter α or β represents the type of glycosidic bond and the numbers represent the carbon of the monosaccharides which form the bond. The core structure of simplified formula $Man_3GlcNAc_2$ is common for all structures and is shown inside the dashed boxes.

Two levels of glycoprotein heterogeneity are associated with glycosylation. The first is termed *macroheterogeneity* and denotes the variation resulting from presence or absence of oligosaccharide structures at specific sequons on the amino acid backbone of the protein. Macroheterogeneity arises from perturbations in the synthesis of the precursor oligosaccharide or it not being adequately added to the nascent polypeptide chain during translation. The second type of glycosylation-associated variability occurs due to different levels of oligosaccharide processing within the Golgi apparatus and is named *microheterogeneity*. Microheterogeneity occurs due to variations in the sugar trimming and addition reactions that occur within the Golgi apparatus. Concentration of substrates (nucleotide sugar donors), enzyme availability and residence time of the glycoprotein within the Golgi apparatus all influence the extent with which these reactions occur. There are five types of sugar trimming and addition reactions in the Golgi apparatus depending on the enzymes that catalyse them and the substrates involved. These reactions can be seen in Figure 8.

There are two families of mannosidases (ManI and ManII) in the Golgi, and their purpose is to catalyse the removal of mannose residues from the N-linked precursor oligosaccharide. There are five types of N-acetylglucosamine transferases (GnTI, GnTII, GnTIII, GnTIV and GnTV). These enzymes use uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) as substrate and add N-acetylglucosamine (GlcNAc) residues to the glycopeptide oligosaccharide. Galactosyltransferase (GalT) is the enzyme which adds galactose (Gal) residues to the glycan by using uridine diphosphate galactose (UDP-Gal) as its substrate. The final type of enzyme that participates in the processing reactions within the Golgi apparatus is sialyltransferase (SiaT). This enzyme uses

cytosine monophosphate N-acetylneuraminic acid (denoted as either CMP-Neu5Ac or CMP-SA) as substrate and adds sialic acid residues (Neu5Ac or SA) to the asparagine-bound oligosaccharide.[64, 65]

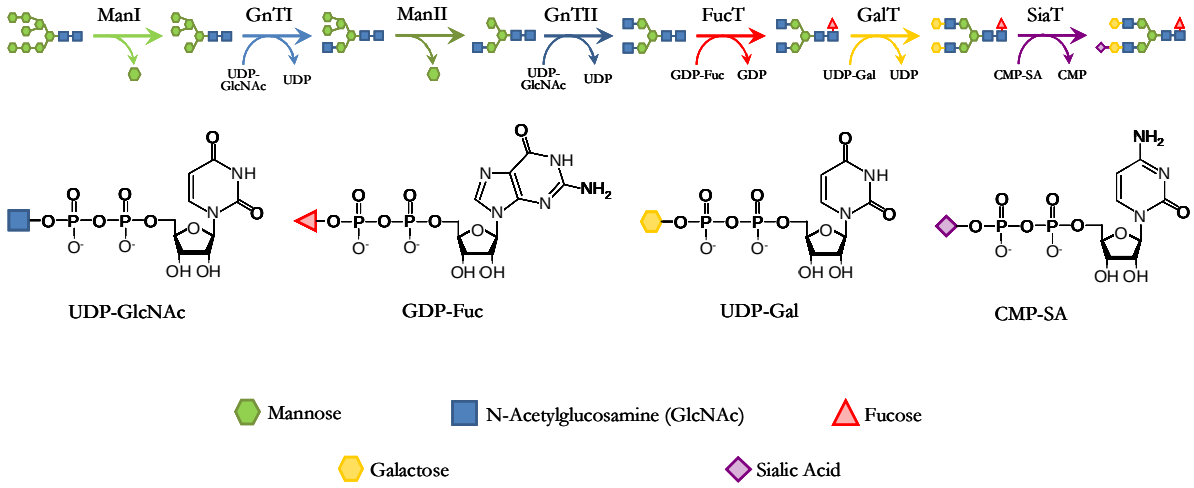


Figure 8. Oligosaccharide processing reactions of the Golgi apparatus[29]
The oligosaccharide structures are trimmed by mannosidase reactions. Also, monosaccharides are added to the oligosaccharide by action of specific enzymes. The substrates for each of these reactions (UDP-GlcNAc, GDP-Fuc, UDP-Gal and CMP-SA), and the by-products (UDP, GDP and CMP) can be seen in this figure.

2.3.1. Nucleotide sugar metabolism

The substrates for the sugar addition reactions described in the previous section are the protein-bound oligosaccharide chain and the nucleotide sugar donors (NSD). As substrates for these reactions, their availability within the cell can affect both the macro and microheterogeneity of N-linked oligosaccharides. The availability and concentration of nucleotide sugars in the cell's cytoplasm determines the rate at which the precursor oligosaccharide is synthesised on the membrane of the endoplasmic reticulum. If nucleotide sugar availability in the cytoplasm is low, the rate at which the precursor oligosaccharide is produced will also be low. When this occurs, the precursor oligosaccharide may be available in insufficient quantities for it to be added onto the N-glycosylation site(s) of the protein, thus affecting glycoprotein macroheterogeneity. Moreover, the concentration of nucleotide sugars within the Golgi apparatus determines the extent of processing reactions therein, thus affecting glycoprotein microheterogeneity as well. Mammalian cells are able to produce all of the required nucleotide sugars through metabolism by using simple carbon sources, such as glucose and glutamine, as inputs. Figure 9 presents an overview of the metabolic network that is used by mammalian cells to produce NSDs. The metabolic synthesis of nucleotide sugars is closely related to central carbon metabolism and energetic state of the cells. In Figure 9, we see that

the nucleotide sugar metabolism consists of a series of enzyme-catalysed reactions whereby extracellular components, such as glucose, galactose, mannose and others, are sequentially modified to produce the NSDs. An interesting feature of the metabolic pathway is that all NSDs can be produced from a single carbon source, such as glucose, and an organic nitrogen source (such as glutamine) and that the entire pathway is highly interrelated.

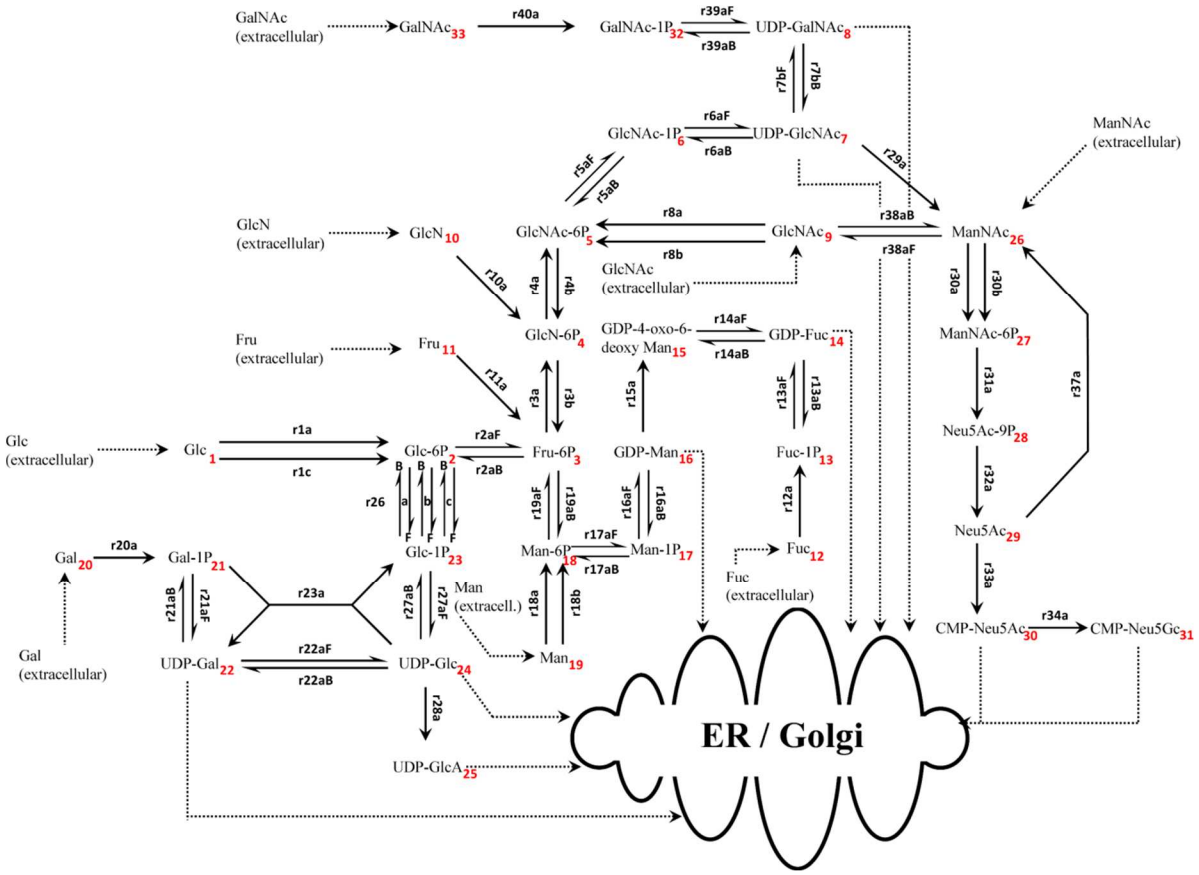


Figure 9. Nucleotide sugar metabolism in mammalian cells[66, 67]
The solid arrows represent reactions while the dashed arrows represent transport into the cell (glucose, galactose, fructose, mannose, GlcN, fucose, ManNAc) and from the cytosol towards the endoplasmic reticulum or Golgi apparatus (UDP-GlcNAc, GDP-Fuc, UDP-Gal, CMP-Neu5Ac, UDP-GlcA, UDP-GalNAc). The numbers in red enumerate the number of components in the pathway.

2.4. Effects of glycosylation on mAb structure and function

All IgG-based mAbs present a conserved N-linked glycosylation site at asparagine 297 of the Cγ2 domains of their heavy chains. Numerous reports both for polyclonal and monoclonal antibodies have shown that the oligosaccharides bound to this site have a profound effect on the antibody Fc 3D structure and effector functions.[25-27] Additionally, the variable region of IgG-based mAbs may also contain glycosylated N-linked sequons. It has been found that in many cases where this occurs, the oligosaccharide structures of these sites influence the affinity of the mAb for its respective antigen.[59, 68-71] However, the presence of oligosaccharides in these

fragments is relatively uncommon – between 15 and 20% of human polyclonal IgG contain Fab glycans.[72, 73] Hence, this review will mainly focus on glycosylation of the Fc fragments of mAbs. Given this tight glycosylation-structure/function relationship, it is clear that adequate glycosylation of IgG-based mAbs is essential towards their use as therapeutics. The following section will review some of the efforts that have been made to analyze the impact mAb oligosaccharide profiles have on their structure and function.

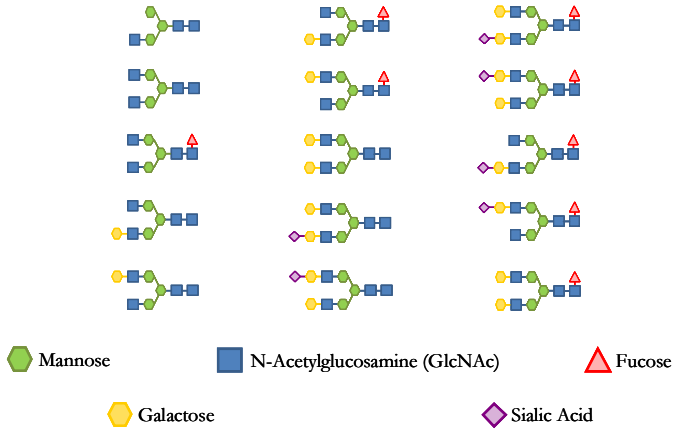


Figure 10. Common oligosaccharide structures on the Fc of mAbs
The most common glycans on the Fc of mAbs are complex, biantennary structures with varying degrees of fucosylation, galactosylation, and sialylation.[74-76]

Both heavy chains of the mAb are in close proximity to each other due to disulphide bonding and intermolecular interactions. This spatial constraint reduces the affinity of the glycosyltransferase enzymes, which are responsible for the addition of sugar residues to the growing oligosaccharide chain, for their substrates.[77, 78] It is for this reason that the Fc oligosaccharides of mAbs are not as diverse as for other glycoproteins. This does not mean, however, that there is not a high degree of heterogeneity in these structures. Many studies have reported that the most common oligosaccharide structures bound to the Fc of monoclonal antibodies are of the complex, bi-antennary type with varying degrees of core fucosylation, bisecting GlcNAc, galactosylation and sialylation. The most commonly observed oligosaccharide structures on the Fc of IgG-based mAbs[74-76] are schematically shown in Figure 10.

Given the distribution of oligosaccharides bound to the Fc of mAbs, many research groups have focused on the influence these structures have on mAb function both *in vivo* and *in vitro*. Many anti-cancer mAbs rely on their ability to generate ADCC for the removal of cancerous cells.[27, 49, 51] One of the findings that has attracted substantial interest by drug developers is that absence of core fucosylation of the oligosaccharide structures can increase ADCC dramatically – up to 50 fold.[79-82] Capitalizing on these findings, groups have developed methods for reducing or completely inhibiting fucosylation of mAbs.

The early efforts to increase ADCC by glycan engineering were based on over-expressing genes for GnTIII, which is the sugar-transfer enzyme that adds bisecting GlcNAc to the core mannose of the oligosaccharide.[83] At that point it was thought that the presence of bisecting GlcNAc on IgG1-based mAbs contributed to enhanced ADCC. However, it was later observed that the addition of bisecting GlcNAc residues increased ADCC by competing with the addition of fucose to the oligosaccharide structures.[81] Additional attempts involving over-expression of competing enzymes to reduce oligosaccharide fucosylation were made when GnTIII and mannosidase II (ManII) were over-expressed simultaneously to yield low-fucose Fc oligosaccharides.[84] The other tendency has been to either silence[85, 86] or knock out[87, 88] the fucosyltransferase genes in mAb-producing cell lines.

It is worth noting that there is a risk in genetically modifying cell-lines. It is possible that extensive engineering of the endogenous (native) glycosylation pathways in cells could lead to negatively impacting their growth profile or protein productivity. With this in mind, additional strategies based on disrupting other pathways involved in N-linked glycan biosynthesis have been developed. Reduction in core fucosylation of mAbs has been achieved by supplementing culture medium with fucose analogues that inhibit the activity of GDP-mannose dehydratase. This enzyme is responsible for synthesizing GDP-fucose which, in turn, is the direct precursor for glycan core fucosylation; thus, inhibition of its activity leads to reduced fucose content on mAb Fc glycans produced under these conditions.[89] A similar strategy that has been employed is to add castanospermine (CS) to culture medium. CS is a mannosidase inhibitor and, as such, prevents cleavage of mannoses from early glycoproteins, thus yielding lower fucosylation.[90] Although fucosylation is prevented with the CS strategy, high mannose structures are produced. It has been widely reported that mAbs with high mannose glycans have reduced serum half-life in patients, an effect likely due to recognition by the mannose receptor of macrophages.[91-93] Given the synthesis of undesirable high mannose glycans, applicability of the CS strategy is yet to be confirmed.

Other interesting findings have revealed that mAb Fc glycans also influence CDC induction either by binding to the C1q component of the complement[94, 95] or by interaction with the mannose binding lectin (MBL).[96, 97] Specifically, it has been shown that absence of galactose on mAb Fc glycans yields lower affinity of this fragment for the C1q component of the complement cascade.[25, 98] It has also been reported that non-galactosylated Fc glycans interact with the MBL, and can thus activate the complement cascade.[96, 97] Similarly to the glycosylation engineering attempts made to increase ADCC, close attention should be given to glycosylation profiles of mAbs when their therapeutic mechanism relies on CDC.

In contrast to glycans suitable for eliciting cytotoxic effector functions, the oligosaccharide structures of mAbs used for the treatment of autoimmune disorders should modulate the immune response of patients. Studies on therapeutic intravenous IgG (IVIG) have found that the presence of terminal sialic acid residues on the oligosaccharides of their Fc fragments has great influence on their anti-inflammatory effect.[54-56] IVIG must be administered in high doses (1-2 g/kg) to achieve anti-inflammatory effects. This has been found to correlate with the fraction of IVIG that contains α2,6-bound terminal sialic acid residues on their Fc oligosaccharides.[54] Again, certain cell lines have been engineered to ensure that sialic acid addition to glycoproteins is adequate.[99-101]

Viewing the information presented in this section under the QbD scope, it is possible to consider different glycan profiles as integral components of the *quality target product profile* (QTPP) of mAbs. Following the definition of QTPP, the ideal glycosylation profile of an ADCC-based therapeutic mAb would contain no core fucosylation. Similarly, a mAb whose therapeutic function relies on CDC would be fully galactosylated and a mAb developed for immune response modulation would be sialylated. Furthermore, given the importance Fc glycosylation has on mAb safety and therapeutic efficacy, the composition and distribution of these carbohydrate structures must be considered *critical quality attributes* (CQAs) of these molecules. Defining glycans as CQAs of mAbs implies that their composition and distribution must be monitored and controlled within ranges that ensure product safety and therapeutic efficacy throughout process development and manufacture.

2.5. Effects of bioprocess conditions on protein glycosylation

Having defined Fc glycosylation as a CQA of mAbs, the next step in QbD is to determine which manufacturing process inputs affect the CQAs and – mechanistically – how this occurs. Monoclonal antibodies are produced by culturing mammalian cells, and as is common for these processes, nutrient availability, dissolved oxygen, supplements, temperature, operation schedule and culture modes can all influence product quality. The following section will present the extent at which these process inputs have been reported to influence the composition and distribution of N-linked carbohydrates on therapeutic glycoproteins. It will also provide mechanistic descriptions of how these inputs have been proposed to impact glycan composition and distribution. Other reviews on these topics are available.[28, 29, 102]

2.5.1. Nutrient availability

Several studies have demonstrated that carbon source and its availability has great influence on therapeutic protein glycosylation.[103] The impact of carbon source availability has been observed to affect both the site occupancy (*macroheterogeneity*) and glycan profile (*microheterogeneity*) of glycoproteins. Experimental findings indicate that when monosaccharides other than glucose are used as the main source of carbon for cell growth, the glycoform is affected.[104] For example, Tachibana et al.[105] observed differential glycosylation on the Fab of a monoclonal antibody due to substitution of glucose by fructose, mannose and galactose. The effects of carbon source on protein glycosylation are likely due to the differential efficiency with which cells channel monosaccharide inputs through the nucleotide sugar pathway. This, in turn, explains the effects that complete carbon source replacement has on both macro and microheterogeneity, as reviewed by Goochee & Monica.[103] However, quantitative information on the effect of carbon source on intracellular nucleotide sugar pools and protein N-glycosylation is still lacking, and few direct links have been established.

Studies on the impact of carbon availability on protein glycosylation have mainly involved glucose as the primary carbon source and glucosamine as a supplement. Hayter and collaborators[106] found that when glucose concentrations varied from 0.3 mM to 0.8 mM within a chemostat, the fraction of fully glycosylated interferon gamma (IFN-γ) went from around 50% to 70%, while the proportion of non-glycosylated IFN-γ went from 20% to below 5%. However, it was not until key research by Nyberg and collaborators[107] that the decrease in site occupancy of IFN-γ under glucose and glucosamine starvation was explained. They found that site occupancy issues occurred most likely due to low biosynthesis of UTP and glucosamine-phosphate, which are both substrates for the production of UDP-GlcNAc. In turn, UDP-GlcNAc is a fundamental substrate for the production of the precursor oligosaccharide on the membrane of the endoplasmic reticulum. If UDP-GlcNAc is not available, the precursor oligosaccharide cannot be synthesized, and hence, the first step of N-glycosylation cannot proceed.

The findings by Nyberg et al.[107] were complemented when Wong et al.[108] observed that despite controlling glucose and glutamine concentrations to low values (0.7 mM of glucose and 0.1 mM of glutamine) within a fed batch system, no site occupancy issues arose. Through their findings, this group concluded that only extreme glucose starvation (concentrations below 0.35 mM) reduced protein productivity and generated site occupancy issues. However, they did observe changes in the microheterogeneity of their model glycoprotein. These consisted in lower sialylation and increased production of high mannose oligosaccharide structures. These negatively affect IgG effector functions,[91, 94] as well as increasing mAb clearance rates from

human sera, which further reduces their therapeutic efficacy.[91, 92, 109] An additional consideration comes from the fact that highly sialylated oligosaccharides on the Fc of mAbs increases their anti-inflammatory capacity.[54, 56] It is advantageous for mAbs relying on inflammatory responses (ADCC and CDC) to contain low levels of sialylation, whereas for anti-inflammatory mAbs high proportions of sialylated Fc oligosaccharides are desired in order to increase their therapeutic efficacy.

Both Nyberg et al.[107] and Wong et al.[108] proposed that the effect of nutrient limitations on protein glycosylation was most likely due to a decline in intracellular availability of UDP-GlcNAc. This can be explained through the glycosylation process within the Golgi apparatus and nucleotide sugar metabolism. The initial steps of N-glycan processing are sequential trimming reactions whereby mannose residues are cleaved from the $Man_8GlcNAc_2$ oligosaccharide to yield $Man_5GlcNAc_2$, as can be seen in Figure 7. At this point, the first sugar transfer reaction occurs, which is the GnTI (*N*-acetylglucosaminyltransferase I) catalysed addition of a GlcNAc residue to the α-1,3 branch of the core oligosaccharide. The substrates for this reaction are the oligosaccharide chain bound to the protein and UDP-GlcNAc. If this reaction does not occur, the subsequent reactions of glycan processing in Golgi cannot proceed, and the resulting oligosaccharide structure will be of the high mannose variety.

The reduction in sialylation can be explained by the biosynthetic pathway of CMP-Neu5Ac, which is the substrate required for addition of sialic acid (Neu5Ac). The primary substrate for this pathway is ManNAc, which is, in turn, produced through the enzymatic modification of UDP-GlcNAc. Again, if low levels of UDP-GlcNAc are available within the cell, very little ManNAc will be produced through UDP-GlcNAc, and thus very low levels of CMP-Neu5Ac will be available as substrates within the Golgi apparatus. Furthermore, it has been shown that addition of ManNAc to cell culture media has considerably increased the sialylation of produced glycoproteins.[110] Contradictorily, no such effects were observed when UDP-GlcNAc was added to the medium, but this has been attributed to the fact that high cytosolic UDP-GlcNAc concentration limits the transfer of CMP-Neu5Ac into the Golgi apparatus.[111] These internal control loops give evidence of the sophisticated levels of intracellular organization surrounding the protein glycosylation processes.

2.5.2. Dissolved Oxygen

Another process parameter linked to mAb glycosylation is dissolved oxygen availability. The observed effects of dissolved oxygen tension appear to vary depending on glycoprotein and cell line. High values of dissolved oxygen tension (DOT) (100% DOT) have been reported to

decrease the sialic acid content of an erythropoietin-antibody crystallisable fragment fusion protein (EPO-Fc) produced in Chinese hamster ovary cells (CHO),[112] whereas Chotigeat and collaborators[113] observed increased sialylation of human follicle stimulation hormone (hHFS) expressed in the same cell line at a dissolved oxygen tension (DOT) of 90%. Moreover, Gawlitzek et al.[114] and Trummer et al.[112] observed that reduced dissolved oxygen (0% to 10% DOT) had no observable impact on sialylation.

Table 2. Effect of dissolved oxygen tension (DOT) on protein N-glycosylation[29]

PRODUCT	CELL LINE	DOT	EFFECT
Human follicle stimulating hormone (hFSH)	CHO	10 - 90%	Maximum sialylation at 90% DOT[113]
Interleukin-2-N-glycosylation variant	BHK-21	0 and 100%	No effect on glycosylation at 0% DOT[114]
IgG1 mAb	Murine hybridoma CC9C10	0, 50 and 100%	Galactosylation increases with DOT[115]
EPO-Fc	CHO	10, 30, 50, 70, 90 and 100%	No effect observed from 10 to 90%. Loss of sialylation at 100% DOT[112]

Additionally, it has been reported that mice, when grown under hypoxic conditions, produce transferrin with increased core fucosylation and IgG with reduced core fucose residues.[116] High values of dissolved oxygen have also been reported to increase the galactose content of mAb glycans.[115] This group observed that higher galactosylation occurred at a DOT of 100% and was significantly lower at DOT of 10%. The extent of galactosylation impacts the effector functions generated by mAbs[94] and has even been described as possibly being the optimal mAb glycoform for FcγR binding.[26] As mentioned earlier, FcγR binding is fundamental for eliciting effector functions, and thus is crucial for effectiveness of the mAb as a therapeutic.

Despite the efforts that have been made to understand the interplay between dissolved oxygen and glycosylation, the mechanisms by which this occurs have yet to be defined.[102] The main hypothesis as to why galactosylation increases with dissolved oxygen concentrations is that the intracellular concentration of uridine diphosphate galactose (UDP-Gal) increases with dissolved oxygen concentrations. Tentatively, it could be possible to relate increased UDP-Gal concentrations at high dissolved oxygen values through nucleotide sugar metabolism. Kunkel and collaborators[115] found that high values for dissolved oxygen generated shifts in the energy metabolism of the cultured cells, and that these shifts included higher flux of glucose through the glycolysis pathway, of which the first intermediary is glucose-6P. Glucose-6P serves as an early precursor of UDP-Gal biosynthesis, and the increase in its production could, at least in part, account for the higher availability of UDP-Gal for oligosaccharide processing within Golgi.

2.5.3. Ammonia and pH

The effect ammonia concentration and pH have on protein glycosylation is closely related. Ammonia concentration and pH have been reported to have a synergic effect on protein N-glycosylation macroheterogeneity. Borys and collaborators[117] found that when increasing ammonia concentration from 0 to 9 mM, the fraction of heavily glycosylated mouse placental lactogen-I (mPL-I) went from 90% to 25% at a pH of 8.0, from 90% to 65% at a pH of 7.6 and from 90% to 80% at a pH of 7.2.

Increasing concentrations of ammonia (from 1 mM to 15 mM of NH_4Cl) have been observed to reduce the galactosylation and sialylation of a tumour necrosis factor-IgG fusion protein up to 40%.[118] It was concluded that accumulation of NH_4^+ affects N-glycosylation by increasing the pH of the Golgi compartments to levels at which the GalT (β1,4-galactosyltransferase) and the SiaT (α1,3-sialyltransferase) enzymes have lower activities. It has also been reported that high pH values cause mislocalization of the ManII, GnTI, GalT and SiaT glycosyltransferase enzymes along the Golgi apparatus.[119, 120] This mislocalization, in turn, has been associated with lower galactosylation and sialylation of N-linked glycans.[120]Additional findings by Gawlitzek et al.[121] showed that ammonia concentrations of 15 mM increased branching on the N-glycans of a recombinant interleukin-2 produced in BHK-21 cells. This was observed to correspond with an increase in availability of UDP-GlcNAc and UDP-GalNAc. Both reports by Gawlitzek and collaborators[118, 121] are consistent given that there is an inverse relationship between UDP-Gal and UDP-GlcNAc. The more UTP that is consumed within the cell to produce UDP-GlcNAc, the less there is to produce UDP-Gal.

Borys and collaborators[117] observed that glycosylation was relatively unaffected in a pH range between 6.9 and 8.2 at low ammonia concentrations. However, they did find decreased site occupancy at pH values outside that range. Yoon et al.[122] found that the sialic acid content decreases from 68% to 40% between a pH range of 6.8 and 7.8. This is likely due to decreased GalT and SiaT activities at high Golgi pH values reported by Gawlitzek et al.[118] The relationship between galactosylation and sialylation with pH has also been observed by Muthing et al.[123] This group found that culture medium pH values of 7.4 yielded the highest proportion of fully galactosylated oligosaccharides in a murine IgG3 monoclonal antibody, whereas the lowest levels of sialylation occurred at a pH of 6.9. These results are also consistent and possibly due to decreased GalT and SiaT activities at high pH,[118] but also may be due to glycosyltransferase mislocalization in Golgi.[120]

2.5.4. Cell culture viability, growth phase and temperature

The manner in which temperature affects protein N-glycosylation has been related to cell viability and cell growth stage.[124-126] These research groups have proposed that decreased temperatures maintain cell viability over longer periods of time, and that protein glycosylation occurs more efficiently under these high-viability conditions. This is consistent with previous research that has shown that the genes involved in the synthesis of the precursor oligosaccharide on the membrane of the endoplasmic reticulum are up-regulated during the transition between growth stage 0 (G0) and growth stage 1 (G1).[127]

Additionally, Ahn and collaborators (2008) found that at temperatures below 32°C, branching and sialic acid content of EPO produced in CHO cells decreased when compared to the glycoforms produced at 37°C. This correlates with previous observations where the intracellular pool of UDP-GlcNAc and UDP-GalNAc were found to decrease at lower temperatures.[128] This implies that there is an optimum temperature for N-linked glycosylation where production of the precursor oligosaccharide is the highest permitted by availability of UDP-GlcNAc.

2.5.5. Serum and lipid supplements

The effect of lipid supplements on N-linked glycosylation has been reported by Jenkins and collaborators.[129] This group observed that by supplementing the culture medium with a mixture of bovine serum albumin (BSA) and lipids, site occupancy increased greatly, when compared to their control experiments. They proposed that this increase in site occupancy was either due to absence of possibly toxic oxidised lipids which are found in conventional BSA preparations or due to the availability of additional lipids which favour glycosylation. Jenkins et al.[129] also observed that the duration of the growth phase of the culture increased when a lipid supplement was added to the medium during the late log phase of the culture. Similarly to temperature, a prolonged growth phase could also account for increased N-glycosylation site occupancy.

Another cell culture factor that has been observed to affect N-linked glycosylation is the use of chemically defined or serum-supplemented culture media. The desire to avoid animal-derived components in cell culture media has led several research groups to analyse the effects of supplements on N-glycosylation of therapeutic mAbs. Lifely and collaborators[79] produced the IgG-based mAb CAPMPATH-1H by culturing CHO cells in either serum supplemented or serum free media. They found that the major oligosaccharides bound to CAMPATH-1H in both cultures were the same; however, they also found that the product derived from serum free media had

less mono and bi-galactosylated species. Lifely et al.[79] observed that the relative amount of sialylated species was higher for the serum free culture than for its non-supplemented counterpart. Gawlitzek et al.[114] produced recombinant interleukin-2 in BHK cells in absence or presence of fetal calf serum, and found that the serum free cultures increased the core fucosylation and sialylation levels on the glycan structures of their product. The latter is consistent with the findings by Lifely and collaborators.[79] An additional study quantified the effect of serum supplementation, serum-free preparations and chemically-defined media on the oligosaccharide profiles of a mAb produced by hybridoma cell lines.[130] They observed that serum-supplemented cultures produced broader distributions of glycans when compared to the serum-free and chemically-defined media. They also observed clear differences in the oligosaccharide profiles derived from each of their culture media. They found higher galactosylation when they used serum-supplemented and chemically-defined media compared to when they used serum-free media. Moreover, they observed reduction of sialylation but increases in the amount of core fucosylation in serum-free and chemically-defined media. As seen above, the effects of serum supplementation on the oligosaccharide profiles of glycoproteins has been reported to be consistent throughout. However, because some of the components of serum preparations are unknown, the mechanisms by which these effects occur are not fully understood.

2.5.6. pCO_2 and osmolality

The effect of CO_2 partial pressure (pCO_2) and osmolality of culture media on the oligosaccharide moieties of a mAb have been reported.[131] They observed that the macroheterogeneity of an IgG2-based mAb was unaffected at elevated values of pCO_2 (up to 250mmHg) and osmolality (up to 476 mOsm/kg). They also found that, while controlling osmolality at 320 mOsm/kg, elevating pCO_2 increased mAb galactosylation, but at controlled pCO_2 (either at 40 mmHg or 195 mmHg) high osmotic stress (435 mOsm/kg) negatively impacted galactose content by 5% and 25% respectively. They concluded by proposing that the reduction in galactosylation was due to increased pH of the Golgi apparatus. As mentioned previously, other research groups have proposed that the GalT (β1,4-galactosyltransferase) enzyme has reduced activity at high Golgi pH values.[118, 132] Elevated pH of the Golgi apparatus due to elevated pCO_2 and hyperosmotic stress may be the underlying reason for the effects observed by Schmelzer & Miller.[131] However, these authors mention that measurements of Golgi pH under these conditions are still lacking.

2.5.7. Stirring speed

Another very important process parameter is stirring speed. Adequate stirring within cell culture environments ensures that all cells receive equal amounts of nutrients, oxygen and heat. However, excessive stirring can cause the cells in culture to become stressed, and thus affect their viability, mAb productivity and mAb quality (including glycosylation). Only the effect of shear stress on site occupancy has been investigated.[133] This group found that while increasing stirring speed from 40 to 200 rpm (with corresponding Reynold's number (Re) values of $1.87x10^3$ and $9.31x10^3$), the fraction of underglycosylated Tissue-Type Plasminogen Activator (tPA) increased from 31% to 72%.

They proposed that at high stirring rates (high shear stress), cells respond by up-regulating membrane synthesis, increasing the rigidity of their membranes by additional synthesis of cholesterol and overall increase in protein synthesis, all at the expense of lower doubling rates. Of these, the factor to which they attribute lower site occupancy is the higher rate of protein synthesis. They concluded that at higher protein synthesis rates, the retention time of glycoproteins within the endoplasmic reticulum decreases, and because of this, N-glycan site occupancy is negatively affected.

2.5.8. Culture modes and operation schedule

Much of the research that has aimed to elucidate the effects of process conditions have been carried out in different cell culture systems. These culture systems have mainly been batch bioreactors,[79, 107, 110, 112, 117, 118, 124, 129, 131-136] continuous bioreactors[107, 114, 121, 126, 137] and fed-batch bioreactors.[107, 108, 124, 136, 138] Additionally, more complex systems have also been used, such as hollow fibre bioreactors,[79] biphasic bioreactors[139] and microcarrier-immobilised cells in fluidised bed bioreactors.[114] It is more suitable to discuss the effects batch cultures, perfusion cultures and fed-batch systems have on product N-glycosylation because these are the culture modes that predominate in the biopharmaceutical industry and because the more complex culture systems have been used for research purposes.

The relationship between product N-glycosylation and cell culture mode can directly be linked through the metabolic effects presented previously in this section. Each of these three culture modes implies different cell growth and viability, nutrient consumption and metabolite production profiles. In a simple batch system, N-glycosylation issues arising from cell growth stage, cell culture time, nutrient availability and metabolite production would all be expected to occur, although they could be relatively offset by selecting optimal batch times or by reducing temperature to ensure a favourable growth stage for protein productivity and glycosylation.[125,

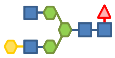
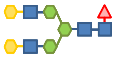
[126, 137] In a fed batch reactor, nutrient availability issues would be resolved, but the rest of the drawbacks would still be present. For these systems, accumulation of metabolic by-products and the effects of cell culture time on N-glycosylation would be predominant. However, these issues could be addressed through the same means as for normal batch systems. The major advantages for batch and fed batch systems are their simplicity, their flexibility to be used for the manufacture of several different products and their relatively high yields and final product concentrations.[140] Contrastingly, continuous bioreactors offer more control over most of the cell culture conditions. Nutrient depletion is addressed by continuous addition of culture media; metabolic by-product accumulation is resolved by continuous removal of cell culture medium, and cell culture age, cell viability and cell growth stage can all be manipulated by the dilution rate of the continuous system and cell recirculation. For example, Nyberg et al.[107] reported that low dilution rates reduce N-glycosylation site occupancy due to nutrient limitations. Ahn and collaborators[126] also observed the effect of dilution rates on the microheterogeneity of a mAb. They found that galactosylation decreased with lower dilution rates (from 0.2 to 0.3 d^{-1}) and proposed that this effect occurred because the concentration of metabolites, such as NH_4^+, inhibited mAb galactosylation as had been previously described.[118, 132] Moreover, Ahn et al.[126] observed an increase in galactosylation at higher dilution rates (between 0.5 and 0.8 d^{-1}) which could be explained by higher nutrient availability, a cell growth stage more favourable for N-glycosylation or lack of inhibitory metabolite (NH_4^+) effects under those conditions.

2.5.9. Cell line

The choice of production cell line also has a significant effect on the microheterogeneity of mAb Fc glycosylation. The expression of glycosyltransferase enzymes and glycosidases varies from one cell line to another.[141] Since these enzymes catalyse the monosaccharide trimming and addition reactions that occur in Golgi, their concentration and distribution along the secretory pathway determine the extent of oligosaccharide processing and overall glycan distribution of secreted proteins. The cell lines which are more commonly used for the production of mAbs are CHO and NS0, although other cell lines, such as Sp/0, Y0 and human lymphoma clones, have also been employed to a lesser extent. As the most common hosts for mAb production, the glycosylation profiles produced by NS0 and CHO have been compared frequently.[45, 79, 142-144] All these studies have observed that while CHO cells produce bi -galactosylated glycans where each galactose residue has a β–1,4 bond with the preceding GlcNAc on each antenna of the oligosaccharide (row 3, Table 3), NS0 cell lines generate bi-galactosylated structures with two galactose residues on one of the antenna. This occurs because murine-derived cell lines encode genes for glycosyltransferase enzymes which can add galactose residues with an α-1,3

glycosydic bond to a preceding galactose residue (row 4, Table 3). Significance of these α-1,3-bound galactose residues is highlighted by a report that found these motifs produced anaphylactic reactions in patients.[145] Beck and collaborators[143] also found that despite CHO and NS0 generating the same predominating oligosaccharide structures, the relative abundance of these three is different in both cell lines (rows 1 to 3, Table 3)

Table 3. Differences between mAb glycosylation profiles between the CHO and NS0[143]

NAME	STRUCTURE	CHO	NS0
G0F		61%	15%
G1F		36%	53%
G2F		3%	26%
G2F (α-1,3)		No α-1,3 Gal bonds are produced	This bond is produced

Raju et al.[141] carried out a study in which they compared the serum IgG oligosaccharide structures of several species including human, rat, mouse and guinea pig. They found clear species-specific differences in the type of sialic acid residues on the oligosaccharides. They observed that human oligosaccharides only contained N-acetyl neuraminic acid (NANA), while mouse glycans contained N-glycolylneuraminic acid (NGNA) residues and rat glycans contained both NANA and NGNA. Additionally, this group found that the proportion of monogalactosylated isomers varied considerably from species to species. They attributed these differences to a species-dependent specificity of the β1,4-galactosyltransferase enzymes for either one or the other antenna on the complex oligosaccharide structures.

Lifely and collaborators[79] (1995) also used the Y0 cell line to produce a version of the CAMPATH mAb. They found that this cell line introduced bisecting GlcNAc to the central mannose of the core oligosaccharide structure. Stadlmann et al.[45] compared the glycoforms of commercial mAbs produced in these expression systems and found similar discrepancies between the oligosaccharide profiles produced by CHO and NS0. This group also measured the oligosaccharide profiles of an anti-HIV mAb generated by a human lymphoma cell line, and they observed that its oligosaccharides were more heavily galactosylated and sialylated. In addition, this group also determined the IgG Fc glycosylation patterns of guinea pigs and observed that the predominating oligosaccharides were mono and di-sialylated structures with bisecting GlcNAc.

Another promising tendency in the biopharmaceutical sector has been to explore the production of biological therapeutics in more robust expression systems, such as bacterial, yeast, insect or plant cells. From the preceding paragraph, which mentions glycosylation pattern variation among different mammalian cell lines, it can be inferred that the differences should increase as the evolutionary distance of the production cell lines diverges from mammals. Interestingly, a form of N-linked glycosylation has been described in bacteria of the *Campylobacter* genus.[146-148] However, the N-linked carbohydrates produced by these bacteria are relatively simple (only three constituting monosaccharide species).[146-149] Speculatively, it is possible that higher complexity of glycan structures is achievable exclusively in eukaryotes given that they have internal compartmentalisation (i.e., presence of the Golgi apparatus).

Accordingly, native yeast cells are capable of glycosylating proteins; however, they lack many of the transferase enzymes of the Golgi apparatus, and thus generally produce only high mannose oligosaccharide structures.[101, 150, 151] Plant cells express most of the required enzymes; however, they have additional transferases that can add non-mammalian sugars such as xylose, mannose and fucose to complex oligosaccharides.[152] Finally, certain insect cell lines also contain many of the glycosyl transferase enzymes necessary to produce glycoforms similar to those found in mammalian systems.[153-155] However, many differences have been found between the resulting glycoforms produced by these and mammalian cell lines. Most insect cell lines lack critical enzymes which participate in the biosynthetic pathway of the nucleotide-sugar precursor CMP-SA (cytosine mono phosphate sialic acid). Because of this, these cell lines are incapable of adding sialic acid residues to N-linked oligosaccharides.[154]

Insect cell lines have also been reported to lack certain transport proteins of nucleotide sugars (the co-substrates of the sugar-addition reactions within the Golgi apparatus) on the surface of the Golgi membrane. This reduces the concentration of co-substrates of the sugar addition reactions which, in turn, results in the synthesis of incompletely processed oligosaccharide structures by these cell lines.[156] Additionally, the low activity or absence of certain glucosyl transferase enzymes (such as GnTII with ~1% of the activity of mammalian GnTII) in the Golgi apparatus,[157] along with the presence of enzymes which limit (such as GlcNAcase) or impede the synthesis of human-type glycoforms has also been reported within insect cell lines.[158]

Despite the broad differences between the glycosylation patterns of alternative expression systems and mammalian cell lines, many promising advances have been achieved. Examples of these advances are the production of mammalian-like oligosaccharides in yeast cells,[150, 151, 159, 160] galactosylated oligosaccharides in insect cell lines,[161] galactosylated oligosaccharides in

plants,[162] controlled glycosylation of mAbs in plants[152] and optimised glycosylation of mAbs in plants.[163]

Table 4. Impact of bioprocess inputs on protein N-glycosylation[23]

Process Condition	Product/Cell Line: Effect	Proposed cause(s)
Nutrient availability	mAb/human hybridoma: change in microheterogeneity when glucose is substituted by fructose, mannose and galactose	Differential efficiency with which the cells consume the different carbon sources[105]
	IFNγ/CHO: ↓ site occupancy on at glucose<0.3mM	Low biosynthesis of UTP and glucosamine 6- phosphate[106, 107]
	IFNγ/CHO: ↓ sialylation and ↑ high mannose glycans at 0.35mM <glucose<0.7mM and glutamine at 0.1mM	Decline of UDP-GlcNAc synthesis[108]
Dissolved O_2 (These effects depend on cell line and product)	EPO-Fc/CHO: no effect between 10-90% DOT, ↓ sialylation at 100% DOT	Mechanism not defined[112]
	hFSH/CHO: maximum sialylation at 100%DOT	Mechanism not defined[113]
	IL-2 glycovariant/BHK-21: no effect at 0% DOT	Mechanism not defined[114]
	IgG1 mAb/murine hybridoma CC9C10: ↑galactosylation with ↑ DOT	UDP-Gal increases with DOT due to increased flux through glycolysis[115]
Ammonia and pH (The effect between NH_4^+ and pH is synergic)	mPL-I/murine: ↓ site occupancy on at 9mM NH_4^+	Mechanism not defined[117]
	TNF-IgG/CHO: 40% ↓ in galactosylation and sialylation	↑pH inhibits GalT and SiaT activities[118] ↑pH causes GnTI, ManII, GalT and SiaT to mislocalize in Golgi[119, 120]
	IL-2/BHK-21: ↑ branching in at 15mM NH_4^+	↑Concentration of UDP-GlcNAc and UDP-GalNAc reduces the concentration of UDP-Gal[121]
	EPO/CHO: ↓ sialylation of at 10mM NH_4^+	↑Concentration of UDP-GlcNAc decreases transport of CMP-Neu5Ac into Golgi[111, 132]
	mPL-I/murine: 28%↓ sialylation at 6.8<pH<7.8	↑pH inhibits GalT and SiaT activities. ↑pH causes GnTI, ManII, GalT and SiaT to mislocalise in Golgi[117]
	Highest levels of IgG3 mAb/murine hybridoma galactosylation at pH=7.4, lowest levels of sialylation at pH=6.9	↑pH inhibits GalT and SiaT activities. ↑pH causes GnTI, ManII, GalT and SiaT to mislocalize in Golgi[123]
Viability, growth phase and temperature	EPO/CHO: ↓ branching at 32°C<T<37°C	↓ temperature decreases UDP-GlcNAc and UDP-GalNAc concentration[126]
Serum and lipid supplements	IFNγ/CHO: ↑ branching with lipid supplementation[129] IgG1/CHO: ↑ sialylation with serum-free medium[79] IL2/BHK-21: ↑ fucosylation and ↑ sialylation with serum-free medium[114] mAb/murine hybridomas: broader glycan distribution and ↑ galactosylation with BSA supplementation[130]	Mechanisms not understood because serum composition is unknown
pCO_2 and osmolality	IgG2/SP2/0: unaffected up to 250mmHg pCO_2 ↑ galactosylation at 320mOsm/kg and high pCO_2 25% ↓ in galactosylation at 195mmHg pCO_2 and 435mOsm/kg	Osmotic stress induces ↑pH which inhibits GalT and SiaT activities and causes them to mislocalize[131]
Stirring speed	tPA/CHO: 72% ↓ site occupancy at 40<S<200rpm	Shear stress increases protein syhtnesis wich reduces ER retention time[133]
Culture modes	I IFNγ/CHO: low dilution rates yield low site occupancy 0.5<D<0.8day^{-1}.	Concentration of metabolites inhibits GalT and SiaT activity and mislocalises these compounds[107]

This section has reviewed the bioprocess inputs that have been reported to impact protein N-linked glycosylation. It has also discussed many of the mechanism by which these inputs influence glycan composition and distribution. A summary of the impact of bioprocess conditions on N-linked glycosylation is presented in Table 4.

In the QbD context, this information should be used to define the design space of allowable process inputs that will ensure appropriate product quality. For example, the threshold for minimum glutamine availability during cell culture described by Nyberg et al.[107] and Wong et al.[108] can be used to establish a set of feeding strategies that ensure adequate site occupancy and antennarity of N-glycans. However, a fundamental consideration while establishing the design space under QbD is how combinations of different process inputs impact product quality. If the feeding strategies for glutamine were not to consider ammonia accumulation and the negative effects this would have on product quality,[118] the design space would be too lax and would not provide quality assurance. Therefore, all known combinations of process inputs should also be used to establish a better constrained design space that will lead to product quality assurance.

Although the interactions between process inputs and protein glycosylation have been reported, several of the underlying mechanisms responsible for these effects are still to be described mechanistically (e.g. dissolved oxygen, the synergy between ammonia and pH, serum supplementation and shear stress). It is therefore impossible to develop mechanistic mathematical models relating such inputs with protein glycosylation. Conversely, the mechanisms that have been described to a fuller extent are those which associate nutrient availability with protein N-linked glycosylation. More specifically, the mechanisms by which glucose and glutamine availability impact glycosylation have been clearly drawn and associated with cellular metabolism and the availability of key metabolites (nucleotide sugar donors – NSDs) for the glycan processing reactions in Golgi. For this reason, and given our objective of developing a *mechanistic* mathematical model that relates bioprocess conditions with mAb N-linked glycosylation, glucose and glutamine availability throughout cell culture were defined as the process conditions to link with mAb Fc N-linked glycosylation in this work.

2.6. Previous mathematical models for protein glycosylation

Despite the relevance of N-linked glycosylation, few attempts have been made to model this phenomenon. Concretely, four models have been formulated over the past decade of which one addresses the attachment of the precursor oligosaccharide to the nascent polypeptide,[164] and

the other three have focused on the oligosaccharide processing reactions within the Golgi apparatus.[31-33] In the context of this research project, it has seldom been reported for monoclonal antibodies to have considerable macroheterogeneity. Therefore, this review will focus on the mathematical models that have been developed to address glycan microheterogeneity.

Umaña & Bailey, 1997[31]

The first N-glycosylation model was developed by Umaña & Bailey in 1997. This group approximated the Golgi apparatus compartments to a series of four continually-stirred tank reactors connected in series. Within this framework, they defined a series of 33 oligosaccharides which participate in 33 reactions catalysed by eight different enzymes. In order to define the kinetics for each of the reactions, Umaña & Bailey did a thorough revision of the literature available on the properties of the enzymes they considered for their work. The purpose of their model was to gain qualitative insight on the distribution of oligosaccharide structures represented by the model, and to simulate the effects of over-expression of certain enzymes within the system. They found that over-expression of GnT enzymes (III, IV and V) generated clear variations on the distribution of oligosaccharide structures. Additionally, they observed considerable qualitative oligosaccharide profile changes due to changes in protein productivity which, in the case of CSTR reactors, would affect the residence time of the glycoprotein within the Golgi compartments.

Being the first attempt to mathematically model the glycosylation process within the Golgi apparatus, and that the purposes of their model were of a more qualitative nature, Umaña & Bailey made a series of simplifying assumptions. Among these simplifying assumptions, they considered few oligosaccharide structures and reactions and neglected the addition of fucose or sialic acid in their model. Additionally, when they developed their model little information on the mechanisms by which the enzymes work was available, which led Umaña & Bailey to use oversimplified enzyme kinetics. Finally, and likely also due to the aims of their mathematical model, they assumed no variation in the availability of nucleotide sugar donors within the compartments of the Golgi apparatus. This, while being useful for gaining qualitative knowledge on the distribution of oligosaccharides, isolated their system from extracellular conditions, thus impeding their model to represent external influences on the calculated oligosaccharide structures.

Krambeck and Betenbaugh, 2005[32]

Eight years went by until the next mathematical model for protein N-linked glycosylation was developed. This second effort was made by Krambeck and Betenbaugh in 2005. They also

considered the Golgi apparatus as a series of four CSTRs in series, but their model was more ambitious than Umaña & Bailey's since it considered the addition of core fucose, galactose, poly N-lactosamine and sialic acid into their model. The addition of these species increased the number of reactions to 22,871 and the number of possible oligosaccharide structures to 7,565. Krambeck and Betenbaugh based the kinetic parameters of their model on Umaña & Bailey's, but had to add parameters to account for the additional enzymes and reactions. As part of their strategy, this group adjusted enzyme concentrations within the Golgi apparatus to match obtained glycosylation profiles with experimental data. Once this was done, they increased protein productivity and observed the differences in the oligosaccharide profiles and to what these differences were due to. Finally, they evaluated which metabolic engineering strategies could be implemented to ensure that, despite increased productivity, oligosaccharide profiles of their protein would remain unchanged. Given the sheer size and complexity of their model, Krambeck and Betenbaugh also included assumptions into their modelling framework. They considered that all kinetic mechanisms were the same for all oligosaccharide processing enzymes. However, they improved the reaction rate expressions by adding competitive inhibition terms which account for different substrates competing for a single enzyme. Again, since the aims of this mathematical model were to assess possible genetic engineering strategies for adequate glycosylation, this group also assumed the concentration of nucleotide sugar donors to be constant throughout the Golgi, uncoupling glycosylation from cellular metabolism.

Hossler et al., 2007[33]

The most recent attempt to model protein N-linked glycosylation was published in 2007. Again, the aim of this modelling attempt was to gain insight into the physiology of glycoprotein-producing cells and to evaluate possible cellular engineering techniques to minimise glycoprotein microheterogeneity. Based on recent findings supporting the Golgi maturation model over the stationary or classical Golgi model,[165, 166] Hossler and collaborators modelled the system as 4 PFR reactors in series. This group considered 326 glycan species within their model, and do not report the number of reactions that participate in their system. Hossler et al. obtained kinetic parameters from the same sources the previous two studies had. Despite the similarities with the previous two models, this group generated additional insight into the modelling of this process by comparing the performance of two models, 4 PFRs and the 4 CSTRs. When comparing the results obtained from these models with experimental data, they found that the 4 PFRs present results closer to those obtained with the 4 CSTR model, which is further supports the Golgi maturation model. An additional strategy this group adopted was to optimise the proportion of terminally-processed oligosaccharide structures by adjusting the enzyme concentration and distribution along the different compartments of the Golgi apparatus and the

concentration of nucleotide sugar donors within each compartment. Finally, this group performed a sensitivity analysis on their model in order to assess which of its variables had higher impact on the calculated oligosaccharide profiles. Hossler et al. considered that all enzymes function with the same kinetic mechanism, and also included competitive inhibition terms into their enzyme catalysed reaction rate expressions. This group also assumed a constant concentration for nucleotide sugar donors within the compartments of the Golgi apparatus, which prevented accounting for metabolic effects on protein N-linked glycosylation.

The modelling efforts presented above have all contributed towards understanding the N-linked glycosylation process within the Golgi apparatus. However, because the objective of these research groups was to assess possible cell engineering strategies with their models, they do not consider cellular metabolism, and therefore, uncouple the glycosylation process from bioprocess conditions. However, it is clear from preceding sections of this thesis that a variety of bioprocess inputs influence protein N-linked glycosylation. In light of these limitations, the objective of this work was to bridge the gap and produce a mathematical model that mechanistically and quantitatively links nutrient availability with the glycosylation associated quality of mAbs by including cell culture dynamics and cellular metabolism.

2.7. Techniques for mAb glycosylation analysis

A crucial element required for all steps of QbD implementation is process analytical technology (PAT). As was discussed previously, PAT is "a system for designing, analysing and controlling manufacturing through timely measurements of CQAs and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality."[167] In order to implement QbD for the development of processes that integrate N-glycosylation quality into the mAb product, accurate and high-throughput methods to determine the glycosylation-associated CQAs of mAbs are fundamental. In line with this perspective, the need for rapid acquisition of information throughout the development of manufacturing processes has led to the adaptation, optimization and development of novel techniques for the analysis of glycoproteins and their corresponding oligosaccharides. This section will briefly discuss the techniques that have been used to determine, on a qualitative and quantitative basis, the oligosaccharide structures of mAbs. For further background, two reviews on the overall strategy for glycan preparation and analysis are available.[168, 169] Broadly, there are three levels of information pertaining to the oligosaccharides of glycoproteins:

Level 1: Identification of glycosylation sites (macroheterogeneity)

This first level of investigation identifies the possible glycosylation sites of the mAb. In the case of IgG-based mAbs, the general procedure of glycosylation site identification is simpler than for other glycoproteins which have higher degrees of macroheterogeneity. As described earlier, each IgG-based mAb molecule contains two conserved N-linked glycosylation sites on its Fc fragment – located on the Cγ2 domains of the Fc fragment (Figure 4).

The other possible N-linked glycosylation sites (appropriate Asn-X-Ser or Asn–X–Thr sequons) on the IgG molecule may occur on the variable regions of its Fab fragments. It has been estimated that between 10–20% of all IgGs contain Asn–X-Ser or Asn–X–Thr sequons in the variable regions of their Fabs.[72, 73] Figure 11 presents two ways in which the protein backbone of mAb molecules can be cleaved. If N-glycosylation sites were to be found on the variable regions of the mAb, these could be identified by these digests. Reduction and acidification[45, 170] and papain digestion[143, 171] have been performed on mAbs to analyze their macroheterogeneity. Additional glycosylation sites may appear in three arrangements: on the variable region of the light chain (V_L domain), on the variable region of the heavy chains (V_H domain) or on both. For the two first cases, mAb digestion with papain yields two Fab fragments and a single Fc fragment. Isolating each of these permits distinguishing the source of the oligosaccharide chains (whether they belong to the variable regions or to the constant region). For the third case ,which is less common, the mAb would have to be sequentially digested first by papain, and then

by chemical means to determine which oligosaccharide structures are present on each peptide fragment.

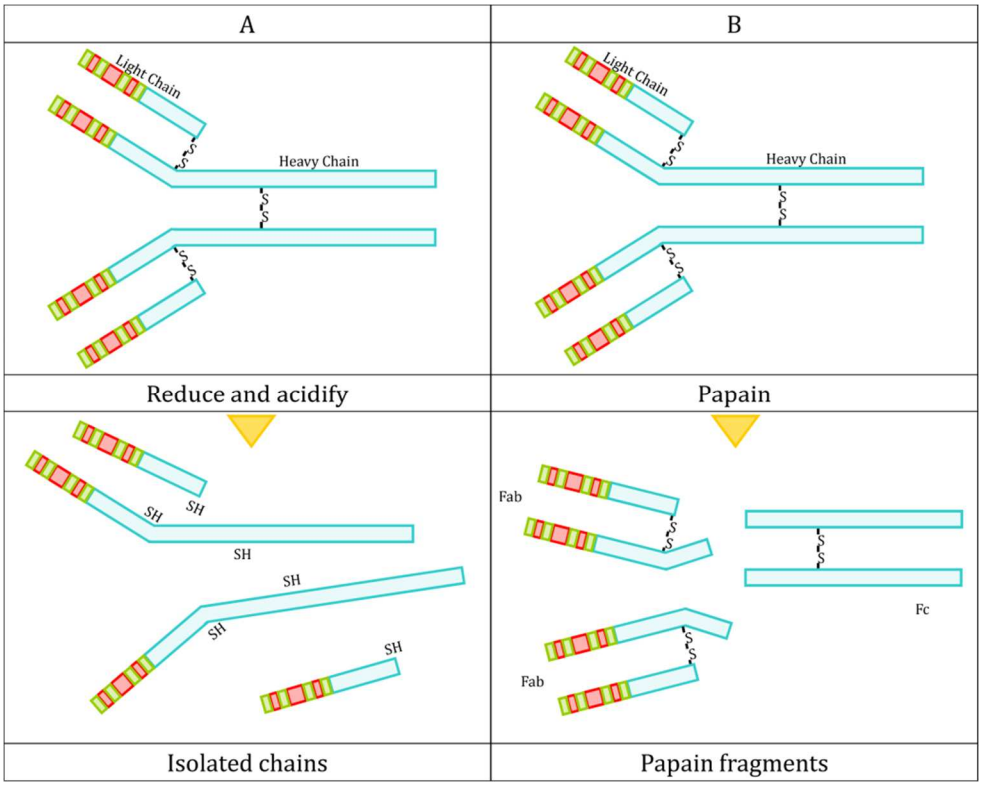


Figure 11. Cleavage of the peptide backbone of mAbs[29]
Reduction and acidification yield four peptide chains – two heavy chains and two light chains (A). Papain digestion produces three fragments – two Fab and one Fc (B).

Another attractive strategy for finding additional glycosylation sites on IgG-based mAb samples is to perform collision induced dissociation (CID) experiments through mass spectrometry.[45, 171] This method consists in randomly fragmenting glycoproteins while they are inside an ion trap (IT) cell of an MS instrument through bombardment with ions. The resulting mass spectrum of the fragmented sample yields information on the glycosylation sites of the molecule.

Level 2: Glycan profiling (microheterogeneity)

The second level of mAb glycoform characterization is to detect and quantify the different oligosaccharide structures bound to each of its glycosylation sites (mAb microheterogeneity due to glycosylation). Once the possible glycosylation sites have been identified and separated, the glycans are ready for analysis. At this point, two strategies can be used. The first investigates glycopeptides, which requires the protein backbone to be cleaved usually with trypsin to yield low molecular weight glycopeptides, which are then analyzed by the techniques described later in this section.[45, 58, 144, 172] The second strategy is to remove the oligosaccharides from the protein

backbone. This can be achieved by chemical means (e.g. hydrazinolysis); however, enzymatic cleavage is more frequently used. The most common enzyme employed for this task is Peptide N-glycosidase F (PNGase F), which is known to have broad specificity and releases all N-linked glycans from the peptide backbone of the IgG-based mAbs. The released oligosaccharides or glycopeptides are then ready to be analyzed by any of the different techniques presented below. The objective of the analytical techniques is to identify which oligosaccharide structures are present and their abundance within the mixture.

Level 3: Glycan structural analysis

The third and most detailed level of mAb glycosylation is structural analysis. Given that the linkages between the constituting monosaccharides of glycans may vary, these must also be characterized to obtain the intramolecular structures of the oligosaccharides bound to mAbs. In order to generate structural information, a series of sequential enzyme-catalyzed reactions that remove monosaccharides from the oligosaccharide, must be performed.[59, 78, 79, 109, 130, 142, 173] The key to finding the types of bonds that occur between the different monosaccharides are the enzymes that are used. Certain glycosidases have such high specificity that they will be able to cleave an α-1,3 glycosidic bond, but not an α-1,6. If the sequence of the reactions is planned properly, and the products (both monosaccharides and oligosaccharides) are quantified accurately by any of the analytical techniques presented below, the bond structure of the original glycan can be obtained. An additional strategy that has been used to analyze the structure of mAb glycans has been CID.[45, 58, 144, 174-176] Similarly to glycopeptides, cleaved oligosaccharides can be ionized, trapped and collided with ions to generate oligosaccharide fragments. On these fragments, so-called parent ion experiments can be performed to obtain structural details on the linkages of the glycans.

So far, only general strategies for glycosylation analysis have been discussed. In the next section, the specific analytical techniques that have been used for the analysis of mAb glycans are discussed. We categorise these analytical techniques into two groups: mass spectrometric techniques; and separation-based techniques. A brief overview of the individual analytical techniques will be presented in the following sections.

2.7.1. Mass Spectrometry (MS)

Mass spectrometric techniques are characterized by the three main phenomena by which they function. MS is based on a) ionizing the molecules which compose the sample, b) separating them as a function of their mass to charge ratio, and c) detecting and quantifying them. The detection modules of current MS systems have been optimized for sensitivity and accuracy, and

hence do not have much impact on the quality of data. For this reason, this section will only address the most commonly used ionization and separation methods used in MS-based analytical techniques for mAb glycan characterization.

Ionization Techniques

The purified and, in some cases, derivatized glycan or glycopeptide sample must be ionized for analysis by MS. Weak ionization techniques should be used for the analysis of oligosaccharide structures because they are complex and fragile. The two main ionization techniques used for the characterization of mAb oligosaccharides are:

Matrix Assisted Laser Desorption Ionisation (MALDI)

This method is commonly used to ionize glycans that have been cleaved from the protein backbone of the mAb. Prior to ionization by this method, it is common for the cleaved oligosaccharides to be derivatized by permethylation.[142, 174] This derivatization of the oligosaccharide structures presents considerable advantages, namely that it avoids fragmentation of the sample during the ionization process, it allows for quantification of both neutral and charged (sialic acid-containing) oligosaccharide species simultaneously and improves the detection sensitivity of molecular ions.[177] Once the oligosaccharides are prepared, they are embedded into a low molecular weight, UV absorbing crystalline matrix, which, in the case of glycans is usually 2,5-dihydroxybenzoic acid (2,5-DHB). A pulsed laser is focused onto the sample/matrix mixture where most of the energy from the laser pulse is absorbed by the matrix, and enough energy is transmitted to the sample molecules to ionize them. The process by which the ionized sample is released from the matrix is still not fully understood, but is believed to be similar to flash evaporation. MALDI is known as a low energy ionization method because it causes little or no fragmentation while producing singly charged ions.

Electrospray Ionisation (ESI)

This ionization method can be applied to released oligosaccharides, but has also been used for analysis of glycopeptides. The sample, while in solution, is injected into a narrow capillary. Pressure is applied to the injection end of the capillary so that its content is dispersed as an aerosol of small droplets. The tip of the sprayer is generally coated with gold and a high voltage is applied to it so that when the sample is sprayed charge is conferred to the droplets. Once this occurs, the aerosol of highly charged droplets is passed through a drying gas to aid solvent evaporation. The charged droplets become smaller until the sample molecules are solvent-free and charged. This ionization method is also low energy, but in contrast to MALDI, it may generate multiply charged ions.

Mass Analysers

Once the oligosaccharides are ionized, they are transferred towards the mass analyzer of the instrument. The mass analyzers separate the species within the sample according to their mass to charge ratio. The most common mass analyzers used for mAb glycan characterization have been:

Quadrupole (Q) Mass Analyzers

Quadrupole mass analyzers, which are also named quadrupole mass filters, have low resolution, and because of this, two or more of them are used in series or are also used in conjunction with other mass analyzers to obtain higher resolving power. In the specific case of mAb glycan characterization, several quadrupole mass filters are arranged in series to perform CID (MS/MS) experiments or are also coupled in tandem arrangements to TOF mass analyzers.[45, 144, 178-180] Q mass filters use a quadrupolar electrical field, which consists of both radio frequency (RF) and direct current (DC) components to separate ions. At given field intensities, ions with certain *m/z* ratios become resonant and are thus trapped within the field, while other ions with different *m/z* ratios do not resonate and escape the field. The quadrupolar electrical field can be manipulated by changing the magnitudes of its RF component. When coupled to TOF mass analyzers, quadrupole mass filters are used as sieves in which the ions with a desired range of *m/z* are condensed and sent towards the TOF analyzer while all other ions outside the set range leave the field and never reach the TOF system.

Time of Flight (TOF) Mass Analyzers

These analyzers are the most commonly used for the characterization of mAb glycans. Linear TOF analyzers consist of an electrical field-free tunnel through which the ionized species travel at a given velocity. In turn, the velocity of each charged molecule that travels through the TOF analyzer is a direct function of their mass and charge. For example, if a mixture contains two different oligosaccharide structures A and B with molecular weights of 1,400 g/mol and 1,600 g/mol respectively, and both are singly charged, their mass to charge ratio *m/z* will be 1,400 and 1,600. Within the TOF analyzer, the velocity of A will be higher to that of B because it has lower mass, and will reach the detector in less time than B. A contrasting example would occur if A is singly charged, but B is doubly charged. The *m/z* value for A would remain as 1,400, but the one for B would be 1,600/2 = 800. In this case, B would have higher velocity along the TOF tube and reach the detector earlier than A. Linear TOF systems have poor resolution and mass accuracy. The modern TOF analyzers contain systems called reflectrons which can reverse the flight path of ions. Reflectrons increase the flight distances within the TOF analyzer, thus

increasing resolution, and they also correct kinetic energy distribution effects, which greatly improves mass accuracy as well.

2.7.2. Separation-based methods for glycan analysis of mAbs

The separation techniques for oligosaccharide analysis are somewhat analogous to MS techniques in the sense that the components of the sample mixture are separated and later quantified. However, the principle by which the sample components are separated does not depend on the *m/z* ratio. Instead, different properties of the sample components (such as size, polarity and ionic interactions) are used to differentially separate them so that they can be detected. Additionally, the means by which glycans are detected by these techniques influences the accuracy and sensitivity of the method. There are four predominating separation-based techniques that have been used for mAb glycan characterization:

High Performance Anion Exchange Chromatography coupled to Pulsed Amperometric Detection (HPAEC-PAD)

In HPAEC-PAD, which is a type of adsorption liquid chromatography, the sample molecules are differentially adsorbed onto a positively charged solid matrix by ionic interactions. Alkaline buffer is added to the sample to ensure that its pH is higher than the pKa of the oligosaccharides which are to be separated, thus conferring them a net negative charge. The high pH sample is then loaded into the column leaving the negatively charged analytes to interact with the positively charged column matrix. The elution step consists of injecting a solution of so-called gradient ions onto the column so that the oligosaccharide molecules are differentially desorbed from the matrix. This differential desorption of the oligosaccharides from the matrix gives them different elution times, thus permitting them to be separated for detection. Pulsed amperometric detection is based on measuring the current generated by oxidation or reduction reactions on the surface of an electrode.[181]

This detection method is highly sensitive, but in order to directly measure concentration of oligosaccharides, reference compounds must be used to generate calibration curves. This is because the current that is produced by a single mole of analyte varies from compound to compound, and is not proportional to their concentration within the sample mixture. For example, two OH groups of galactose can be oxidized under the detection conditions, but only one OH of mannose is oxidized. Therefore higher response would be measured for galactose than mannose even when they appear at the same concentration. This problem can be overcome by producing calibration curves for each of the compounds that are being analyzed. When PAD is used to detect complex oligosaccharides, it may be difficult to produce the

calibration curves due to lack of reference compounds. In these cases, only relative quantification is possible.[182]

RP and HILIC Liquid Chromatography (LC)

Liquid chromatography techniques separate the components of a mixture based on the differential partition of components, due to hydrophilic and hydrophobic interactions, between a mobile phase and a stationary phase. In the context of mAb separation and quantification, detection is usually made by photometric detectors which measure either UV absorption or fluorescence. This section shall first discuss the different separation modes that are used in LC and will conclude with a brief overview of how oligosaccharides are detected when analyzed by these methods.

Reversed-phase high performance liquid chromatography (RP-HPLC)

Reversed-phase liquid chromatography refers to when the mobile phase is polar (hydrophilic) and the stationary phase is non-polar (hydrophobic). Depending on the polarity of the different analytes within the mixture, they interact more strongly with either phase and thus are separated for eventual detection and quantification. In RP-HPLC, the dissolved sample is injected at high pressure (~1,500 psi) into a column packed with non-polar media (commonly, C_{18}) octadecylsilyl columns). It is worth noting that oligosaccharide structures, being composed of monosaccharide (carbohydrate) units, tend to be quite polar due to the large amount of hydroxyl groups that are substituted along their carbon backbone. If these hydroxyl groups were to be left intact, the oligosaccharide structures would interact very weakly with the non-polar stationary phase and elute almost immediately with the polar mobile phase, yielding little or no separation. Because of this, oligosaccharides that are to be separated by RP-HPLC, are generally derivatized by permethylation. This process substitutes the hydrogen on the hydroxyl group of the constituting carbohydrates with methyl groups, thus reducing their polarity and greatly increasing stability (especially of sialylated oligosaccharides if MS analysis will be done on the effluent of the LC system). By reducing their polarity, permethylation increases interaction of the oligosaccharides with the non-polar (hydrophobic) stationary phase, and thus, their separation is more efficient. Once the oligosaccharide molecules are separated by HPLC, they are usually detected by conventional UV absorption methods.

Hydrophilic interaction liquid chromatography (HILIC-HPLC or HILIC-UPLC)

Hydrophilic interaction chromatography is neither normal phase (NP) nor reverse phase (RP) chromatography. NP chromatography is the opposite of RP because in it, the mobile phase is non-polar and the stationary phase is polar. In contrast to both, hydrophilic interaction

chromatography uses a relatively polar mobile phase (acetonitrile) and a polar stationary phase (e.g. silica derivatized with amide groups). Instead of having the analytes separate as a result of their differential partition between polar and non-polar media, HILIC separates compounds as a function of their differential capacity of producing hydrogen bonds with the polar media of the stationary phase. HILIC techniques for the characterization of oligosaccharides have been established for conventional HPLC systems (1,500 psi, column packed with particles between 3 and 5 μm). However, the advent of new technologies for producing columns packed with smaller particles (1 to 2 μm) has permitted systems which operate at much higher pressures (14,000 psi). At these pressures, the process is named ultra performance liquid chromatography (UPLC). UPLC systems have improved resolution, speed and sensitivity compared to their lower pressure counterpart, HPLC.[183] Additionally, new UPLC columns for glycan analysis are being developed and tested. These columns have been used to completely separate and quantify mAb glycans (including structural isomers) in 30 minutes[183] which is a considerable gain compared to the 50 minute run time recently achieved by rapid resolution RP-HPLC.[184]

As was mentioned earlier, when LC techniques are used to analyze oligosaccharides, an important factor is detection and quantification. Mass spectrometry can only yield relative quantification of oligosaccharides (e.g. fraction of species A with respect to the most abundant species detected), and in some cases PAD detectors have the same limitation. In contrast, detection and quantification in LC is commonly done either by UV absorbance or fluorescence, and both of these photometric methods ensures direct quantification of the analyzed molecules as long as the measured signal falls into the linear range of the detectors.

Paradoxically, carbohydrates have no chromophores (usually conjugated double bonds), and thus are undetectable by UV or fluorescence photometers. The strategy that has been developed to overcome this issue is to label oligosaccharides (or their constituting carbohydrates) with UV-absorbing molecules. In the specific case of N-linked glycans, derivatization techniques have been developed through which a single oligosaccharide is bound with a single label molecule, and these methods have focused on reductive amination chemistry. The GlcNAc residue at the base of free complex oligosaccharides contains an exposed reducing carbonyl group. The carbon on this reducing group is reacted, under reducing conditions, with the amino group of the label molecule, thus generating a stable, secondary amine. These types of reactions ensure that each oligosaccharide is bound to a single label. As expected, the most commonly used labels for glycan quantification contain primary amine groups and are: 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), 3-(acetylamino)-6-aminoacridine (AA-Ac), 2-aminopyridine (AP), 8-aminopyrene-1,3,6-trisulfonate (APTS) and 1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB).[185, 186]

Table 5. Commonly used UV and fluorescent labels for quantification of glycans[29]

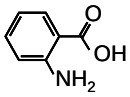
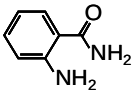
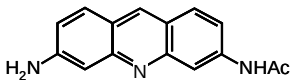
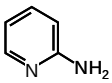
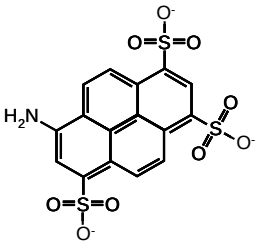
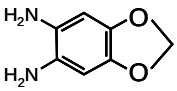
NAME	FORMULA	ABSORBANCE/FLUORESCENCE
2-aminobenzoic acid (2-AA)		UV: 217, 247 and 335 nm λ_{exc}: 360 nm λ_{em}: 425 nm
2-aminobenzamide (2-AB)		UV: 213 and 248 nm λ_{exc}: 330 nm λ_{em}: 420 nm
3-(acetylamino)-6-aminoacridine (AA-Ac)		UV: 272, 377 and 432 nm λ_{exc}: 382 or 445 nm λ_{em}: 520 nm
2-aminopyridine (AP)		UV: 240 nm λ_{exc}: 310 nm λ_{em}: 380 nm
8-aminopyrene-1,3,6-trisulfonate (APTS)		λ_{exc}: 315 nm λ_{em}: 512 nm
1,2-diamino-4,5-methylenedioxybenzene dihydrochloride (DMB)		λ_{exc}: 373 nm λ_{em}: 448 nm

Table 5 shows the structures of these commonly used labels along with the maximum of their spectral peaks both for absorbance and fluorescence. The labels can then be detected by common spectrophotometers or by fluorescence inducing spectrophotometers. These labels allow oligosaccharides to be detected with sensitivities 500 to 1,000 times higher than the non-labeled glycans.[187] Analytical technique sensitivity is crucial for the implementation of QbD to mAb manufacturing processes because a) in these processes, the amount of material that is to be analyzed is limited and b) higher sensitivity of the analytical method allows monitoring of oligosaccharides that are present in lower abundances, and thus be included in the development of the manufacturing process.

Capillary electrophoresis with laser induced fluorescence (CE-LIF)

This method is quite different from the liquid chromatography techniques described above. The principle by which capillary electrophoresis separates the components of a mixture is by their size to charge ratio. The sample, while dissolved in an appropriate buffer, is fed into a capillary, which has an electrical potential applied to its ends. The voltage induces the components of the sample to migrate at rates which depend on their charge and size. In many cases, the capillary is filled with a gel which enhances separation resolution by sieving the components as a function of their size. Detection in CE systems is quite similar, if not identical, to fluorescence detection in LC methods. Therefore, when analyzed by this technique, it is necessary for the oligosaccharides

to be labeled with fluorescent molecules as well. A commonly used label for CE-LIF is APTS.[178, 188, 189] CE-LIF has high resolving power in smaller time frames given the nature of the separation, and similarly to LC methods, this technique is also very sensitive given its use of fluorescent labels.

Fluorophore-assisted carbohydrate electrophoresis (FACE)

FACE is a gel-based analytical technique developed for the separation and quantification of carbohydrates and structures. The mechanism by which FACE separates carbohydrates is by their charge and size. (165).[190] The principle by which this technique works is similar to CE, and also requires labeling with fluorescent compounds. However, FACE usually entails fluorescent labeling with compounds that are charged, such as ANTS or APTS, to ensure that the species is charged for adequate migration through the gel. FACE, over time, has been replaced by other techniques, such as CE-LIF and LC because they have been found to have higher resolving power, more reproducible detection and higher throughputs due to their automation. The replacement of FACE by CE and other techniques can be inferred from Table 6. The three cases in which FACE has been used for mAb oligosaccharide characterization were Kunkel et al.,[115] Leibiger et al.,[59] and Saba et al.[175]

2.7.3. Combination of analytical techniques (LC-MS and CE-MS)

A very important advantage of separation-based analytical techniques in the context of mAb glycan characterization is that they can be coupled to MS via an ESI interface. An underlying issue of all non-MS oligosaccharide characterization methods is the identification of all separated species. As mentioned previously, the amount and purity of standard oligosaccharides is scarce, thus when employing a single technique for oligosaccharide characterization, there is a relative degree of uncertainty concerning both absolute quantification and identification of the oligosaccharide structure. The manner in which these limitations have been overcome is by combining two or more of these techniques. In mAb glycan analysis (and usually for all other glycoprotein characterization), the most common example of this combination has been the on-line connection of liquid chromatography or capillary electrophoresis techniques to mass spectrometry through an ESI interface. This combination of analytical techniques permits separation and absolute quantification of oligosaccharide species by LC or CE (through fluorescence or absorption) and identification (or identity confirmation) of each species which has been separated by mass spectrometry. Examples of the combination of these techniques towards the characterization of mAb glycans are CE–LIF–ESI–MS[189, 191] and RP–HPLC–ESI–MS.[45, 75, 142-144, 172, 175, 178, 180, 190, 192, 193]

2.7.4. Multimethodological approach

The need for obtaining all three levels of glycan information has led to the proposal of multimethodological approaches for glycoprotein oligosaccharide characterization (168). In these approaches, a combination of analytical techniques is performed for the full characterization of oligosaccharides.

Level 1, which corresponds to the determination of glycosylation sites of the glycoprotein (macroheterogeneity), can be covered by generating glycoprotein digests (as shown in Figure 11) and feeding the digests into a tandem CE–MS or LC–MS system. These systems not only have the resolution necessary to distinguish the mass of the protein backbone to which the oligosaccharide is bound, but they can also determine differences as to which oligosaccharide structure is bound to each fragment of protein backbone.

Level 2, which concerns the oligosaccharide structures and concentrations (glycan profile), can be covered by applying HPAEC, CE–LIF or LC–fluorescence/absorbance for the separation and quantification of each oligosaccharide species. These techniques are coupled to MS in order to identify the oligosaccharide species. Additionally, these results (both the quantification and the identification of the oligosaccharide profiles) can be confirmed by monosaccharide composition analysis, which consists in cleaving all the glycosidic bonds within the oligosaccharide, thus producing a pool of monosaccharides which can then be analyzed by CE–LIF, LC–fluorescence/absorbance, HPAEC–PAD, MALDI–TOF or ESI–MS. The concentrations of monosaccharides and their relative proportions within the pool should match their stoichiometric abundance within the oligosaccharides which were measured previously.

Finally, *Level 3*, which corresponds to oligosaccharide structural analysis and linkage characterization, can be obtained by sequential glycosidase digestions of the oligosaccharides, and identifying and quantifying the products of each digest by CE–LIF, LC–fluorescence/absorbance, HPAEC–PAD, MALDI–TOF or ESI–MS. Another approach for obtaining level 3 information is by performing fragmentation analyses with MALDI–Ion Trap–MS or ESI–Ion Trap–MS. Ion traps are similar to quadrupole mass analyzers since they can be tuned to retain oligosaccharide ions with a set *m/z* range. However, they are designed to trap said ions instead of sieving them as the quadrupole mass analyzer does. Once the charged oligosaccharide ions are trapped, they can be fragmented by collision with other ions. The manner in which the ionized gycans are fragmented yields structural information on the original oligosaccharide structure. Table 6 presents the techniques for mAb glycosylation analysis that have been used over the past fifteen years.

Table 6. Analytical techniques for mAb glycosylation[29]

YEAR	MAB FORMAT	PRODUCTION CELL LINE	OLIGOSACCHARIDE CHARACTERISATION	REFERENCE
1993	Chimeric IgG1, IgG2, IgG3 and IgG4	CHO and mouse J558L	HPAEC-PAD for relative glycan quantification Neutral oligosaccharides labelled with 2-AP for RP-HPLC for quantification	(134)
1994	Human IgG1	NS0	HPAEC-PAD for relative glycan quantification	(138)
1995	CAMPATH-H1: Humanized IgG1κ, anti-CDw52	CHO, Y0 and NS0	HPAEC-PAD for relative glycan quantification sequential exoglycosidase treatment and MALDI-TOF-MS for structural analysis	(79)
1997	CAMPATH-H1: Humanized IgG1κ, anti-CDw52	CHO and NS0	Permethylated glycans were directly injected into ESI-Q-MS for CID structural characterisation MALDI-TOF-MS for glycan identification and relative quantification exoglycosidase treatment and MALDI-TOF-MS for structural analysis	(142)
1998	Murine IgG2b	C3H/SW cell line	RP-HPLC-ESI-Ion Trap-MS/MS for glycopeptides identification and relative quantification	(58)
1998	Murine IgG1	CC9C10 cell line	ANTS-labelling for FACE HPAEC-PAD for relative glycan quantification	(115)
1999	Human CBGA1 IgG1λ	Human-mouse hybridoma	ANTS-labelling for FACE HPAEC-PAD for relative glycan quantification exoglycosidase treatment-HPAEC-PAD for structural analysis	(59)
1999	Chimeric IgG1κ (Rituximab)	CHO	APTS-labelling and CE-LIF for glycan quantification	(194)
1999	Chimeric IgG1	CHO	MALDI-TOF-MS for glycan identification and relative quantification	(83)
2000	Chimeric IgG1-TNFR (immunoadhesin tumor necrosis factor) fusion protein	CHO	MALDI-TOF-MS for glycan identification and relative quantification HPAEC-PAD for relative monosaccharide quantification OPD (O-Phenylenediamine)-labelling and RP-HPLC for Neu5Ac quantification	(118)
2000	Murine IgG1κ	Sp2/0	HPAEC-PAD for relative glycan quantification	(195)
2000	Chimeric IgG1	CHO	2-AB labelling and NP-HPLC for glycan quantification Exoglycosidase treatment and NP-HPLC for structural analysis	(109)
2001	Murine IgG1 (MGR48)	Tobacco plants	MALDI-TOF-MS for glycan identification and relative quantification	(196)
2001	Humanized IgG1	NS0	MALDI-TOF-MS for glycan identification and relative quantification Permethylated glycans and RP-HPLC-ESI-Q-TOF-MS/MS with CID for structural analysis	(174)
2001	Humanized IgG1	GSNS0	2-AB labelled for NP-HPLC for glycan quantification Exoglycosidase treatment and NP-HPLC for structural analysis	(78)
2002	Murine IgG1κ	Sp2/0 (CC9C10)	APTS-labelling for FACE HPAEC-PAD for relative glycan quantification PMP-derivatised for MALDI-TOF-MS for glycan identification NP-HPLC-ESI-Q-TOF-MS/MS for structural analysis	(175)
2004	Murine IgG1	Murine hybridoma	2-AB labelling and NP-HPLC for glycan quantification Exoglycosidase treatment and NP-HPLC for structural analysis	(173)
2004	Murine IgG1κ	Tobacco plants, Sp2/0	2-AB labelling and NP-HPLC for glycan quantification Exoglycosidase treatment and NP-HPLC for structural analysis	(154)

2005	Murine IgG1κ	N/A	APTS-labelling and CE-LIF for intact oligosaccharide and free monosaccharide quantification MALDI-TOF-MS for glycan identification RP-HPLC-ESI-Q-TOF-MS for site occupancy analysis	(178)
2005	Humanized IgG1κ (7H2HM and A2CHM)	7H2HM produced in CHO and A2CHM produced in NS0cell lines	MALDI-TOF-MS for glycopeptide identification and relative quantification RP-HPLC-ESI-TOF-MS for glycopeptide structural analysis and relative quantification Papain digestion and hydrophobic interaction LC for site occupancy analysis	(143)
2006	Human IgG1κ	Aquatic plant *Lemna minor*	2-AA labelling and NP-HPLC-ESI-Q-TOF-MS for glycan quantification and identification HPAEC-PAD for relative monosaccharide composition MALDI-Q-TOF-MS for glycan identification	(163)
2006	Humanized IgG1k (trastuzumab), Chimeric IgG1k (rituximab) and humanized IgGk (palivizumab)	CHO and NS0	2-AA labelling and CE-LIF for glycan quantification Oligosaccharide fractions separated by RP-HPLC and MALDI-Q-TOF-MS/MS for structural analysis	(176)
2006	Murine IgG1	Murine hybridoma	HPAEC-PAD for relative glycan quantification	(136)
2006	Recombinant IgG	N/A	RP-HPLC-ESI-Q-TOF-MS/MS for intact mAb structural analysis MALDI-TOF-MS for glycan identification 2-AA labelling and HPAEC-PAD for relative glycan quantification	(180)
2007	Human IgG1κ	Human myeloma cell lines	RP-HPLC-ESI-TOF-MS with orthogonal acceleration for glycoprotein structural analysis 2-AP labelling and NP-HPLC for glycan quantification	(75)
2007	IgG1 and IgG2	CHO	2-AB labelling and RP-HPLC-ESI-Q-TOF-MS for quantification and on-line glycan identification RP-HPLC-ESI-IT-MS for structural analysis	(192)
2007	Murine IgG1	BCF2 hybridoma	APTS labelling and CE-LIF for glycan quantification Exoglycosidase treatment and CE-LIF for structural analysis	(130)
2007	Chimeric IgG1κ (Cetuximab)	Sp2/0	2-AA labelling and NP-HPLC for glycan quantification MALDI-Q-TOF-MS for glycan identification MALDI-Q-TOF-MS/MS for structural analysis	(179)
2008	Chimeric IgG1κ (rituximab), humanized IgG1κ (omalizumab, trastuzumab, anti-HIV) expressed in CHO, Chimeric IgGκ (cetuximab, infliximab) expressed in Sp/0, humanized IgG1κ (daclizumab) expressed in NS0 and anti-HIV IgG1κ expressed in human cell line	CHO, Sp/0, NS0 and human cell lines	RP-HPLC-ESI-Q-TOF-MS/MS for glycopeptide structural analysis and relative quantification Borohydrate-reduced and Porous graphitic carbon (PGC)-HPLC-ESI-Q-TOF-MS for glycan identification 2-AB labelling and NP-HPLC for quantification	(45)
2008	N/A	CHO	Papain digestion of mAbs and RP-HPLC-ESI-Q-TOF-MS/MS for glycopeptide identification, site occupancy and relative quantification	(171)

2008	Humanized IgG1κ (A2CHM, A3BHM) and murine IgG1κ (7C10)	NS0, CHO and murine hybridoma	RP-HPLC-ESI-TOF-MS for relative glycoprotein identification and relative quantification ESI-IT-MS/MS for glycan structural analysis Nano-RP-HPLC-ESI-IT-MS/MS for glycopeptide structural analysis and relative quantification	(144)
2008	IgG1	N/A	APTS labelling and CE-LIF-ESI-TOF-MS for glycan quantification and identification	(189)
2008	IgG1κ	Mouse hybridoma	Intact mAb and RP-UPLC-ESI-Q-Ion Mobility (IM)-TOF for glycoprotein identification and relative quantification Reduced mAb and RP-UPLC-ESI-Q-IM-TOF for site occupancy, glycoprotein identification and relative quantification Nano-RP-HPLC-ESI-Dual Ion Trap-MS/MS for glycopeptide structural analysis	(170)
2009	Human IgG1κ	NA	HPAEC-PAD for relative monosaccharide quantification HPAEC-PAD for relative glycan quantification APTS labelling and CE-LIF for glycan quantification RP-HPLC-ESI-IT-MS for glycan identification and relative quantification RP-HPLC-ESI-IT-MS for glycopeptide identification and relative quantification	(172)
2009	N/A	N/A	3-AA-Ac labelling for NP-UPLC-ESI-TOF-MS for glycan quantification and identification (30 min)	(183)
2009	Humanized IgG1κ (trastuzumab), Chimeric IgG1κ (rituximab) and humanized IgGκ (palivizumab)	CHO and NS0	9-fluorenylmethyl chloroformate (Fmoc-Cl) labelling and CE-LIF-ESI-IT-MS/MS for glycan quantification, identification and structural analysis	(191)
2009	IgG1, IgG2, IgG3 and IgG4	CHO	2-AB labelling and RP-HPLC-ESI-Q-TOF-MS for glycan quantification and identification (45 min)	(184)
2009	Humanized IgG1κ (trastuzumab)	CHO	Intact mAb and RP-UPLC-ESI-Q-Ion Mobility (IM)-TOF for glycoprotein identification and relative quantification Reduced mAb and RP-UPLC-ESI-Q-IM-TOF for site occupancy, glycoprotein identification and relative quantification Nano-RP-HPLC-ESI-Dual Ion Trap-MS/MS for glycopeptide structural analysis	(197)
2010	Human IgGs (HA78 anti-hepatitis A; 2G12 anti-HIV)	*Arabidopsis thaliana*	RP-HPLC-ESI-Q-TOF-MS/MS for glycopeptide structural analysis and relative quantification Borohydrate-reduced and Porous graphitic carbon (PGC)-HPLC-ESI-Q-TOF-MS for glycan identification 2-AB labelling and NP-HPLC for quantification	(198)
2011	N/A	N/A	Zwitterionic-HILIC-HPLC-ESI-Q-TOF of unlabelled glycans for relative quantification and identification HILIC-HPLC-ESI-Q-TOF of 2-AB labelled glycans for quantification and identification	(199)
2011	Monoclonal human IgG1; monoclonal mouse IgG1; polyvalent human IgG (Redimune©); 16 licenced mAbs	Human; mouse myeloma; CHO; NS0; Sp2/0	APTS labelling and CE-LIF for glycan quantification	(200)
2012	Humanized anti-EGFR IgG1	NS0	2-AB labelling and NP-HPLC-ESI-Q-TOF for glycan quantification and identification Exoglycosidase treatment for glycan structural analysis MALDI-TOF-MS for glycan identification	(201)

There have also been multiplexed analysis methods where parallel use of analytical techniques greatly improves N-glycan profiling.[202, 203] These methods have relied on performing parallel lectin microarrays and mass spectrometry to identify specific glycoforms of different proteins at high-throughput.[202] An interesting example of glycan analysis using multiplexed analytical techniques was reported recently by Mittermayr et al.[203] This group found that when using HILIC-UPLC in parallel with CE-LIF, complete structural annotation of the glycans bound to human polyclonal antibodies was performed in less than 20 minutes. They attributed the short time in glycan structure annotation to the convenient orthogonality of both analytical techniques.

2.7.5. Glycosylation databases

The techniques described in the previous section produce rather complex datasets. Because of this, a major challenge underlying glycan analysis is data interpretation. As was mentioned in previous sections of this thesis, MS generates spectra where observed peaks correspond to different mass to charge ratios (m/z). By knowing – or in some cases, assuming – the possible molecular ion charges (z) produced by the ionisation method, the peaks can be assigned to a molecular weight (MW). The obtained MWs must then be compared to glycans that have either been previously observed or that are feasibly produced by known glycan biosynthetic pathways.[204]

Similarly, glycan data generated with the separation-based methods described in the previous sections must also be interpreted. The peaks generated by these methods must be assigned to a glycan structure. Because separation efficiency may vary on a day-to-day basis and is also heavily dependent on the analytical equipment itself, peak assignment is challenging. In order to address this issue, methods where the retention times of observed peaks are correlated with those obtained for an internal standard (2-AB labelled dextran ladder), have been developed.[205, 206] This methodology allows for accurate assignment of glycan species, the product of sequential glycosidase digestions and – rather importantly – has made results generated with different equipment comparable.[168, 206]

Underlying the interpretation of glycan analysis data is the need of detailed databases where all possible (or previously observed) glycan species are reported in such a way that measured glycan data can be readily compared and interpreted. In addition to the databases, software that annotates experimental data by comparing it to the available glycan profiles from databases is also necessary. Several glycan databases and glycan analysis software tools have been developed and made available for glycan data interpretation. Currently available glycan

databases and software have been reviewed recently,[207, 208] and therefore, this section will briefly summarise the most relevant ones to this thesis.

CCSD (CarbBank)(http://www.glycome-db.org/getDownloadPage.action?page=carbbank)[209]

This was the first carbohydrate database and was established in the early 1990s. Its latest release contains >50,000 records which link carbohydrate structure, the organism that synthesised it, along with the analytical methods that were used to determine the structure. Maintenance and additions to the database were discontinued in 1997, but it is still available as part of newer carbohydrate databases.

GlycomeDB (http://www.glycome-db.org/)[210, 211]

This is the latest glycan database, and links all the major publicly-available databases. This compiled database contains over 35,000 unique glycans along with taxonomic annotations. GlycomeDB database is updated weekly by retrieving new data that has been uploaded to its associated databases.

KEGG Glycan (http://www.genome.jp/kegg/glycan/)[212, 213]

This database provides data on glycan structures and associates them with the metabolic pathways involved in their biosynthesis. This database includes >10,000 glycan structures and is able to associate each of them with the necessary genome, transcriptome and proteome required for their synthesis.

GlycoBase 3.0 (http://glycobase.nibrt.ie/glycobase/show_nibrt.action)[214]

This database contains experimental HPLC and UPLC elution times for >380 2AB-labelled N-glycans with respect to the dextran ladder internal standard. As was mentioned previously, the normalisation of the retention time with respect to the dextran ladder eliminates day-to-day and equipment-associated variation.

GlycoWorkBench (http://code.google.com/p/glycoworkbench/)[204, 215]

This is an open source software package that has been developed to assign peaks obtained through MS glycan analysis. GlycoWorkBench links to the CarbBank, GlycomeDB and Glycosciences.de databases to search for all the experimentally reported glycans that have a given molecular weight. Based on the glycoprotein-producing organism and the glycoprotein itself, the user selects the most feasible glycan structure associated with the measured molecular weight.

Glycosciences.de (http://www.glycosciences.de/)[216, 217]

This is a web-based application for glycosciences in general. It is constituted both by glycan databases as well as tools for glycome analysis. Its main database is derived from CCSD and SugaBase. Based on these, Glycosciences.de contains glycan structure, experimental NMR, MS and crystallographic data along with taxonomy. The main software tools available within Glycosciences.de include applications that calculate the main fragments that would be generated in CID-MS spectra (*GlycoFragment*), applications that calculate the possible torsion angles of glycans (*GlyTorsion*) and software that analyses the presence of amino acids in the vicinity of carbohydrate residues (*GlyVicinity*).

With the bioinformatics tools presented in this section, many of the difficulties associated with the complex data produced by glycan analysis can be overcome. Depending on the analytical method that is selected for glycan analysis, different databases would be used in order to identify which glycans are present – and in what proportions – in a given sample. Without the availability of these tools, glycan analysis would be cumbersome and less reliable.

2.7.6. High throughput methods

An important element of PAT within the QbD perspective is *timeliness* of the measurements. Ideally, product CQA measurements should be taken on-line to obtain dynamic datasets and ensure optimal control over the process. In the case of small molecule therapeutics, on-line measurement of product CQAs has been achieved with relative success.[218-220] In contrast, on-line measurement for process inputs and CQAs of biologics is challenging, and is even more so in the case of glycosylation-associated CQAs. Although much progress has been made in reducing runtime of N-glycan analysis,[183] sample preparation prior to the analytical techniques remains a bottleneck for on-line monitoring of these CQAs. Moreover, if glycan analysis is to be used on-line or at-line to inform manufacturing operations, fast, small-scale, accurate and, preferably, automated methods would be required. Initial attempts were based on microtitre plate assays where capture, reduction, tryptic digestion and fractionation of N-linked glycopeptides were performed prior to MALDI-TOF MS glycan analysis.[221] Although this methodology required small volumes, it was still laborious and lengthy.

More successful methods for high throughput preparation of N-glycans have been reviewed recently.[222] This review discusses two methods that rely on HILIC UPLC for quantification of 2-AB labelled N-linked glycans with subsequent structural elucidation of these species performed by sequential exoglycosidase treatment and weak anion exchange (WAX) HPLC. In the context of small-scale high-throughput sample preparation, two of the methods reviewed by Doherty et

al.[222] In these methods, all steps required for sample preparation (isolation, reduction and alkylation, glycan release, labelling and purification) are performed in a 96 well plate format and have been fully automated with the use of liquid-handling robotic systems.[205, 223]

Automated, small-scale approaches for sample preparation and subsequent analysis of mAb N-glycans represent the forefront in developing on-line monitoring of glycan-associated mAb QCAs in the QbD context. These methods, combined with highly sensitive and high-throughput analytical techniques, are continuously contributing to a fuller understanding of the mechanistic relationships between bioprocess inputs and mAb glycosylation-associated quality. In addition, these high-throughput methods should eventually allow for improved bioprocess control and aid in building glycosylation-associated quality into mAbs throughout process development and manufacture.

Chapter 3

Golgi N-linked glycosylation model

3.1. Definition of the mathematical model

This chapter will present the definition of a dynamic mathematical model which intends to mechanistically describe the process of mAb Fc glycosylation. The four models that have been previously developed to represent N-linked glycosylation have been reviewed in section 2.6 of this thesis. All previous models have generated a great deal of insight into the N-linked glycosylation process. However, throughout this chapter, we hope to show that further considerations that were included into this model yield enhanced performance over previous ones, and that these additional features contribute to a fuller mechanistic description of the N-linked glycosylation of mAbs.

The three main elements of the mathematical model presented herein are:

i) The Golgi apparatus functions under the so-called cisternal maturation regime;
ii) The rate expressions for the enzymatic reactions that occur within the Golgi apparatus have been derived from the kinetic mechanisms that have been reported for each enzyme;
iii) The protein-mediated transport of NSDs into the Golgi apparatus has been included to provide a link between cellular metabolism and the Golgi glycosylation process. These elements are presented in more detail in the following sections.

In addition, the model presented in this section assumes that the first step of the N-linked glycosylation process of mAbs is completed successfully, and therefore, has been developed to only represent the glycosylation reactions which occur in the Golgi apparatus. As was mentioned previously, the close proximity of both Cγ2 domains in IgG molecules limits the space

available for the oligosaccharide structures, which in turn, limits their branching and complexity.[77] Previous reports show that the oligosaccharide bound to the Fc of mAbs are limited to complex, bi-antennary structures that are usually core-fucosylated, and have varying degrees of galactosylation and low levels of sialylation.[76, 224] For these reasons, the reaction scheme for mAb Fc glycosylation has been reduced to the one presented in Figure 12. The structure and constituting monosaccharide linkages of the oligosaccharides considered in this model are presented in Figure 7.

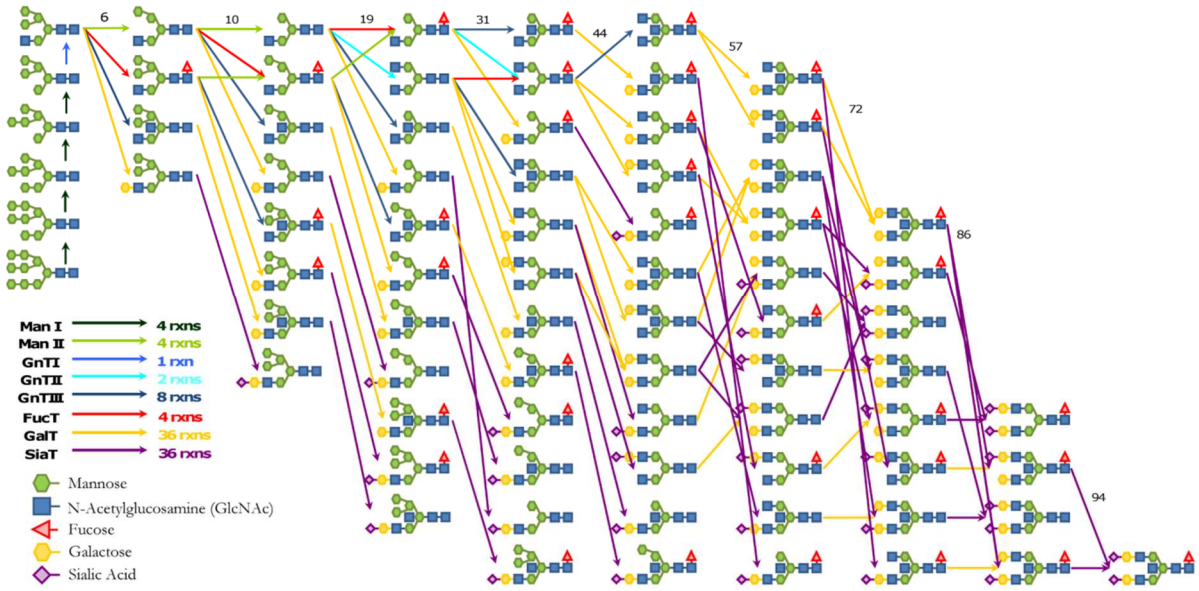


Figure 12. mAb Fc N-linked glycosylation reaction scheme[225]
The scheme begins with a Man_9 oligosaccharide which is acted upon by ManI. Once Man_5 is produced, GnTI adds GlcNAc to the α-1,3 arm of the core oligosaccharide structure. After this step, the reaction scheme diverges to 77 species generated by 95 possible reactions which are catalysed by eight different enzymes. Finally, the scheme converges to produce a fully processed biantennary oligosaccharide.

3.2. Cisternal maturation and material balances

Most of the current experimental information surrounding the Golgi apparatus favours the cisternal maturation model over the previously accepted static Golgi model.[165, 166, 226, 227] Cisternal maturation postulates that the Golgi compartments are synthesized from components of the endoplasmic reticulum and transit towards the secretory components of the cell.[228, 229] During their movement, Golgi-resident proteins are recycled from subsequent compartments to preceding ones through retrograde transport mediated by COPI coated vesicles.[226, 227] Based on cisternal maturation, the Golgi apparatus has been modelled as a single PFR, as illustrated in Figure 13.

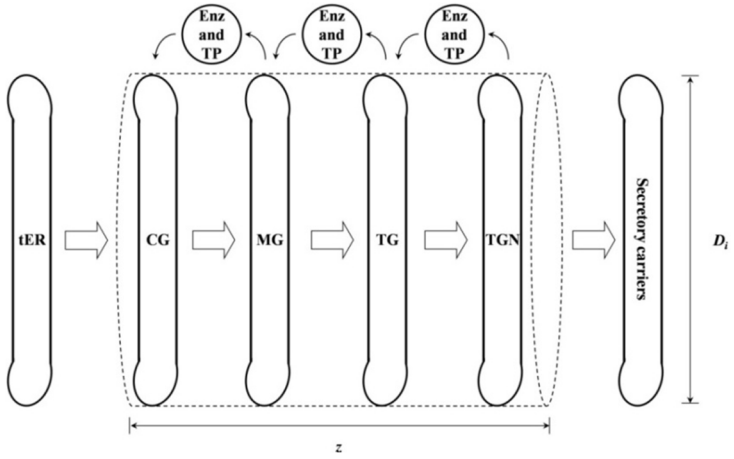


Figure 13. Cisternal maturation of the Golgi apparatus viewed as a PFR[225]
A single cisterna is produced from components of the endoplasmic reticulum (tER) and enters the Golgi space, which is defined herein as an imaginary tube (dotted line). The cisterna transits through the Golgi space, and while this occurs, its resident proteins (TPs and glycosylation enzymes) are transported to preceeding compartments as depicted by the vesicles shown at the top. Once the compartment reaches the end of the tube, its contents are transferred to carriers that are responsible for secreting the glycoprotein to the extracellular environment.

Although the compartments are discrete sections, their movement along the z axis is continuous, thus making it possible to consider the system as a single, continuous PFR. Assuming no axial dispersion within the compartment, constant flow of cargo through the complex (constant linear velocity), and constant diameter of the Golgi compartments, the time-dependent material balance for the components inside the PFR are shown in equations 1, 2 and 3.

- Oligosaccharide balances:

$$\frac{\delta OS_i}{\delta t} = F_{i,in} - \frac{4q}{\pi D_i^2}\frac{\delta OS_i}{\delta z} + \sum_{j=1}^{N.R.} \nu_i r_j \qquad \text{...Eq. 1}$$

- NSD balances:

$$\frac{\delta Pc_k}{\delta t} = F_{k,in} - \frac{4q}{\pi D_i^2}\frac{\delta Pc_k}{\delta z} + \sum_{j=1}^{N.R.} \nu_{k,j} r_j \qquad \text{...Eq. 2}$$

- By-product balances:

$$\frac{\delta B_k}{\delta t} = -F_{k,in} - \frac{4q}{\pi D_i^2}\frac{\delta B_k}{\delta z} + \sum_{j=1}^{N.R.} \nu_{k,j} r_j \qquad \text{...Eq. 3}$$

An integral element of the cisternal maturation model is that the resident proteins of the Golgi apparatus (glycosylation enzymes and transport proteins) are recycled from subsequent

compartments to preceding ones.[226, 227] In the present model, this has been incorporated by defining enzyme concentration profiles along the length, z, as three parameter normal functions (Equations 4 and 5):

- Enzymes: $$E_j(z) = E_{j,max} e^{-\frac{1}{2}\left(\frac{z - z_{j,max}}{\omega_j}\right)^2} \quad \text{...Eq. 4}$$
- Transport proteins: $$TP_k(z) = TP_{k,max} e^{-\frac{1}{2}\left(\frac{z - z_{k,max}}{\omega_k}\right)^2} \quad \text{...Eq. 5}$$

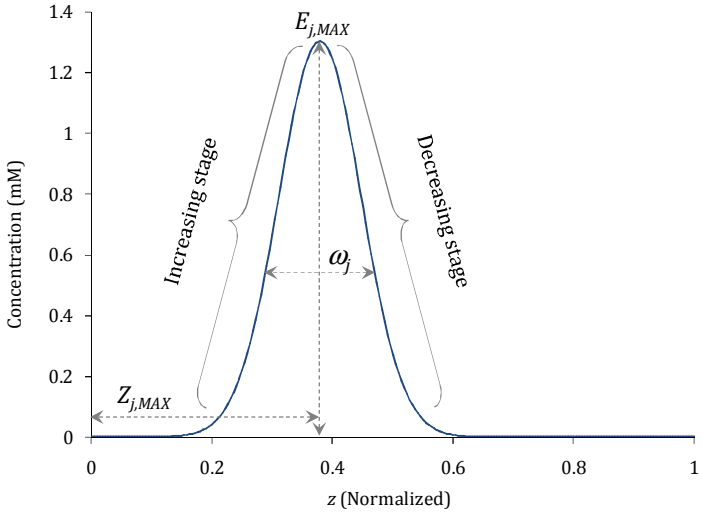


Figure 14. Definition of Golgi resident protein concentration profiles
The Golgi resident protein concentration profiles are defined by three parameters ($E_{j,MAX}$, $z_{j,MAX}$ and ω_j). $E_{j,MAX}$ represents the peak concentration of the resident protein, $z_{j,MAX}$ corresponds to the location of the peak concentration along the Golgi space and ω_j represents the breadth of the concentration.

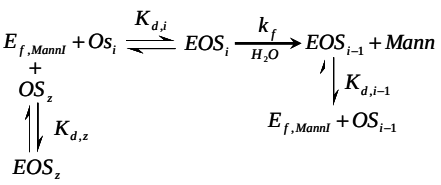
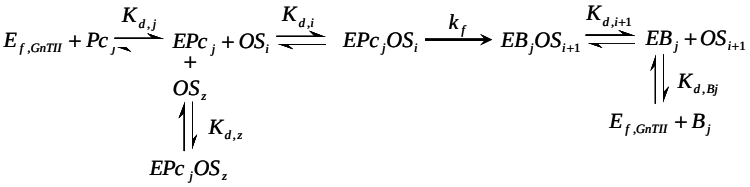
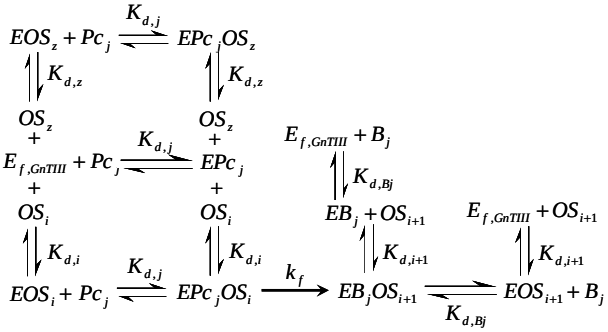
These functions represent the inherent behaviour of enzyme concentration when recycling is considered: there would be a stage during which the enzyme accumulates while it is being fed into a given compartment from the subsequent one. The enzyme concentration would then reach a transitory peak concentration after which the enzyme would begin to be transferred to preceding compartments, thus generating a decline stage. The stages of Golgi resident protein accumulation and depletion, as well as a description of each of the three parameters that define Eqs. 4 and 5, are presented in Figure 14.

3.3. Enzyme kinetics

The previous models for N-linked glycosylation considered simplified enzyme kinetics, where only competitive inhibition was considered. Moreover, all previous models considered the same

kinetic mechanism for all of the enzyme-catalyzed reactions. The kinetic expressions in the present model were expanded by considering the structural and kinetic mechanisms by which these enzymes act. Table 7 shows that three kinetic mechanisms prevail among the enzymes involved in mAb N-linked glycosylation. Each reaction mechanism defines how the substrate(s) bind to the enzyme, and which step of the process is the catalytic one. Additionally, but not independently to the catalytic mechanism, different types of inhibitory effects may occur within these systems. Given that many different oligosaccharides can be modified by the same enzyme, it is reasonable to assume that competitive inhibition occurs. Moreover, product inhibition has been reported for many of these reactions.[230-232] The reaction rates are presented as equations 6 through 8. The full derivation is included in Appendix I.

Table 7. Kinetics of N-linked glycosylation enzymes

Kinetic Mechanism	Reaction Mechanism	Enzymes
Michaelis Menten with competitive and product inhibition		ManI[233, 234] ManII[235]
Sequential-order Bi-Bi with competitive and product inhibition		GnTI[236] GnTII[237] GalT[231]
Random-order Bi-Bi with competitive and product inhibition		FucT[230] GnTIII[232] SiaT[238]

Michaelis-Menten kinetics:

$$r_j = \frac{k_{f,j}\left[E_j\right]\left[OS_i\right]}{K_{d,i}\left(1+\frac{\left[OS_i\right]}{K_{d,i}}+\sum_{z=1}^{NC}\frac{\left[OS_z\right]}{K_{d,z}}+\frac{\left[OS_{i-1}\right]}{K_{d,i-1}}\right)} \quad \text{...Eq. 6}$$

Sequential-order Bi-Bi kinetics

$$r_j = \frac{k_{f,j}\left[E_j\right][Pc_k][OS_i]}{K_{d,k}K_{d,i}\left(1+\frac{[Pc_k]}{K_{d,k}}+\frac{[Pc_k]}{K_{d,k}}\frac{[OS_i]}{K_{d,i}}+\frac{[Pc_k]}{K_{d,k}}\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}}+\frac{[B_k]}{K_{d,Bk}}\frac{[OS_{i+1}]}{K_{d,i+1}}+\frac{[B_k]}{K_{d,Bk}}\right)} \quad \text{...Eq. 7}$$

Random-order Bi-Bi kinetics

$$r_j = \frac{k_{f,j}\left[E_j\right][Pc_k][OS_i]}{K'_{d,k}K_{d,i}\left(1+\frac{[Pc_k]}{K_{d,k}}+\frac{[OS_i]}{K_{d,i}}+\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}}+\frac{[Pc_k]}{K'_{d,k}}\frac{[OS_i]}{K_{d,i}}+\frac{[Pc_k]}{K''_{d,k}}\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}}+\frac{[B_k]}{K_{d,Bk}}+\frac{[OS_{i+1}]}{K_{d,i+1}}+\frac{[B_k]}{K_{d,Bk}}\frac{[OS_{i+1}]}{K_{d,i+1}}\right)}$$

...Eq. 8

3.4. Nucleotide sugar donor transport

NSDs (UDP-GlcNAc, GDP-Fuc, UDP-Gal and CMP-Neu5Ac) are the co-substrates for the sugar addition reactions of N-linked glycosylation and their availability inside the Golgi cisternae has been reported to affect the N-linked glycosylation process.[78, 107, 108, 162] NSDs are synthesized in the cytosol of cells (with the exception of CMP-Neu5Ac, which is produced in the nucleus) and their transfer into the Golgi apparatus is mediated by transport proteins (TPs). All four TPs considered in this paper, act by an anti-port mechanism whereby the NSDs are transported into the Golgi cisternae while their dephosphorylated nucleotide component (which accumulates in the Golgi lumen through the glycosylation reactions) is transported out of the cisternae.[239] For example, the TP for UDP-GlcNAc removes a single uridine monophopsphate (UMP) molecule for every UDP-GlcNAc molecule it introduces into Golgi.[240] The same occurs for all other transport proteins considered in this work.[241-243] These transport mechanisms have also been characterized quantitatively. In all cases, transport has been found to follow saturation kinetics, which depends on the concentration of the NSD on the cytosolic side of Golgi and on the respective dephosphorylated nucleotide concentration on the luminal side.[241-243]

To generate a first link between cellular metabolism and N-linked glycosylation, the transport of NSDs has been included into our model. Using the information presented above, and considering that the interaction between transport substrates and the TP is analogous to enzyme-substrate interactions, the expression describing transport of NSDs into the Golgi apparatus is defined as equation 9. The derivation of this equation is presented in Appendix II.

$$F_{k,in} = k_{T,k}\left[TP_k\right]\left(\frac{A_G}{V_G}\right)\left(\frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}+\left[Pc_k^{Cyt}\right]}\right)\left(\frac{\left[BP_k^{G.L.}\right]}{K_{BP}^{G.L.}+\left[BP_k^{G.L.}\right]}\right) \quad \text{...Eq. 9}$$

3.5. Golgi glycosylation parameter values

The overall cell volume considered in this work is 2,500 μm^3.[244] Combining this with structural information of the Golgi apparatus reported by Ladinsky et al.,[245] the volume of the Golgi apparatus was calculated as 25 μm^3. Assuming only four compartments and using an aspect ratio (diameter:thickness) of 60 to 1 (approximated from Ladinsky et al.[245]), the obtained internal diameter of the tube presented in Figure 13 is 7.82 μm. The overall length of the tube (length of the Golgi stack) is calculated as 0.52 μm. With this information, the surface area of a single compartment was calculated as 990 μm^2. The specific protein productivity was taken as 387 pmol (10^6 cell)$^{-1}$ day^{-1}.[246] This value falls within the range reported by Bibila and Robinson.[247] When considering that each mAb molecule contains two N-linked glycosylation sites, the glycan productivity for this system is 773 pmol (10^6 cell)$^{-1}$ day^{-1}, which also lies within the range of 100 to 1,500 pmol (10^6 cell)$^{-1}$ day^{-1} used in previous models.[31-33]

Umaña and Bailey estimated the residence time of proteins within the Golgi apparatus as 22 min.,[31] which seems reasonable considering that Hirschberg and Lippincott-Schwartz[248] reported an ER/Golgi residence time of 40 min with a protein flux equivalent to 574 pmol (10^6cell)$^{-1}$ day^{-1}. The volumetric flowrate (q) through the Golgi space was defined as 1.12 μm^3 min^{-1} and the concentration of glycoprotein that is transferred from the endoplasmic reticulum to the Golgi apparatus (*i.e.* the concentration of glycoprotein at z = 0) was defined as 600 μM of glycoprotein. The ensuing concentration of glycan entering the Golgi apparatus is 1,200 μM. Enzyme kinetic parameters and NSD transport parameters are presented in Table 8 and Table 9, respectively. Details for the calculation of these parameter values are presented in Appendix III.

3.5.1. Parameter estimation for enzyme concentration profiles

The localization of the enzymes along the length of Golgi (z) is based on experimental information.[62] This group determined the distribution of ManII, GnTI, GalT and SiaT by immuno-labeling cryo-sections of HeLa cells, but left the distribution of GnTII and FucT unknown. This methodology generated a series of "snap-shots" of enzyme distribution related to the static compartment model of Golgi. The mathematical model presented herein attempts to dynamically describe cisternal maturation of Golgi, for this reason, the data presented by

Rabouille et al.[62] were further detailed by an optimization-based approach. The parameters that define enzyme distribution were estimated by finding the minimum amount of total enzyme (the integral of enzyme concentration over the length of Golgi) necessary to achieve full oligosaccharide processing, including 50% sialylation.

Table 8. Enzyme turnover and dissociation constants

Enzyme turnover rates

Enzyme	V_{max} (μmol mg^{-1} min^{-1})	Organisms	MW (g/mol)	k_f (min^{-1})
ManI	3.677	Human, purified.[249]	230,000[250]	888
ManII	6.750	Rat liver, purified.[234]	285,000[234]	1,924
GnTI	19.80	Rabbit liver, purified.[251]	51,600[252]	1,022
GnTII	27.50	Rat liver, purified.[237]	200,000[237]	1,406
GnTIII	5.52	Rat liver, purified.[253]	114,000[253]	629
FucT	4.37	Porcine brain, purified.[254]	66,530[254]	291
GalT	19.80	Rat liver, purified.[255]	44,040[256]	872
SiaT	8.20	Rat, purified.[257]	59,930[258]	491

Enzyme Dissociation constants

Enzyme	Substrate	$K_{d,i}$ (μM)	$K_{d,k}$ (μM)	Notes and References
ManI	Man_9	60.5	----	Estimated.[233]
	Man_8	110.0	----	Human, purified.[249]
	Man_7	30.8	----	Estimated.[233]
	Man_6	74.1	----	Estimated.[233]
ManII	Man_5	200.0	----	Purified human adenocarcinoma.[259]
	Man_4	100.0	----	Estimated.[32]
FucT	$CoreGlcNAc_2$	25.0	46.0	Porcine brain, purified.[254]
	α-1,6 arm occupied	1.2		Estimated.[32]
GnTI	Man_5	260.0	170.0	CHO, purified.[260]
GnTII	$CoreGlcNAc_1$	190.0	960.0	Rat liver, purified.[237]
GnTIII	$CoreGlcNAc_2$	21.0	420.0	Rat GnTIII expressed in insect cells.[232]
GalT	$CoreGlcNAc_2$ (α-1,3 arm)	430.0	65.0	Rat liver, purified.[261]
	$CoreGlcNAc_2$ (α-1,6 arm)	130.0		
	$CoreGlcNAc_2Gal_1$ (α-1,6 arm)	6,280.0		
SiaT	Asialo glycoprotein	330.0	50.0	Rat expressed in COS-1 cells.[262]

Table 9. Transport protein parameters

Transport protein	V_{max} (pmol mg_{prot}^{-1} min^{-1})	$k_{T,k}$ (min^{-1})	$K_{Pc,k}^{Cyt}$(μM)
TP_1 (UDP-GlcNAc/UMP)	915.0	1,084	7.13[240]
TP_2 (GDP-Fuc/GMP)	1.4	230	7.5[243]
TP_3 (UDP-Gal/UMP)	22.0	689	2.4[242]
TP_4 (CMP-Neu5Ac/CMP)	335.0	397	1.3[241]

Table 10. Intracellular, cytosolic and Golgi NSD concentrations

NSD	Total intracellular concentration (μM)	Concentration in Golgi lumen (μM)	Concentration in cytosol (μM)
UDP-GlcNAc	36.7[263]	1,057	26.4
GDP-Fuc	20.0[264]	576	14.4
UDP-Gal	9.2[263]	265	6.6
CMP-SA	33.0[110]	950	23.7

The parameters that define enzyme distribution were estimated by seeking the minimum amount of total enzyme (the integral of enzyme concentration over the length of Golgi) necessary to achieve terminal oligosaccharide processing, including 50% sialylation. The parameter values were obtained through a multiple-shooting control vector parameterization algorithm, which i) chooses the duration of each control interval and the values of the control and differential variables at the start of each control interval, ii) solves the dynamic model over the time horizon to determine the time variation of all variables in the system and with them, obtains the values for the objective function, the constraints and the differences between computed values and the ones chosen by the optimiser at the start of each interval. The optimiser finally revises the first step and repeats the procedure until the objective function is minimised, the constraints are satisfied and the differential variables are continuous at the beginning of each control interval.[265]

The calculated enzyme concentration profiles along with the estimated parameters that define them are presented in Figure 15A. The enzymes appear in a sequential fashion, but overlap considerably, which is consistent with previous findings.[62, 227] The order in which the enzymes appear is related to the reaction sequence itself. For example, the early mannose-trimming reactions occur at early stages of the reaction scheme, thus ManI appears first. Co-localization of enzymes was found to have an impact on the reaction kinetics. Specifically, in simulations where ManI did not overlap with GnTI, higher overall enzyme concentration was necessary to achieve full removal of the mannose residues from the precursor oligosaccharide (GnTI pushes the mannose trimming reactions forward and, to an extent avoids product inhibition of ManI activity).

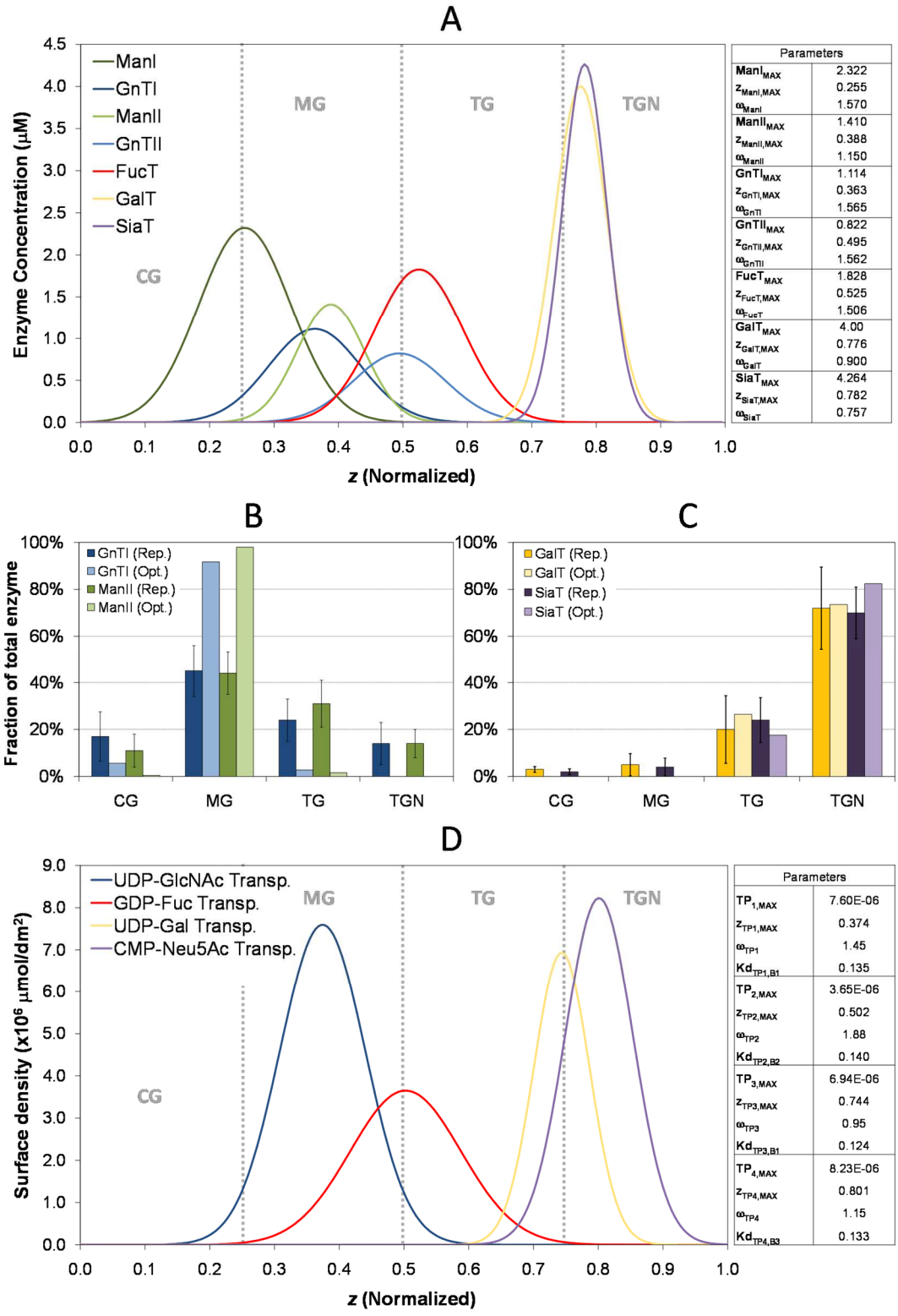


Figure 15. Resident protein profiles along the Golgi Apparatus[225]
A: Calculated glycosylation enzyme concentration profiles along the normalised length of Golgi. B and C: Comparison of the optimisation results for enzyme distribution (Opt.) with those reported by Rabouille et al.[62] (Rep.) D: Calculated TP concentration profiles. The vertical lines in (A) and (D) present the regions of the Golgi space. The tables alongside A and D present the parameters that were estimated and used to calculate the presented profiles.

Similar effects were observed relating FucT with GnTII and GalT with SiaT. The total concentration of enzymes (the integral of the presented profiles over the length of Golgi, *z*) ranged between 3.2 μM and 9.1 μM, which is consistent with the glycosylation enzyme concentration range of 1 μM to 35 μM estimated through multiplying the membrane-bound protein content of 38,000 proteins μm^{-2} reported by Quinn et al.[266] by the molar fraction of Golgi glycosylation enzymes reported by Gilchrist et al.[267]

A comparison between the estimated enzyme parameters and those reported by Rabouille et al.[62] is presented in Figure 15B and C. It is clear from Figure 15B that the experimental profiles are broader, whereas the simulation profiles are narrower and mainly localized in the medial Golgi. This deviation is likely due to the underlying assumptions of the model (symmetric enzyme concentration profiles and minimum enzyme required for full oligosaccharide processing), but also because of the "snap-shot" representation of the Rabouille et al.[62] data. In Figure 15C, good agreement can be seen between the experimental and estimated profiles. The enzyme profiles presented in Figure 15 also satisfy previously reported properties: i) GnTI co-localizes with ManII in the medial Golgi, ii) GnTI, ManII, GalT and SiaT coexist in the trans Golgi region, and iii) GalT and SiaT co-localize in the trans Golgi (TG) and the trans Golgi network (TGN) regions.[62, 120, 227] Hence, it is possible to consider the estimated enzyme concentration profiles as representative of the system.

3.5.2. Parameter estimation for transport protein concentration profiles

The transport protein concentration profiles were also assumed to follow normal functions so that the effect of Golgi resident protein recycling would be represented. Low availability of NSDs for the Golgi glycosylation reactions has been linked to cell metabolism and not to deficiency of their transport into the Golgi[107, 118, 268] with the with the exception of Rijcken et al.,[111] who observed an interesting control mechanism by which UDP-GlcNAc was observed to control transport of CMP-Neu5Ac into Golgi. This implies that when NSDs are sufficiently available in the cytosol, the transport of NSDs into the Golgi is very close to the rate of accumulation of respective by-products which, in turn, is the rate of consumption for each NSD. By assuming that the rate of by-product dephosphorylation is much faster than their accumulation, the objective function presented as equation 10 was defined and minimized through a multiple-shooting control vector parameterization algorithm[265] to estimate the missing TP parameters.

$$OF_k = \int_0^1 \left(F_{k,in} - V \sum_{k=1}^{N.R.} r_k \right)^2 dz \qquad \text{...Eq. 10}$$

A comparison between the rate of consumption of NSDs and the rate of their transport is provided in Appendix IV. Both sets of data overlap showing minimization of the objective function presented in equation 10. Results for the parameter estimation are presented in Figure 15D. It has been previously reported that the UDP-GlcNac TP closely co-localizes with ManII,[269] the UDP-Gal transporter with GalT[270] and the CMP-Neu5Ac transporter with SiaT.[271] A comparison of Figure 15A with Figure 15D shows that these observations are satisfied. In addition, the estimated cumulative concentration of transport proteins along the Golgi apparatus ranges between 900 μmol μm^{-2} and 1,700 μmol μm^{-2} which is well within the total membrane-bound protein density of 38,000 reported for the Golgi apparatus of baby hamster kidney cells.[266]

3.6. Parameter estimation and model simulation

Most of the enzyme kinetic parameters were obtained through experiments performed using free oligosaccharides as substrates (see Table 8). However, Tabas and Kornfeld[233] also performed kinetic experiments using an immunoglobulin M-derived glycopeptide. Their results showed that at low glycopeptide concentrations, the kinetic parameters are similar to those of the free oligosaccharide, but at high glycopeptide concentration, the rates diverge, implying a change in affinity of the enzyme for the substrate. This evident change in enzyme-glycoprotein affinity along with the spatial constraints associated to the mAb Fc oligosaccharides served as the basis to consider the dissociation constants as the adjustment criterions when trying to match experimentally-obtained mAb Fc oligosaccharide profiles.

Two recent publications reporting the oligosaccharide profiles of commercial mAbs[45, 191] were used as target data so that unknown kinetic parameters of the model could be estimated. The data were introduced into gPROMS version 3.4.0, which is based on a SRQPD sequential quadratic programming code.[265] Parameter estimation in gPROMS v.3.4.0 is based on the maximum likelihood formulation that attempts to determine values for the uncertain physical and variance model parameters which maximize the probability that the model will predict the values obtained from experimental measurements.

Parameter estimation was only performed for the dissociation constants of GnTII, FucT and GalT because very little or no high mannose, hybrid, bisecting GlcNAc or sialylation were

reported by Stadlmann et al.[45] and Nakano et al.[191] The $K_{d,i}$ for GnTIII and SiaT were defined as $1x10^{12}$ to represent that the oligosaccharide structures of the mAb Fc are inaccessible for GnTIII and SiaT. Also, the dissociation constant for the addition of galactose to the α-1,6 arm of the biantennary oligosaccharide when the α-1,3 arm is already occupied (KdiGalTa2B) was defined as 50-fold that of KdiGalTa1A based on the previously reported branch specificity of GalT.[261] The estimated parameters are presented in Table 11 and the terminal oligosaccharide profiles calculated with these parameters are presented in Figure 18. Additional dynamic results showing the oligosaccharide profile distributions along the length of the Golgi apparatus (z) are presented in Figure 16.

Table 11. Estimated dissociation constants with present model[225]

Dissociation Constant (μM)	Herceptin	Rituxan	Remicade	Erbitux
$K_{d,i}^{GnTII}$	1,572	1,180	462	366
$K_{d,i}^{FucTA}$	711	593	176	145
$K_{d,i}^{FucTB}$	467	465	169	72
$K_{d,i}^{GalTa1A}$	143	166	24	5.4
$K_{d,i}^{GalTa1B}$	23	105	29	24
$K_{d,i}^{GalTa2A}$	59	87	149	510
Deviation from experimental data (%)	1.9	10.4	0.5	2.2

In three cases, the experimental data is reproduced to within 2.5% total deviation. The exception is Rituxan, where the net deviation is 10%. The main source for the larger deviation is the bi-galactosylated, core fucosylated oligosaccharide 57 shown in Figure 18. Three of the six dissociation constants that were estimated correspond to the specificity of GalT for different arms of the oligosaccharide. We believe that the higher deviation associated with OS_{57} in the Rituxan data is due to the constraints generated by the relationship between the three GalT dissociation constants. If the values for any of the $K_{d,i}^{GalT}$ values change, they affect the synthesis of more processed oligosaccharides and yield higher deviations associated to species OS_{31} and $OS_{44/45}$. In reality, each oligosaccharide species should have its own affinity for the different enzymes due to enzyme-substrate specificity. We believe that, for the Rituxan data, a deviation of 10% is acceptable considering that the number of GalT-associated $K_{d,i}$ parameters has been reduced from a maximum of 36 given by our reaction scheme (see Figure 12) to only three.

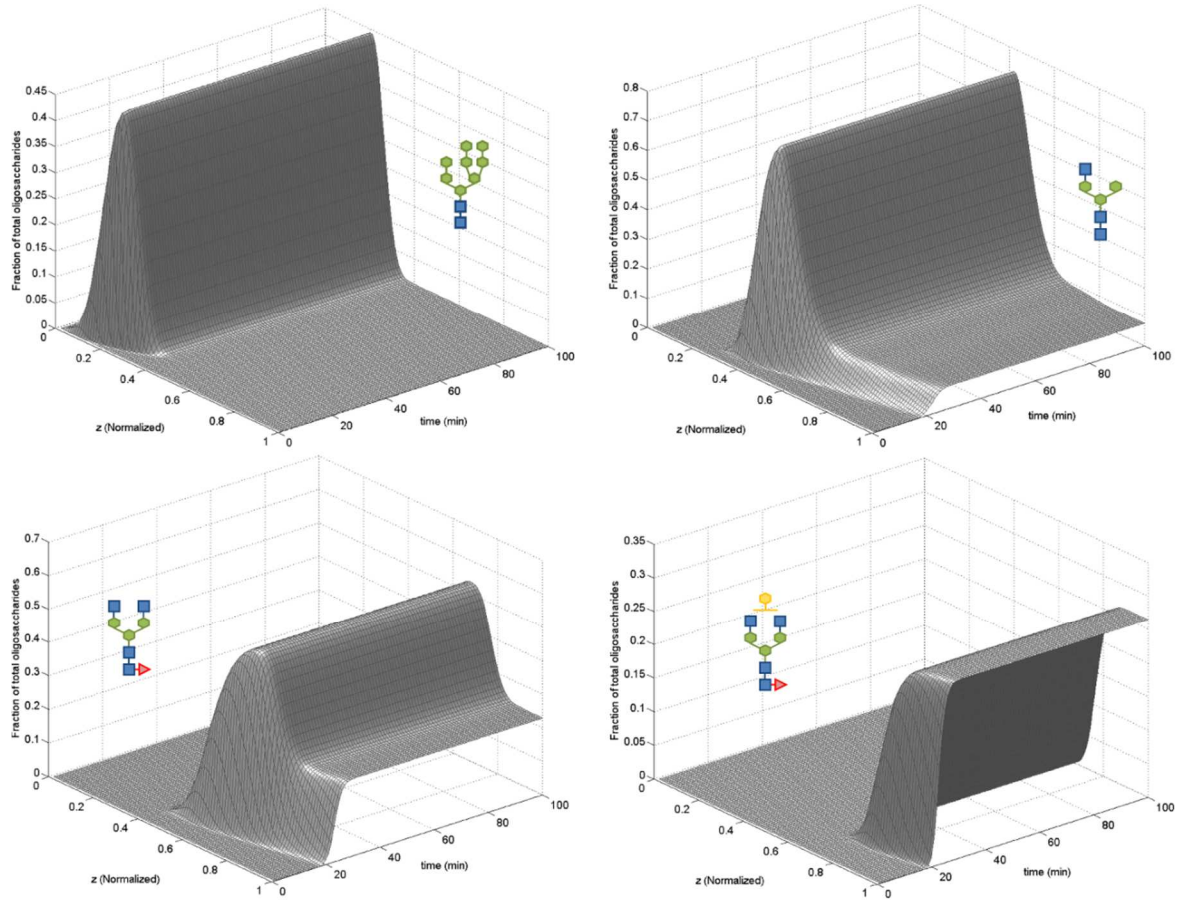


Figure 16. Calculated spatial and temporal oligosaccharide profiles within Golgi. (A) shows the molar fraction profile of an early oligosaccharide as a function of time and distance along the Golgi. (B) corresponds to the profile of an intermediate oligosaccharide and (C) and (D) present the profiles of terminal oligosaccahrides.

3.7. Comparison with previous models

3.7.1. Effect of kinetic mechanisms

As mentioned in previous sections, the present model builds on the knowledge generated by previous models of the Golgi N-linked glycosylation process. In the three previous models,[31-33] all enzymes involved in the sugar addition reactions of the process were assumed to act with a random- order kinetic mechanism. However, literature suggests that GnTI, GnTII and GalT function with a sequential-order kinetic mechanism.[231, 237] When comparing the rate expressions derived for both mechanisms, we find that there are additional terms in the denominator for the random-order kinetic expression (compare the denominator of equations 7 and 8). These terms correspond to additional enzyme-substrate complexes (EOS_i, EOS_z and EOS_{i+1}) that occur in random-order kinetics, but are absent in sequential-order kinetics.

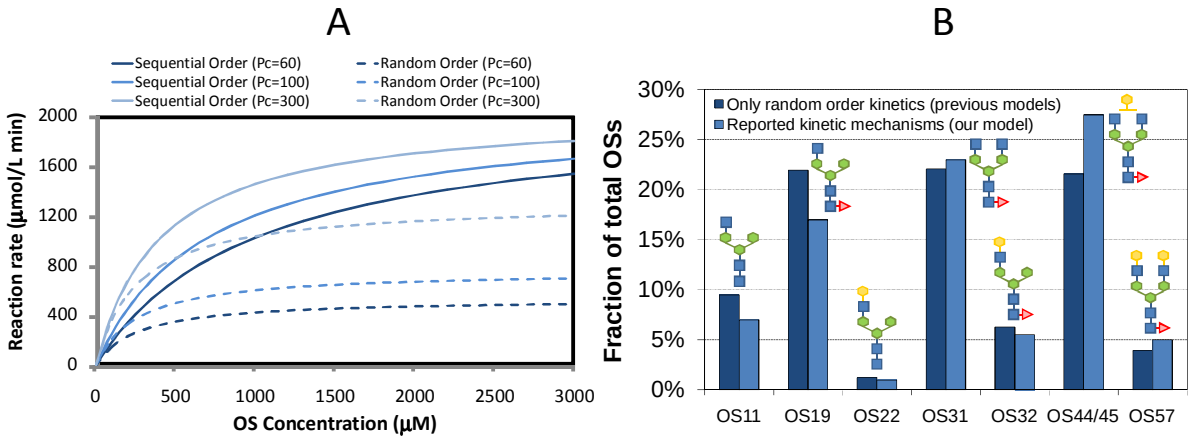


Figure 17. Comparison of present model with previous ones[225]
(A) Compares the profiles of sequential-order kinetics (solid lines) with those of random-order kinetics (dashed lines) as a function of oligosaccharide concentration at three different NSD concentrations (60, 100, and 300mM). (B) Compares oligosaccharide profiles calculated with only random-order kinetics (■) with those calculated with the correct kinetic mechanism for each enzyme (■).

Figure 17 compares the kinetic mechanisms assumed by previous models (only random-order) with the mechanisms included in this section (combination of sequential-order and random-order kinetics, as shown in T). Figure 17A compares the impact the different mechanisms have on the reaction rate values and Figure 17B shows how considering the appropriate kinetic mechanisms influence the oligosaccharide profiles calculated by the model. From the data, it is clear that the different kinetic mechanisms indeed affect the calculated oligosaccharide profiles. In order to assess the impact of the different kinetic mechanisms on parameter estimation, the dissociation constants for the mAb-bound oligosaccharides were estimated as described in the previous section but with the model containing only random-order kinetics. Successful reproduction of the commercial mAb oligosaccharide profiles was achieved. However, we believe that including the actual kinetic mechanisms for each enzyme generates no further complexity in the model, and therefore, should be used to represent the Golgi glycosylation process with higher realism.

3.7.2. Model performance and parameter estimation strategy

The only previous model that attempted to reproduce experimentally-observed oligosaccharide profiles was that of Krambeck and Betenbaugh.[32] They attempted to match the oligosaccharide profiles for human thrombopoietin by adjusting the total glycosylation enzyme concentration while maintaining the corresponding spatial distributions constant. In contrast, Hossler et al.[33] did not attempt to match experimental data given the scope of their work. Instead, they estimated the glycosylation enzyme concentrations that would maximize terminally-processed

oligosaccharides. Under this methodology, Hossler et al. also assumed enzyme distribution along Golgi as constant.

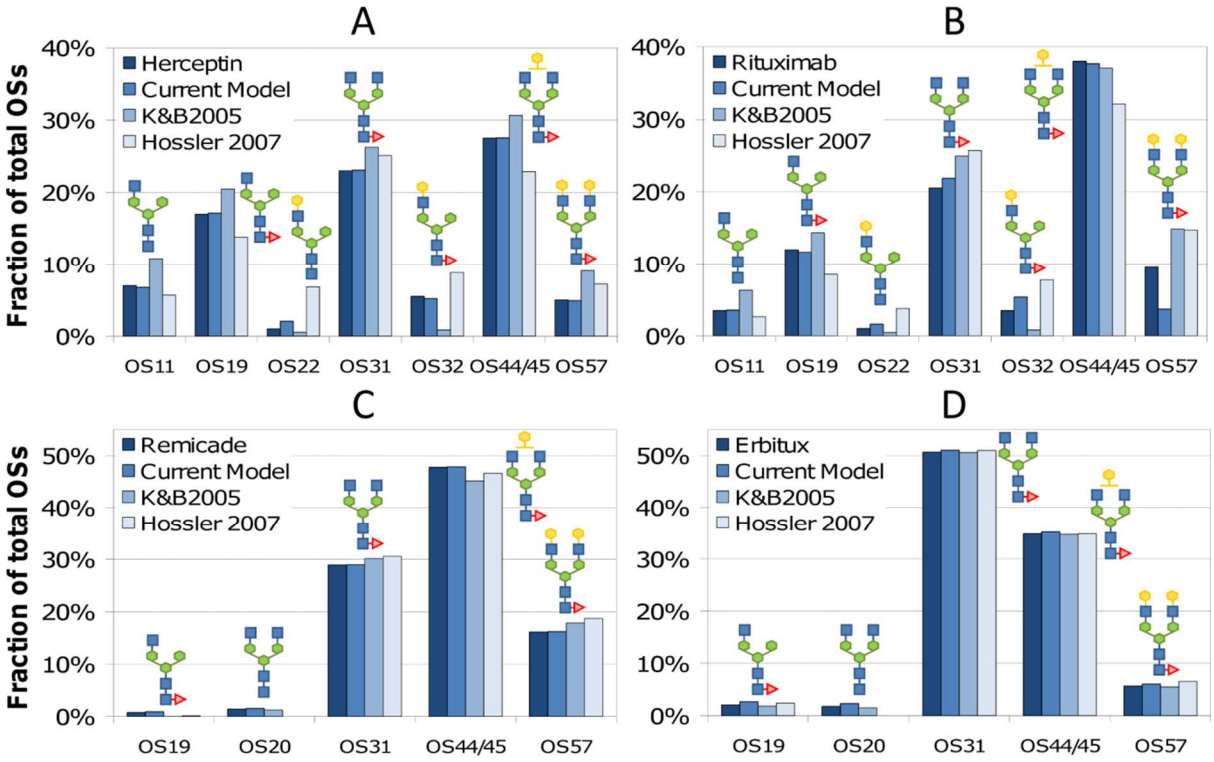


Figure 18. Calculated vs. experimental oligosaccharide profiles[225]
Reported data from Nakano et al.[191] for Herceptin (A) and Rituxan (B) as well as data from Stadlmann et al.[45] for Remicade (C) and Erbitux (D) is shown with the dark bars (■). The subsequent bars correspond to results calculated with the present model (■), the Krambeck & Betenbaugh model (■) and the Hossler et al. model (□).

In order to compare the performance of our model with the previous ones, we defined each of them, including their underlying assumptions and parameters in gPROMS 3.4.0. The sole modification was that the reaction scheme was reduced to the one presented in Figure 12, which is specific for mAb Fc oligosaccharides. When simulating both models with their original parameters, they both calculated unrealistically high proportions of the bi-galactosylated and core fucosylated oligosaccharide 57 (76% for Krambeck and Betenbaugh and 41% for Hossler et al.). These values are five and 2.5 times the value of 16.1% reported for Remicade presented in Figure 18. The large discrepancies are likely due to the combination of parameters and assumptions the previous authors considered and make the results uncomparable with those generated by our model.

In order to produce a more reasonable comparison between the different models, we used the strategy proposed by Krambeck and Betenbaugh of estimating the total enzyme concentration to match experimental oligosaccharide profiles. So that the results are comparable, the same number of parameters was estimated in all cases: six dissociation

constants for the current model (as described in previous sections) and the total concentration of six glycosylation enzymes for the previous models. This approach was implemented in the parameter estimation function of gPROMS version 3.4.0, which has been described in previously in this thesis. The estimated parameters, along with the deviations from the experimental data are presented in Table 12.

Table 12. Estimated parameters with previous models[225]

	K&B Model (μM)				Hossler et al. (μM)			
Enzyme	**Herceptin**	**Rituxan**	**Remicade**	**Erbitux**	**Herceptin**	**Rituxan**	**Remicade**	**Erbitux**
ManI	1.21	1.5	1.67	2.06	20.85	19.91	27.34	27.95
ManII	9.29	6.16	4.07	1.77	6.20	21.88	6.97	9.22
GnTI	2.54	1.58	9.27	3.98	7.72	9.58	2.35	2.40
GnTII	0.18	0.23	1.58	0.36	0.30	0.49	0.64	0.42
FucT	0.82	0.85	0.82	0.81	1.10	1.59	18.89	17.15
GalT	0.014	0.019	0.021	0.012	8.18	10.62	10.46	6.34
Deviation from experimental data (%)	1.9	10.4	0.5	2.2	1.9	10.4	0.5	2.2

It is clear from the data presented in Table 12 that the values obtained for both of the previous models are dissimilar. This is very likely due to differences in the kinetic parameters assumed by each model. Namely, Hossler and collaborators derived the turnover rate for each enzyme by multiplying the reported maximum reaction rate by the molecular weight of the monomeric form of each enzyme, while Krambeck and Betenbaugh assumed the molecular weight to be that of the entire enzymatic complex. Because of this, the estimated enzyme concentrations for the Hossler et al. model are much higher than those estimated for the Krambeck & Betenbaugh model. It is worth noting that due to this assumption by Hossler et al., the total amount of resident protein (both with the original enzyme concentrations and the estimated ones), present in the Golgi is much higher than the maximum membrane bound protein density reported by Quinn et al.,[266] thus deeming these concentration profiles unfeasible. In contrast, in the current model, the total amount of resident protein in Golgi is approximately 60% of the value reported by Quinn et al.[266] This, along with the implication of resident protein recycling from the Golgi maturation model, suggests that our estimated resident protein profiles are more realistic than simple constant concentration profiles for resident proteins.

A comparison of the terminal oligosaccharide profiles calculated with the estimated parameters has been included in Figure 18. All but one case (Erbitux profiles calculated with the Krambeck and Betenbaugh model) show much larger deviations from the experimental data than those obtained with our proposed model and parameter estimation strategy. On average, our strategy of estimating dissociation constants yields deviations which are 4.5 times lower

than those obtained with the Krambeck and Betenbaugh model and 5.5 times lower than those obtained with the Hossler et al. model. The cause of these discrepancies is a combination of model structure and the estimation strategy itself (estimating Kd,is versus estimating enzyme concentrations). In order to isolate these effects, a second series of parameter estimations were performed on the previous models. In this case, the enzyme concentrations were set at the original values assumed by each model and the dissociation constants were estimated. When estimated for the Krambeck & Betenbaugh model, the $K_{d,i}$s produced terminal oligosaccharide profiles that deviate 5 times more than the current model. The results for the Hossler et al. $K_{d,i}$ estimation deviated 9 times more than those presented in Table 11. In both cases, the deviations are higher than those obtained by estimating total enzyme concentration.

These results suggest that the reason why the current model out-performs the previous ones is due to structure and not the estimation strategy. More specifically, the results suggest that the definition of the glycosylation enzyme distributions is why the current model out-performs the previous ones in reproducing experimental oligosaccharide profiles. In order to confirm this, a final set of parameter estimations were performed to obtain the distribution of glycosylation enzymes that could match experimental data. In this case, the estimations yielded results that are comparable with those obtained with the model presented herein (data not shown). However, it is worth noting that, in these cases, 24 parameters were estimated: the concentration of the six glycosylation enzymes in each of the four Golgi compartments. It is possible that good reproduction of the experimental oligosaccharide profiles was achieved due to over-fitting.

This section has compared the performance of the current model with previous ones. The results suggest that the present model along with the methods used to estimate its unknown parameters out-performs the previous ones in reproducing experimental mAb Fc oligosaccharide profiles. Furthermore, if mathematical models for mAb Fc N-linked glycosylation were to be used for bioprocess design, control and optimization following the Quality by Design paradigm, it would be necessary for them to represent all the relevant known phenomena surrounding the process. In this section, we have shown that our additional considerations (appropriate kinetic mechanisms and Golgi protein recycling) have contributed to the fidelity of the model while avoiding significant additional complexity.

3.8. Case study: glycotransferase silencing

Additional simulations were performed in order to further assess the performance of the model against independent sets of data. These simulations were based on the study by Imai-Nishiya

and collaborators,[86] in which the gene encoding for the FucT enzyme was silenced using small interfering RNA (siRNA). As a first step, the dissociation constants for GnTII, FucT and GalT were estimated (as described in preceding sections) to match the parental oligosaccharide profile.

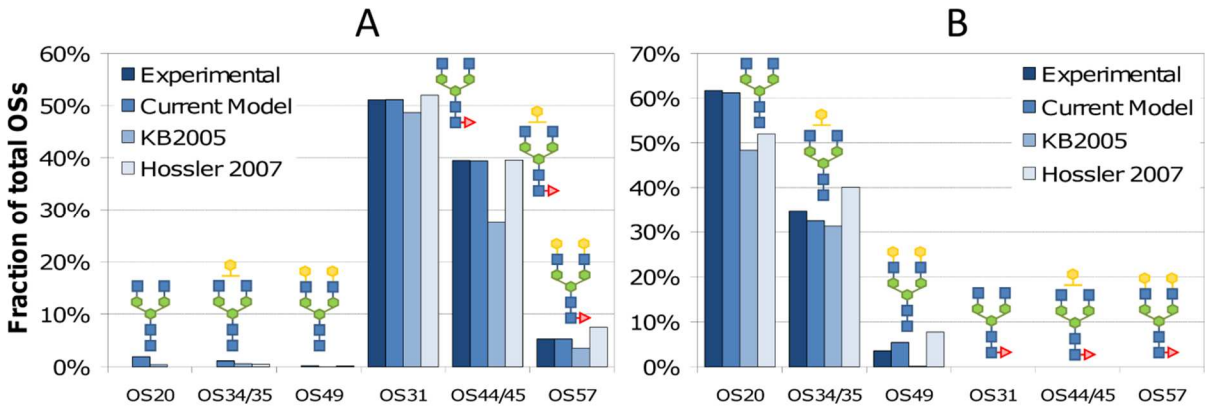


Figure 19. FucT silencing[225]
(A) Shows the experimental[86] and calculated glycan profiles prior to FucT silencing. (B) shows the glycan profiles after FucT silencing. The experimentally observed glycan profiles (■) are compared with those calculated with the current model (■), the K&B2005 model (■) and the Hossler 2007 model (□).

A comparison between the calculated and parental oligosaccharide profiles is presented as Figure 19A, which shows that the profiles obtained through simulation closely match the reported ones yielding a net deviation of 3.4%. Next, the concentration of FucT was set to zero within the model and a simulation was performed. A comparison of the simulated and reported oligosaccharide profiles after FucT silencing is presented in Figure 19B. In this case a slightly higher deviation of 4.7% was obtained. The same procedure was performed with the Krambeck & Betenbaugh and the Hossler et al. model. The Hossler et al. model had a net deviation of 4.0%, which is comparable with the one obtained with the current model. In contrast, the Krambeck & Betenbaugh model had a deviation of 17%. Upon setting the FucT concentration to zero, the Hossler et al. model calculated oligosaccharide distributions with a 19.3% deviation from the experimental ones, while the Krambeck & Betenbaugh model had a 20% deviation. These results show that the current model is capable of reproducing the effect of genetic silencing of a glycosylation enzyme to a very high degree whereas the previous models seem incapable of reproducing the effect accurately. The fact that the current model reproduces this phenomenon so effectively makes it useful for exploring cell engineering strategies for the production of mAbs with a desired oligosaccharide profile.

The model for mAb Fc N-linked glycosylation presented in this chapter has approximated the Golgi apparatus to a single PFR so that cisternal maturation is represented. Also, a series of

optimisation-based strategies have been defined to estimate the unknown parameters that define the resident protein concentration profiles within the Golgi apparatus. These optimisation strategies have yielded results that are in agreement with previous reports on these topics. Further optimisation methods have been implemented to estimate kinetic parameters of the system, and with the obtained parameters, the model has shown the capability of reproducing the experimentally observed glycosylation profiles of commercial mAbs. In addition, this model was compared with previously published ones to observe how its novel features contribute to improved representation of mAb N-linked glycosylation.

Inclusion of the appropriate kinetic mechanisms for the glycosylation enzymes involved in the process has been shown to impact model performance considerably. Furthermore, the manner in which the resident protein profiles were defined and the subsequent strategy for estimating unknown kinetic parameters resulted in an improved reproduction of experimentally observed mAb Fc glycan profiles. These promising results were obtained without adding significant complexity to the model, when compared to previous ones.

Finally, this model was shown to replicate the effects of silencing the fucose transferase enzyme. Accurate reproduction of this complex perturbation not only gives us confidence in the model's performance, but also leads us to conclude that many of the underlying assumptions which were made during its development capture several of the fundamental mechanisms that define the N-linked glycosylation process. This, in turn, is consistent with our overall objective of representing the mechanistic links between process conditions and mAb Fc glycosylation.

Chapter 4

Culture dynamics and NSD model

4.1. Cell culture dynamics model

The overall goal of this research project is to produce a mathematical model that connects nutrient availability with cell growth, mAb productivity, nucleotide sugar metabolism and mAb Fc glycosylation. In order to integrate these elements, a model that describes cell culture dynamics as a function of glucose and glutamine was developed. This culture dynamics model was largely based on Monod-type kinetics, although it was further adapted to account for non-ideal effects that were observed experimentally. This model has four constituting elements, which will be described in the following sections:

4.1.1. Cell growth and death

This element of the culture dynamics model is largely based on previous publications from our research group.[246, 272] Within these models, perfect mixing is assumed, and a total material balance can be defined:

$$\frac{dV}{dt} = F_{in} - F_{out} \qquad \text{...Eq. 11}$$

where V is the total cell culture volume within the bioreactor and F_{in} and F_{out} are the volumetric flowrates in and out of the bioreactor. A balance for viable cells within the bioreactor is defined as:

$$\frac{dVX_v}{dt} = \mu VX_v - k_d VX_v - F_{out}X_v \qquad \text{...Eq. 12}$$

here, X_v is the concentration of viable cells $\left[\frac{10^5\ viable\ cells}{L_{bioreactor}}\right]$, μ is the specific growth rate $\left[\frac{10^5\ v.cells_{prod}}{10^5\ v.cells\ hr}\right]$ and k_d is the specific death rate $\left[\frac{10^5\ d.cells_{prod}}{10^5\ v.cells\ hr}\right]$.

The specific growth rate for mammalian cells is frequently defined as a saturation-type function which depends on the essential nutrients present during cell culture:[273]

$$\mu = \mu_{max}\left(\frac{[Glc_{ext}]}{K_{M,glc} + [Glc_{ext}]}\right)\left(\frac{[Gln_{ext}]}{K_{M,gln} + [Gln_{ext}]}\right) \qquad \text{...Eq. 13}$$

where μ_{max} is the maximum growth rate in $\left[\frac{10^5\ v.cells_{prod}}{10^5\ v.cells\ hr}\right]$, $[Glc_{ext}]$ and $[Gln_{ext}]$ are the extracellular concentrations of glucose and glutamine, respectively and their units are $[mM]$. $K_{M,glc}$ and $K_{M,gln}$ represent the Monod constant for growth with respect to glucose and glutamine, and are expressed in $[mM]$.

For this work, the specific death rate was defined as an inverse function of glucose and glutamine availability, with the following form:

$$k_d = k_{d,max,glc}\left(\frac{K_{d,glc}}{K_{d,glc} + [Glc_{ext}]}\right) + k_{d,max,gln}\left(\frac{K_{d,gln}}{K_{d,gln} + [Gln_{ext}]}\right) \qquad \text{...Eq. 14}$$

$k_{d,max,glc}$ is the maximum death rate under glucose deprivation and $k_{d,max,gln}$ is the maximum death rate under glutamine depletion The units of both are $\left[\frac{10^5\ d.cells_{prod}}{10^5\ v.cells\ hr}\right]$. $K_{d,glc}$ and $K_{d,gln}$ are the inverse saturation death constants for either component, and are expressed in $[mM]$.

Equation 14 represents death of cells as a function of extracellular glucose and glutamine depletion. The terms for each nutrient are added because cells may compensate the lack of one nutrient with the other. Specifically, cells have been reported to use glutamine as their sole source of carbon when they are under glucose deprivation.[274, 275] Furthermore, mammalian cells also have the ability to recycle organic nitrogen species and survive entirely on glucose.[274, 275] When $Glc_{ext} \gg K_{d,glc}$ and $Gln_{ext} \gg K_{d,gln}$, the death rate, $k_d \approx 0$. Inversely, under simultaneous glucose and glutamine deprivation, $Glc_{ext} \ll K_{d,glc}$ and $Gln_{ext} \ll K_{d,gln}$, $k_d \approx k_{d,max,glc} + k_{d,max,gln}$. The more interesting cases are when one of the nutrients is depleted while the other one is in excess: when $Glc_{ext} \gg K_{d,glc}$ and $Gln_{ext} \ll K_{d,gln}$, $k_d \approx k_{d,max,gln}$; and when $Glc_{ext} \ll K_{d,glc}$ and $Gln_{ext} \gg K_{d,gln}$, $k_d \approx k_{d,max,glc}$. Therefore, equation 14 describes cell death as a function of depletion of one essential nutrient or the other, but also has the capability of describing the compounded effect of depletion of both nutrients.

4.1.2. Cell metabolism

The rates at which nutrients are consumed is directly associated with their uptake by cellular metabolism to produce biomass (more live cells), product (mAb) and additional intracellular components. A material balance can be performed on the essential nutrients considered throughout this thesis (glucose and glutamine):

$$\frac{dV[Glc_{ext}]}{dt} = -q_{glc}X_vV + F_{in}[Glc_{in}] - F_{out}[Glc_{ext}] \quad \text{...Eq. 15}$$

$$\frac{dV[Gln_{ext}]}{dt} = -q_{gln}X_vV + F_{in}[Gln_{in}] - F_{out}[Gln_{ext}] \quad \text{...Eq. 16}$$

In equations 15 and 16, q_{glc} and q_{gln} represent the specific glucose and glutamine uptake rates, respectively. The units for these terms are $\left[\frac{mmol_{glc,cons}}{10^5\ v.cells\ hr}\right]$ and $\left[\frac{mmol_{gln,cons}}{10^5\ v.cells\ hr}\right]$. The specific uptake for glucose and glutamine are defined as:

$$q_{glc} = \frac{\mu}{Y_{Xv/glc}} + \bar{A}_{cell}m_{glc} \quad \text{...Eq. 17}$$

$$q_{gln} = \frac{\mu}{Y_{Xv/gln}} + \bar{A}_{cell}m_{gln} \quad \text{...Eq. 18}$$

where $Y_{Xv/glc}$ is the biomass yield coefficient for glucose, expressed in $\left[\frac{10^5\ v.cells_{prod}}{mmol_{glc,cons}}\right]$ and $Y_{Xv/gln}$ is the biomass yield coefficient for glutamine, in units of $\left[\frac{10^5\ v.cells_{prod}}{mmol_{gln,cons}}\right]$. Equations 17 and 18, are composed of two additive terms. The one on the left corresponds to the amount of nutrient that is being consumed to produce biomass, and is therefore proportional to the specific growth rate, μ. The second terms, m_{glc} and m_{gln} are the maintenance coefficients for glucose and glutamine, respectively. The units for these terms are $\left[\frac{mmol_{glc,met}}{dm^2_{cell}\ hr}\right]$ and $\left[\frac{mmol_{gln,met}}{dm^2_{cell}\ hr}\right]$.The additional term $\bar{A}_{cell}$ is the average surface area for a single cell in [dm^2].

Conceptually, these terms represent the amount of glucose and glutamine that is being consumed for all other metabolic activity that is *not* related with cell growth. These terms are key for the development of this model and will be discussed further in subsequent sections of this chapter.

The remaining elements to describe cellular metabolism involve the accumulation of metabolites produced by cells in culture. Specifically, the prevalent metabolites produced through cell metabolism are ammonia and lactate. The corresponding material balances for these species are:

$$\frac{dV[Lact]}{dt} = q_{lact}X_vV - F_{out}[Lact] \quad \text{...Eq. 19}$$

$$\frac{dV[Amm]}{dt} = q_{amm}X_vV - F_{out}[Amm] \quad \text{...Eq. 20}$$

The concentration of each species is given in $[mM]$. Equations 19 and 20 include terms for specific productivity for each of the metabolites, q_{lact} and q_{amm}. The units for these terms are $\left[\frac{mmol_{lact,prod}}{10^5\ v.cells\ hr}\right]$ and $\left[\frac{mmol_{amm,prod}}{10^5\ v.cells\ hr}\right]$. A further description of these terms is:

$$q_{lact} = Y_{lact/glc}q_{glc} \quad \text{...Eq. 21}$$

$$q_{amm} = Y_{amm/gln}q_{gln} \quad \text{...Eq. 22}$$

The terms $Y_{lact/glc}$ and $Y_{amm/gln}$ are the yield coefficients for lactate from glucose and ammonia from glutamine, respectively. Equations 21 and 22 express the accumulation of lactate and ammonia during cell culture as a function of the total amount of glucose and glutamine that is being consumed by the cells.

4.1.3. mAb production

The final element regarding extracellular culture dynamics is the one concerning mAb accumulation. Although groups have previously described synthesis and secretion of mAbs more elaborately,[246, 272, 276] for this work, a simpler approach was made:

$$\frac{dV[mAb]}{dt} = q_{mAb}X_vV - F_{out}[mAb] \quad \text{...Eq. 23}$$

where q_{mAb} is the specific productivity of antibody in $\left[\frac{mg_{mAb,prod}}{10^5\ v.cells\ hr}\right]$. Equation 23 shows that mAb accumulation is proportional to the viable cell density, X_v. The mAb specific productivity, q_{mAb} was, in turn, defined as:

$$q_{mAb} = Y_{mAb/Xv} \quad \text{...Eq. 24}$$

In equation 24, it can be seen that the mAb specific productivity is considered to be constant and equal to the yield coefficient, $Y_{mAb/Xv}$, which is expressed in units of $\left[\frac{mg_{mAb,prod}}{10^5\ v.cells\ hr}\right]$.

Equations 11 through 24 have been defined to capture the extracellular dynamics of cell culture, and have included expressions for cell growth and death, as well as nutrient consumption, metabolite accumulation and mAb productivity. At this stage, however, these equations have not considered mechanistic links between the extracellular environment and what occurs inside the cells that are being cultured. The following section will present which elements of the equations presented above are used to couple extracellular culture dynamics with intracellular accumulation of metabolites.

4.2. NSD metabolism

The accumulation of metabolites inside cells may occur through two phenomena. The first one is transport of species from the extracellular environment into the cells. These processes may occur by purely physical phenomena such as diffusion, but are also commonly mediated by biochemical phenomena such as protein-mediated active transport.[277] The second way in which metabolites may accumulate inside cells is via biochemical synthesis from other nutrients supplied in the medium.

Within the context of this work, it is necessary to describe the accumulation of nucleotide sugars (NSDs) within cells so that their availability can be further connected with the glycosylation reactions that occur in the Golgi apparatus. From the NSD biosynthetic pathway presented in Figure 9 of this thesis, it can be seen that all NSDs can be formed using glucose and glutamine as substrates. The initial step towards mechanistically describing the intracellular dynamics of NSDs is to define how glucose and glutamine are accumulated inside cells.

The intracellular accumulation of glucose and glutamine was defined with the following equations:

$$\frac{d\bar{V}_{cell}[Glc_{int}]}{dt} = \bar{a}_{spec,cell}\bar{V}_{cell}m_{glc} - \bar{V}_{cell}\sum_{n=1}^{N.R.}\upsilon_{glc,n}r_n \quad \text{...Eq. 25}$$

$$\frac{d\bar{V}_{cell}[Gln_{int}]}{dt} = \bar{a}_{spec,cell}\bar{V}_{cell}m_{gln} - \bar{V}_{cell}\sum_{m=1}^{N.R.} \upsilon_{gln,m}r_m \quad \text{...Eq. 26}$$

In equations 25 and 26, $[Glc_{int}]$ and $[Gln_{int}]$ are the intracellular concentration of glucose and glutamine. They are both expressed in $[\mu M]$ (hence the conversion factor of 1e3). $\bar{a}_{spec,cell}$ is the average specific area of the cell in $\left[\frac{dm^2_{cell}}{dm^3_{cell}}\right]$; $\bar{V}_{cell}$ is the average volume of the cell in $[dm^3]$; and r_n and r_m are biochemical reactions where intracellular glucose and glutamine are consumed to produce intracellular NSDs.

Equations 17 and 18 defined two terms that describe the uptake of glucose and glutamine that is not destined for biomass production (m_{glc} and m_{gln}). Based on the conceptual meaning of these terms, they have been used as the inlets for intracellular glucose and glutamine accumulation and were defined as:

$$m_{glc} = k_{T,glc}[Glc_{ext}] \quad \text{...Eq. 27}$$

$$m_{gln} = k_{T,gln}[Gln_{ext}] \quad \text{...Eq. 28}$$

A first-order transport rate expression was defined for both glucose and glutamine based on experimental results that will be discussed in subsequent sections of this thesis. In equations 27 and 28, m_{glc} and m_{gln} have units of $\left[\frac{mmol_{glc,internalised}}{dm^2_{cell}\ hr}\right]$ and $\left[\frac{mmol_{gln,internalised}}{dm^2_{cell}\ hr}\right]$, respectively. The units for extracellular glucose and glutamine concentration are $[mM]$. Consequently, the dimensions of $k_{T,glc}$ and $k_{T,gln}$ are $\left[\frac{mmol_{glc,internalised}}{dm^2_{cell}\ mmol_{glc,ext}\ hr}\right]$ and $\left[\frac{mmol_{gln,internalised}}{dm^2_{cell}\ mmol_{gln,ext}\ hr}\right]$.

4.3. NSD reaction network and model reduction

With equations 25 and 26 describing the intracellular accumulation of glucose and glutamine and considering that all NSDs involved in N-linked glycosylation are synthesised using these nutrients as substrates, the next step in model development is to define how NSDs accumulate intracellularly as a function of glucose and glutamine availability. Figure 9 shows the biosynthetic pathway for NSDs in mammalian cells.

The nucleotide sugar biosynthetic pathway contains 34 components which are produced by 36 enzyme-catalysed reactions, of which 26 are reversible and 10 are irreversible. All of these

components are synthesised from simple carbon and nitrogen sources that are transported from the extracellular environment into the cells. These simple carbon sources may be glucose, glutamine, galactose, fructose, mannose, N-acetylglucosamine (GlcNAc) and fucose, among others. Apart from exceptions where cell culture medium has been supplemented with nucleotide sugars,[278] N-acetylmannosamine (ManNAc),[110] fucose analogues[89] and other carbon sources,[105] the essential carbon and organic nitrogen sources for cell culture are glucose and glutamine. As discussed previously in this thesis, glucose and glutamine will be considered herein as the only nutrients that lead to NSD biosynthesis.

A crucial issue when kinetically modelling intracellular biochemical networks is that data for the concentration of many of the intermediate species involved in the pathway are unknown. Moreover, many of these species are difficult to measure given their low intracellular concentration and, in many cases, extremely high turnover.[279-281]

Given these difficulties, the metabolic pathway for NSDs was heuristically reduced so that unknown species along with uncertain complexity is removed from the final model. The considerations behind this heuristic model reduction are:

1. Glucose and glutamine are considered as the only nutrients that contribute to the pathway.
2. Additional components for NSD synthesis, such as nucleotides (ATP, CTP, GTP and UTP) are considered to be in excess and, thus, do not limit accumulation of NSDs. The reasoning behind this is that nucleotides are used by cells as energy currency and also as substrates for fundamental functions such as DNA and RNA synthesis. Elevated cellular requirement for these components lead us to assume that their availability throughout cell culture will be in excess with respect to NSDs.
3. The number of total species has been reduced from 33 to 9 (Glc_{int}, Gln_{int}, UDP-GlcNAc, UDP-Glc, GDP-Man, UDP-GalNAc, CMP-Neu5Ac, UDP-Gal, GDP-Fuc). This has been done because many of the intermediate species along the pathway cannot be measured. Each species that is included in the model and is not measured can be considered as an unconstrained degree of freedom. These in turn, extend the space of possible solutions of the system and potentially reduce model fidelity.[279, 280]
4. Implicit in assumption (3) is the reduction of reactions from 35 in the original pathway to 7. Each reaction in the reduced set represents the synthesis of NSDs either from glucose and glutamine or a preceding NSD in the pathway. Each of these reactions is shown by a coloured arrow in Figure 20. One reaction in Figure 20 uses glucose and glutamine as substrates to produce UDP-GlcNAc (blue arrow). Another

two reactions consume intracellular glucose to produce UDP-Glc (red arrow) and GDP-Man (green arrow). The remaining reactions produce peripheral NSDs from parent NSDs. UDP-GalNAc (orange arrow) and CMP-Neu5Ac (purple arrow) are produced from UDP-GlcNAc. UDP-Gal is produced from UDP-Glc (yellow arrow) and GDP-Fucose is produced from GDP-Man (red arrow).

5. All reported interrelations between the components considered in the reduced reaction scheme have been incorporated in the reaction kinetics. This will be discussed further in this section.

6. Reversibility of reactions has been neglected because the flux of material in the system is assumed to go from glucose and glutamine towards NSDs. The basis for this assumption is that NSDs are continuously transported into the Golgi apparatus, thus driving the flux of the pathway in that direction.

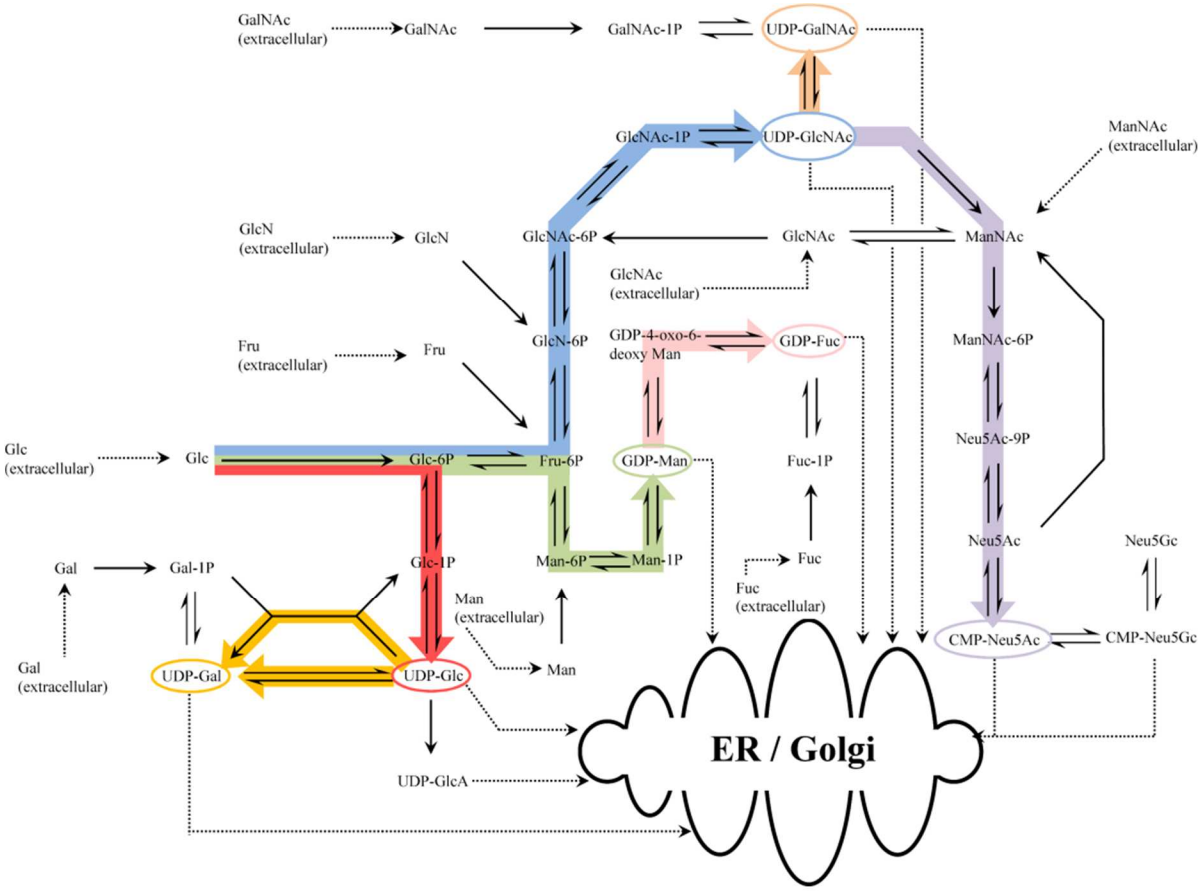


Figure 20. Reduced NSD biosynthetic pathway[23]
The coloured arrows represent the single reactions that were defined from the inputs (glucose and glutamine) to each of the NSDs considered in this work (circled). (■) represents the single UDP-GlcNAc reaction, (□) for UDP-Glc, (■) for GDP-Man, (■) for UDP-GalNAc, (■) for CMP-Neu5Ac, (■) for UDP-Gal and (■) for GDP-Fuc.

Using the assumptions outlined above, and the intracellular glucose and glutamine concentrations as defined by equations 25 and 26, the material balances for intracellular NSDs are:

$$\frac{d\bar{V}_{cell}[UDPGlcNAc]}{dt} = \bar{V}_{cell}(r_{UDPGlcNAc} - r_{UDPGalNAc} - r_{CMPNeu5Ac}) - F_{out,UDPGlcNAc} \quad \text{...Eq. 29}$$

$$\frac{d\bar{V}_{cell}[UDPGlc]}{dt} = \bar{V}_{cell}(r_{UDPGlc} - r_{UDPGal}) - F_{out,UDPGlc} \quad \text{...Eq. 30}$$

$$\frac{d\bar{V}_{cell}[GDPMan]}{dt} = \bar{V}_{cell}(r_{GDPMan} - r_{GDPFuc}) - F_{out,GDPMan} \quad \text{...Eq. 31}$$

$$\frac{d\bar{V}_{cell}[UDPGalNAc]}{dt} = \bar{V}_{cell}(r_{UDPGalNAc}) - F_{out,UDPGalNAc} \quad \text{...Eq. 32}$$

$$\frac{d\bar{V}_{cell}[CMPNeu5Ac]}{dt} = \bar{V}_{cell}(r_{CMPNeu5Ac}) - F_{out,UDPGlcNAc} \quad \text{...Eq. 33}$$

$$\frac{d\bar{V}_{cell}[UDPGal]}{dt} = \bar{V}_{cell}(r_{UDPGal}) - F_{out,UDPGal} \quad \text{...Eq. 34}$$

$$\frac{d\bar{V}_{cell}[GDPFuc]}{dt} = \bar{V}_{cell}(r_{GDPFuc}) - F_{out,GDPFuc} \quad \text{...Eq. 35}$$

The $F_{out}s$ in equations 29 through 35 are defined as a function of their consumption in the glycosylation reactions which occur in the Golgi apparatus. These reactions are, in turn, a function of mAb specific productivity:

$$F_{out,k} = \upsilon_k \left(\frac{1}{\overline{MW}}\right) q_{mAb} \left(\frac{[NSD_k]}{K_{k,glyc} + [NSD_k]}\right) \quad \text{...Eq. 36}$$

Equation 36 was defined to relate the outlet of each NSD with the glycosylation process in the Golgi apparatus and the specific productivity. In equation 36, ν_k is the stoichiometric coefficient of NSD_k consumed by attachment to the mAb-bound glycan $\left[\frac{mmol_{NSD,k}}{mmol_{mAb}}\right]$. $\overline{MW}$ is the

average molecular weight of the mAb in $\left[\frac{mg_{mAb}}{mmol_{mAb}}\right]$. q_{mAb} is the specific productivity of mAb, as defined previously with equation 24, and has units of $\left[\frac{mg_{mAb,prod}}{v.cell\ hr}\right]$. The saturation term on the right hand side of equation 24 was introduced for the outlet term to be consistent with the Golgi NSD transport model defined in equation 9. The saturation constant, $K_{k,glyc}$ has units of $[\mu M]$. These expressions, along with their corresponding parameters will be discussed further in subsequent sections.

With equations 29 through 36 describing the material balances of all considered NSDs, the following step was to define the kinetics whereby these metabolites are produced. As stated in assumption 4 of this section, the number of reactions was reduced so that only the measurable components of the system would be considered. Despite the reasons underlying this assumption, lumping sequential reactions and their corresponding rates may jeopardise model fidelity.[279, 280, 282] In an attempt to avoid losing accurate representation of the system's phenomena and following our objective of capturing the mechanistic behaviour of the process, different kinetic mechanisms for each lumped reaction were defined as a function of reported regulatory behaviour associated with each branch of the pathway.

4.3.1. Lumped NSD kinetic mechanisms

The NSD pathway shown in Figure 20 shares the first two enzymatic reactions with glycolysis – the fundamental metabolic process whereby mammalian cells produce and store energy from glucose.[283] From the intermediary fructose-6-phosphate (Fru-6P) on, glycolysis and NSD metabolism diverge both with respect to biochemical reactions and metabolites. Because of the prevalence of the glycolytic pathway, reactions up to, and including, biosynthesis of Fru-6P have been assumed to occur with unregulated saturation kinetics. From Fru-6P on, the different lumped reactions diverge to produce all NSDs considered in this work. It was therefore necessary to include regulatory features that depend on components considered in the reduced metabolic network defined in Figure 20. The mechanism and regulatory effects for the intermediary reactions along each lumped pathway was retrieved from the BRaunschweig ENzyme Database (BRENDA).[284]

UDP-GlcNAc synthesis

Five enzymes are involved in the synthesis of UDP-GlcNAc from glucose and glutamine. The first reaction is catalysed by glutamine-fructose-6-phosphate transaminase (EC 2.6.1.16) and is the step where the amine group from glutamine is transferred onto Fru-6P. This reaction is reported to have a random-ordered bi-bi mechanism and weak competitive inhibition by UDP-

GlcNAc. The remaining reactions along this lumped pathway have been reported to have ordered ping-pong bi-bi kinetics.[284] Ping-pong mechanisms are those where two irreversible sequential reactions must occur to generate the final product. In ping-pong kinetics an irreversible activation of the enzyme (i.e. by addition of a functional group) must occur before the activated enzyme can catalyse the second reaction.[285] Being sequential reactions, if the rate of one is considered to be much higher than the other, the overall rate expression is reduced to simple Michaelis-Menten kinetics. Within this context, and assuming that the substrates for the fast reactions within ping-pong kinetics are mediated by metabolites that are assumed to be in excess (e.g. nucleotides, as was discussed previously), the kinetics collapse into single substrate Michaelis-Menten expressions. In an analogous fashion, the individual reactions along this branch are sequential, and thus are lumped into a single reaction rate. For these reasons, the overall rate of UDP-GlcNAc synthesis has been defined as a two-substrate, random-ordered rate without inhibitory terms:

$$r_{UDPGlcNAc} = V_{max,UDPGlcNAc}\left(\frac{[Glc_{int}]}{K_{UDPGlcNAc,glc} + [Glc_{int}]}\right)\left(\frac{[Gln_{int}]}{K_{UDPGlcNAc,gln} + [Gln_{int}]}\right) \quad \text{...Eq. 37}$$

where $V_{max,UDPGlcNAc}$ is the maximum turnover rate $\left[\frac{\mu mol_{UDPGlcNAc}}{L_{cell}\, hr}\right]$ and $K_{UDPGlcNAc,glc}$ and $K_{UDPGlcNAc,gln}$ are the saturation constants with respect to glucose and glutamine respectively.

UDP-Glc synthesis

Two sequential reactions are required to synthesise UDP-Glc from glucose-6-P (Glc-6P). The first reaction (from Glc-6P to Glc-1P) is catalysed in parallel by three different enzymes. Two prevalent mechanisms have been described for these enzymes, which are: ordered, ping-pong bi-bi kinetics and single substrate Michaelis-Menten kinetics.[284] Making the same assumptions described for UDP-GlcNAc synthesis (one reaction is much quicker than the other because its substrate is considered to be in excess), this reaction can be simplified to a single substrate Michaelis-Menten kinetic expression.

The second reaction involved in UDP-Glc biosynthesis is from Glc-1P to UDP-Glc. This reaction has been reported to occur by sequential order bi-bi kinetics, where the first substrate to bind to the enzyme is UTP. Again, assuming that nucleotides are in excess within the system, sequential order kinetics may be reduced to single substrate Michaelis-Menten kinetics.

Again, two sequential reactions, whose kinetics have been reduced to simple Michaelis-Menten kinetics, can, in turn, be lumped into a single reaction with the following rate expression:

$$r_{UDPGlc} = V_{max,UDPGlc}\left(\frac{[Glc_{int}]}{K_{UDPGlc,glc} + [Glc_{int}]}\right) \quad \text{...Eq. 38}$$

GDP-Man synthesis

GDP-Man is produced by three sequential reactions that use Fru-6P as the starting substrate. Two of these reactions have been reported to occur by simple Michaelis-Menten kinetics, while the remaining one occurs by a ping-pong bi-bi mechanism. Again, under the rationale that has been presented for the previous NSDs, these three reactions have been lumped into a single substrate Michaelis-Menten kinetic equation:

$$r_{GDPMan} = V_{max,GDPMan}\left(\frac{[Glc_{int}]}{K_{GDPMan,glc} + [Glc_{int}]}\right) \quad \text{...Eq. 39}$$

UDP-GalNAc and CMPNeu5Ac synthesis

UDP-GalNAc is produced by a single epimerisation reaction that uses UDP-GlcNAc as a substrate. This reaction has been reported to have simple Michaelis-Menten kinetics, and thus its rate is defined as:

$$r_{UDPGalNAc} = V_{max,UDPGalNAc}\left(\frac{[UDPGlcNAc]}{K_{UDPGalNAc,UDPGlcNAc} + [UDPGlcNAc]}\right) \quad \text{...Eq. 40}$$

In contrast, CMP-Neu5Ac is synthesised by four sequential reactions using UDP-GlcNAc as a starting substrate. All reactions down this pathway – except one – have been reported to occur with either single substrate Michaelis-Menten kinetics or random-order bi-bi mechanisms. The remaining reaction is catalysed by a bifunctional enzyme (EC5.1.3.14) that sequentially converts UDP-GlcNAc to ManNAc and ManNAc to ManNAc-6P.[286] The ability of catalysing both reactions gives this enzyme interesting properties: it has been reported to have allosteric behaviour with respect to UDP-GlcNAc and inhibitory allorsterism with respect to CMP-Neu5Ac. Allosterism occurs when an enzyme has more than one binding pocket for the substrate(s) and product(s). Binding of susbtrates or products induces conformational changes in the enzyme which, in turn, may expose additional binding sites (positive cooperativity) or, conversely, reduce the available

binding sites (negative cooperativity).[283] Kinetic studies on this enzyme have found that it has a cooperativity coefficient of 0.45 for UDP-GlcNAc and 4.1 for CMP-Neu5Ac.[286] The cooperativity coefficient below one (negative cooperativity) of UDP-GlcNAc implies that binding of this NSD to the enzyme inhibits further substrate binding. This means that excess UDP-GlcNAc inhibits formation of ManNAc and thus, further metabolites down this pathway. The competitive allosterism exhibited by CMP-Neu5Ac, and its reported cooperativity coefficient above one implies that binding of this NSD to the enzyme exposes more sites for further binding (positive cooperativity) which, in turn, inhibits enzyme activity even further. From both allosteric effects, it is possible to conclude that excess UDP-GlcNAc inhibits eventual synthesis of CMP-Neu5Ac, and also that excess CMP-Neu5Ac accumulation strongly inhibits its own synthesis. Because this is the most regulated enzyme in the lumped pathway, we have assumed that it is the slowest reaction, and thus, its rate controls synthesis of CMP-Neu5Ac from UDP-GlcNAc. Being consistent with our attempts to provide mechanistic links between different elements of the glycosylation pathway, the rate for this lumped reaction has been defined as:

$$r_{CMPNeu5Ac} = \frac{V_{max,CMPNeu5Ac}[UDPGlcNAc]^{0.45}}{\left(K_{CMPNeu5Ac,UDPGlcNAc} + [UDPGlcNAc]\right)^{0.45}\left(1 + \frac{[CMPNeu5Ac]}{K_{i,CMPNeu5Ac}}\right)^{4.1}} \quad \text{...Eq. 41}$$

UDP-Gal synthesis

Single reactions are involved in the synthesis of UDP-Gal from UDP-Glc. These have been reported to occur in two different ways. The first is by simple epimerisation of UDP-Glc to form UDP-Gal. These reactions have been reported to follow single substrate Michaelis-Menten kinetics. The other mechanism by which UDP-Gal is produced is through a two substrate ping-pong reaction. As has been done in previous cases, both of these reactions will be lumped into a single one that has simplified Michaelis-Menten kinetics:

$$r_{UDPGal} = V_{max,UDPGal}\left(\frac{[UDPGlc]}{K_{UDPGal,UDPGlc} + [UDPGlc]}\right) \quad \text{...Eq. 42}$$

GDP-Fuc synthesis

Two reactions convert GDP-Man into GDP-Fuc. The first of these reactions converts GDP-Man into GDP-4-keto-6-deoxy-D-mannose through a ping-pong bi-bi mechanism. The second reaction goes from GDP-4-keto-6-deoxy-D-mannose to GDP-Fuc and occurs through a random-order bi-bi mechanism. The ping-pong reaction can be reduced to a simple Michaelis-Menten expression with the considerations that have previously been discussed in this section. The

second reaction, however, has been reported to be non-competitively inhibited by GDP-Fuc.[287] Assuming that this second reaction is slower than the first, and that its secondary substrate (GTP) is assumed to be in excess (as has been the case for all other nucleotides throughout this section), the kinetic expression for this lumped pathway reduces to:

$$r_{GDPFuc} = \frac{V_{max,GDPFuc}\,[GDPMan]}{K_{GDPFuc,GDPMan}\left(1 + \frac{[GDPMan]}{K_{GDPFuc,GDPMan}} + \frac{[GDPMan]}{K_{GDPFuc,GDPMan}}\frac{[GDPFuc]}{K_{i,GDPFuc}} + \frac{[GDPFuc]}{K_{i,GDPFuc}}\right)} \quad \text{...Eq. 43}$$

CMP-Neu5Ac transport into Golgi

A final effect that mechanistically links two of the NSDs considered in this work is the transport of CMP-Neu5Ac into the Golgi apparatus. It has widely been reported that the Golgi transporter for CMP-Neu5Ac is inhibited by presence of UDP-GlcNAc,[111] as was also discussed in section 2.5.1 of this thesis. In an attempt include this effect within our model, the equation corresponding to CMP-Neu5Ac consumption in the Golgi apparatus (Eq. 36) has been modified in the following way:

$$F_{out,CMPNeu5Ac} = \frac{v_k\left(\frac{1}{MW}\right)q_{mAb}\,[CMPNeu5Ac]}{K_{CMPNeu5Ac,glyc}\left(1 + \frac{[CMPNeu5Ac]}{K_{CMPNeu5Ac,glyc}} + \frac{[UDPGlcNAc]}{K_{i,UDPGlcNAc,glyc}}\right)} \quad \text{...Eq. 44}$$

An important consideration behind the model reduction strategy described throughout this section is that, because several sequential reactions have been lumped, all kinetic parameters shown in equations 37-43 are unknown and must be estimated from experimental data. Subsequent sections of this thesis will discuss how the experimental data was produced, how parameter estimation was performed and how the model performs when compared to the experimental data.

This chapter has presented a mathematical model for cell culture dynamics which has been adapted to include a representation of the intracellular accumulation of nutrients and nucleotide sugars. This model has been coupled to the N-linked glycosylation model presented in Chapter 3 of this thesis through intracellular NSD availability and mAb productivity. This coupled model attempts to capture all reported features that *mechanistically* relate extracellular nutrient availability with the intracellular process of mAb N-linked glycosylation.

Chapter 5

Experimental materials and methods

5.1. Cell culture

5.1.1. Cell line and storage

The cell line used for the experiments of this work is the murine hybridoma HFN 7.1 (CRL-1606 from ATCC), which produces an anti-human fibronectin IgG1. HFN 7.1 cells were cultured in a high glucose-containing (4.5g/L) Dulbecco's modified Eagle's medium (DMEM; Gibco, UK) supplemented with 10% v/v foetal bovine serum (FBS; Sigma-Aldrich, UK). For storage, $1x10^7$ cells were suspended in 1mL freeze mix consisting of 95% culture medium and 5% dimethyl sulfoxide (DMSO; Sigma-Aldrich, UK), stored in Nalgene® Mr Frosty at -80°C overnight and subsequently transferred to liquid nitrogen (-160°C). Thawing the cells from liquid nitrogen storage consisted of rapidly agitating the frozen vial in a 37°C water bath for one minute, after which, the contents were suspended in 25mL culture medium and set to incubate in a $75cm^2$ tissue culture flask (TC-75, Nalgene, UK) for 24 hours. One day post thawing, the contents of the TC-75 flask were aspirated and centrifuged at 800rpm for two minutes in a IEC CL30 centrifuge (Thermo Scientific, UK). The cell pellet was re-suspended in 25mL of fresh culture medium and left to incubate for 48 to 72 hours.

5.1.2. Cell culture

Cell culture was performed both in static tissue culture flasks and shake flasks. As mentioned in the previous section, the static tissue culture flasks were mainly employed immediately after thawing. Cells were cultured in suspension once they achieved acceptable viability and growth rates. Suspension cell culture was performed in several working volumes, ranging from 30mL to

400mL in appropriately sized Erlenmeyer flasks (Corning, UK and Nalgene, UK). For all cultures, the cultures were seeded at $2x10^5$ cells mL^{-1} and incubated at 5% v/v CO_2 and at 37°C in a temperature and CO_2-controlled incubator (NUAIRE, UK). The rotation speed for all suspension cultures was maintained at 125 rpm using an orbital shaker (Stuart SSL1, UK).

5.1.3. Extracellular metabolite quantification

Quantification of extracellular glucose, glutamine, lactate and ammonia was performed for all cell culture experiments presented in this thesis. Two millilitre samples were drawn from the cultures, centrifuged at 800rpm for 5min, and finally the supernatant was decanted and stored at -80°C for subsequent analysis. At the time of analysis, the frozen samples were thawed and analysed using a Bioprofile 400 system (Nova Biomedical, UK). Briefly, the Bioprofile 400 system performs automated enzymatic assays on the samples. For glucose, it measures the H_2O_2 released during the enzymatic oxidation of glucose in presence of glucose oxidase. Glutamine is measured through the production of glutamate by the glutaminase-catalysed hydrolysis of glutamine. Lactate is measured through the H_2O_2 produced by the oxidation of lactate in presence of the lactate oxidase enzyme. Ammonia is directly measured by electrochemical means.

5.1.4. mAb quantification

The titre of mAb produced through cell culture was obtained through an indirect sandwich enzyme-linked immunosorbent assay (ELISA). This method was carried out in cell culture treated 96-well plates (NUNC, Denmark) using the following protocol previously described by Kontoravdi et al.:[272]

1. Add 100µL of 1µg/mL anti-human fibronectin antibody from rabbit (Sigma_Aldrich F3648) suspended in coating buffer (0.05M sodium bicarbonate, pH 9.6) to each well and incubate overnight at 4°C on a level surface.

2. Block non-specific binding by adding 250µL of blocking buffer (250mL coating buffer with 1.25g casein Hammerstein grade) to each well and incubate at room temperature for one hour.

3. Empty the plate and wash three times using washing buffer (phosphate buffered saline with 0.05%v/v TWEEN-20).

4. Add 100 µL of 0.2µg/mL human fibronectin (Chemicon international FC010) suspended in PBS to each well. Incubate at room temperature for one hour.

5. Empty the plate and wash three times using washing buffer.

6. Add 100μL of standards and samples containing HFN 7.1 antibody (Pierce Antibodies MA5-12314) suspended in PBS at a dilution of 1:2000 (maximum concentration of mAb per well should not exceed 80μg/L) to each well. Incubate at room temperature for one hour.

7. Empty the plate and wash three times using washing buffer.

8. Add 100μL of 0.64μg/mL anti-mouse Fc antibody from goat (Sigma-Aldrich A0168) suspended in PBS to each well. Incubate at room temperature for one hour.

9. Empty the plate and wash three times using washing buffer.

10. Dissolve one tablet of 3,3',5,5'-tetramethylbenzidine (TMB, Sigma-Aldrich T3405) in 10mL of 0.05M phosphate-citrate buffer, pH 5.0 (substrate buffer). Add 2μL of fresh 30%v/v hydrogen peroxide solution per 10mL of substrate buffer solution immediately before use. Add 100μL of the resulting solution to each well. Incubate at room temperature for 30 minutes.

11. Add 50μL of 2.5M H_2SO_4 solution to each well so that the enzymatic reaction is stopped.

12. Read the plate for absorbance at 450nm on a microplate reader (BioTek Instruments, USA ELX808U).

5.1.5. Cell density and viability

Viable and dead cell density was measured by performing counts on trypan blue treated cells on a Neubauer rulings haemocytometer. The following protocol was followed for this procedure:

1. The surface of the haemocytometer and the cover slip were cleaned using 70% v/v aqueous ethanol. The cover slip was placed on the chamber prior to adding the sample.
2. 20% v/v trypan blue solution was added to the sample so that the end ratio (sample:trypan) was 1:1. Trypan blue is a dye that permeates dead cells while leaving the live cells uncoloured.
3. The trypan blue treated sample was applied to the haemocytometer cell with a pipette.
4. The haemocytometer was placed on the microscope and read at 10x magnification.
5. The count was performed on the centre and corner squares of the counting grid.
6. The total count should lie between 50 and 150 cells. In cases where the count is outside this range, the cell sample was diluted with water prior to trypan blue addition.

7. Cell density was obtained by dividing the count by total volume considered (each grid square has a surface area of 1mm^2 and a depth of 0.1mm). The calculation was:

$$Cell\ density = \frac{cell\ count \times dilution}{5\ squares \times 1e^{-4}mL}[=]\frac{10^4 cells}{mL}$$

5.2. NSD quantification

5.2.1. HPAEC for NSD quantification

Due to the relevance of NSDs for protein glycosylation, several analytical methods have been developed for quantifying their intracellular pools.[263, 288-291] In all methods, two properties of nucleotides and nucleotide sugars have been exploited. Firstly, nucleotides and NSDs contain conjugated double bonds in their nucleotide component, as can be seen in Figure 8. These conjugated double bonds make these molecules optically active, with maximum absorbance wavelengths between 257.4nm and 271.6nm. These wavelengths are readily detectable with UV-vis detectors commonly available on analytical equipment. Secondly, nucleotides and nucleotide sugars are charged molecules at neutral pH because of the phosphate groups they contain.

In all four methods cited, detection of NSDs has been performed by conventional UV-vis detectors and the separation of these compounds has been based on their charge. Tomiya et al.[288] achieved separation and quantification of nine nucleotides and nine nucleotide sugars by using high-performance anion exchange chromatography (HPAEC). Both Kochanowski et al.[263] and Nakajima et al.[291] developed ion-pairing reverse chromatographic methods. The former group was able to resolve eight nucleotides and five NSDs in 47min using a C18 column, tetrabutylammonium bisulphate as the ion pairing agent and methanol as the organic modifier. The latter group resolved twelve nucleotides and eight NSDs in 30min with an Inertsil C4 column, the same ion-pairing agent and acetonitrile as the organic modifier. Feng et al.[290] separated and quantified seven nucleotide sugars and thirteen nucleotides in 40min with a capillary electrophoresis method.

Based on resolving power, apparent robustness and the analytical capabilities of our laboratory, the method developed by Tomiya et al. was selected as our starting point for NSD quantification. Tomiya et al.'s original method was performed on a 10ADvp Shimadzu HPLC under the conditions presented in Table 13.

Table 13. Tomiya et al.[288] HPAEC conditions for NSD quantification

Conditions	Values
Column	CarboPac PA-1 (Dionex, US)
Guard column	PA-1 guard column (Dionex, US)
Flowrate	1mL/min
Eluent 1 (E1)	1mM NaOH
Eluent 2 (E2)	1.0M CH_3COONa in 1mM NaOH
Column Temperature	30°C
Sample Temperature	NA
Gradient	t_{0min} = 20% E2; t_{10min} = 55% E2; t_{25min} = 55% E2; t_{35min} = 80% E2; t_{40min} = 100% E2; t_{50min} = 100% E2
Wavelengths monitored	262.1nm; 271.6nm

The Tomiya et al. method was implemented in our laboratory on a Waters Alliance HPLC (UK) which is composed of a 2695 series separations module coupled to a 2998 model photodiode array detector. Preparation of the eluents was as follows:

E1:

1. 1L of deionised water (18.2MΩ) was sparged with helium for 5 minutes to displace any dissolved CO_2.
2. Once sparged, the appropriate amount of 50% w/v (19.2M) NaOH (Sigma-Aldrich, UK) was added and the liquid gauged to 1L. For 1mM NaOH solutions, 52.1µL of concentrated NaOH solution were added. For 3mM solutions, 156.3mL were added.
3. The solution was mixed and filtered through a 0.2µm filter unit (Nalgene, UK) to eliminate any particulate matter in the solution.

E2:

1. The appropriate amount of >99% sodium acetate (Sigma-Aldrich, UK) was added to 1L of E1 eluent. 82.03g were added to E1 to prepare 1M CH_3COONa solutions and 123.04g were added for 1.5M CH_3COONa solutions.
2. The solution was stirred until the CH_3COONa was completely dissolved and then filtered through a 0.2µm filter unit (Nalgene, UK).

Initially, the gradient and eluent composition reported by Tomiya et al. was used, and produced the chromatogram presented in Figure 21 for a mixture of eight NSDs and three nucleotide standards. In Figure 21, it can be seen that out of eleven standards that were present in the injected mixture, only nine were resolved. Each peak was assigned by comparing the observed retention times with those obtained from single standard injections. Peak assignment was further confirmed by comparing the spectrum of the pure components with those measured in the mixture runs.

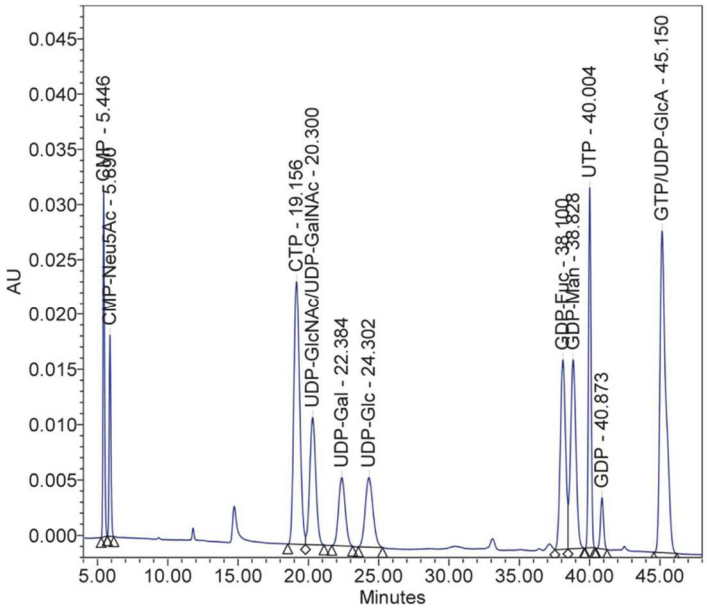


Figure 21. Chromatogram replicating the method by Tomiya et al.[288]

Because not all species were resolved, the method was adapted to the conditions in our laboratory to obtain improved NSD separation. In order to do so, and based on the method development reported by Tomiya and collaborators, the concentration of NaOH in the eluents was varied. Conveniently, the allowable range for NaOH concentration was narrow, with Tomiya et al. having reported that nucleotides and NSDs degrade at NaOH concentrations above 4mM. The NaOH concentration of the eluents was varied between 0.5mM and 3.5mM. With these trials, it was found that at a NaOH concentration of 3mM, resolution of all eleven species was achieved, as can be seen in Figure 22.

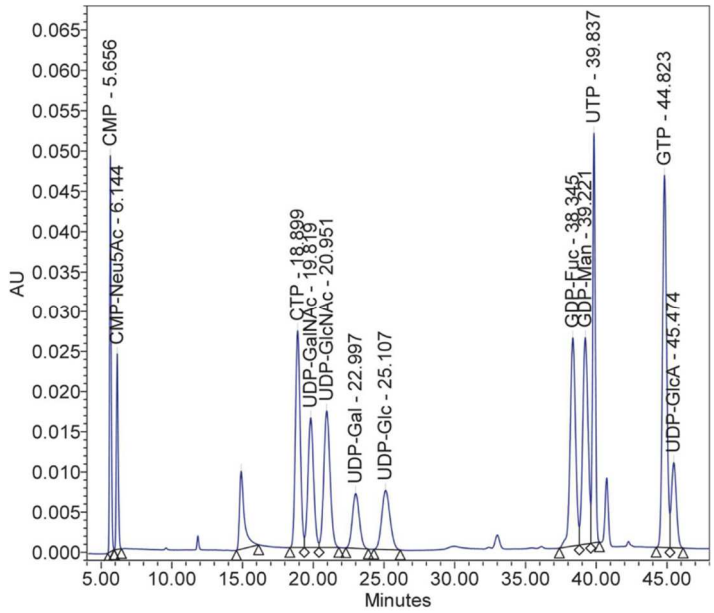


Figure 22. Chromatogram with 3mM NaOH in E1 and E2

As a further effort in optimising the method, we set out to reduce the analysis runtime. As discussed in section 2.7.2 of this thesis, in HPAEC, analytes are adsorbed onto the stationary phase through coulombic interactions. A second anion (CH_3COONa) is then flushed through the system, and due to competing interactions, the analytes are eluted differentially from the stationary phase. Considering the long E2 gradients defined by Tomiya et al. and the principles behind HPAEC, the gradient of E2 was increased in order to reduce analysis time. Having tested several gradient profiles, it was found that the chromatogram contracted in time without compromising resolution of the eleven standards. However, because the gradients that were being tested were exceeding the capabilities of the HPLC, the concentration of CH_3COONa in E2 was increased from 1M to 1.5M. The higher initial concentration of CH_3COONa in E2 allowed for shallower volumetric gradients of E2 while achieving higher CH_3COONa concentrations required to elute the NSDs bound to the column. The optimized conditions for the method are summarized in Table 14. It is worth noting that maximum concentration of CH_3COONa that flows through the column is 1M. This value was not exceeded as it is the maximum recommended CH_3COONa concentration for CarboPac PA1 columns by their manufacturer. A chromatogram showing full separation of eleven NSDs and three nucleotides in 20 minutes is presented in Figure 23.

Table 14. Optimised HPAEC conditions for NSD quantification

Conditions	Values
Column	CarboPac PA-1 (Dionex, US)
Guard column	PA-1 guard column (Dionex, US)
Flowrate	1mL/min
Eluent 1 (E1)	3mM NaOH
Eluent 2 (E2)	1.5M CH_3COONa in 3mM NaOH
Column Temperature	30°C
Sample Temperature	5°C
Gradient	t_{0min}: 20% E2; t_{6min}: 66% E2; t_{18min}: 66% E2; t_{23min}: 20% E2; t_{30min}: 20% E2
Wavelengths monitored	262.1nm; 271.6nm

In Figure 23, it is seen that all injected standards were separated in under 20 minutes after injection. It must be mentioned that baseline separation was not achieved for all species, especially GDP-Fuc and GDP-Man. However, this did not seem to affect accuracy of quantification in these cases and, to our belief, the reduced runtime (<30min) compensates for this issue. Additional parameters regarding reproducibility, sensitivity and quantification linearity will be detailed in the following section.

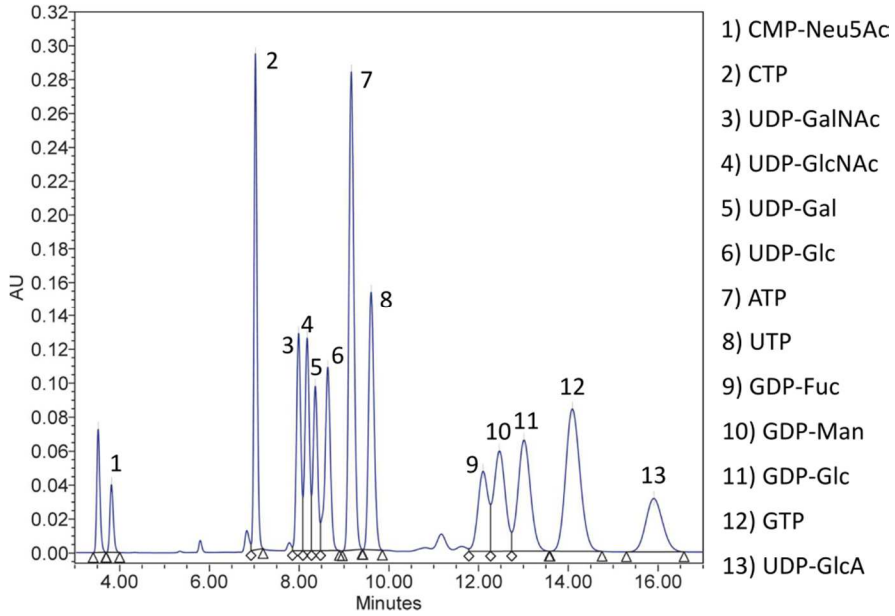


Figure 23. Representative chromatogram of the optimised method

Reproducibility, linearity and sensitivity

The optimised method described in preceding sections was assessed for reproducibility, sensitivity, and quantification linearity. As a means to establish reproducibility, the retention times for each analyte and its variation over 50 injections is presented in Table 15. From the data, we see that retention times vary below 1% of the mean value, except GTP and UDP-GlcA, which have variations of 1.27% and 1.13% respectively. Considering the number of species that are being resolved in less than 30 minutes, we believe these variations are acceptable and confirm the robustness of the separation.

Table 15. Retention time variation over 50 injections

Component	Mean t_{ret} (min.)	Std. Dev. (min.)	RSD
ATP	9.245	0.026	0.28%
CMP-Neu5Ac	3.794	0.009	0.23%
CTP	7.049	0.010	0.14%
GDP-Fuc	12.353	0.072	0.58%
GDP-Glc	13.332	0.090	0.68%
GDP-Man	12.743	0.079	0.62%
GTP	14.697	0.175	1.19%
UDP-Gal	8.424	0.020	0.24%
UDP-GalNAc	8.033	0.016	0.20%
UDP-Glc	8.713	0.024	0.27%
UDP-GlcA	16.521	0.175	1.06%
UDP-GlcNAc	8.229	0.017	0.21%
UTP	9.747	0.044	0.45%

It must also be noted that the 50 injections considered for Table 15 were performed over a period of three months and were performed with different batches of eluents. The data suggests that, with respect to peak identification and separation, the chromatographic method, along with preparation of the different eluents, is indeed consistent and reproducible.

Finally, the range of linear quantification with this HPAEC method was determined by generating a series of calibration curves. These, in turn, were subsequently used to quantify the pools of nucleotides and NSDs from cellular extracts. To generate the calibration curves, known amounts of standards were mixed and serially diluted with deionised water (18.2MΩ). In order to populate the calibration curves further, different volumes (30µL, 25µL, 20µL, 15µL and 10µL) of each dilution were injected. The amount of each nucleotide and NSD was defined to ensure that the measured intracellular pools for each component fell within the range set by the calibration curve for each species. Standards for ATP, CMP-Neu5Ac, CTP, GDP-Fuc, GDP-Man, GDP-Glc, GTP, UDP-Gal, UDP-GalNAc, UDP-Glc, UDP-GlcA, UDP-GlcNAc and UTP were purchased from Sigma-Aldrich (UK). All calibration curves were done in triplicate so that variation in measured concentration was assessed. The area of each peak was plotted against the known amount of each analyte added to the standard mixtures. These curves were fitted to a linear equation with intercept equal to zero:

$$pmol\ injected = m \times measured\ area\ below\ the\ peak$$

Table 16 presents the obtained values for the slope, $\boldsymbol{m}$, along with the coefficient of determination, $\boldsymbol{R^2}$, which is a measure of how well the data fits the linear equation.

Table 16. Calibration curve slope and linearity

Component	m	R^2
ATP	1.13E-03	0.99967
CMP-Neu5Ac	5.82E-03	0.99986
CTP	1.79E-03	0.99968
GDP-Fuc	1.65E-03	0.99580
GDP-Glc	1.53E-03	0.99587
GDP-Man	2.36E-03	0.99977
GTP	1.33E-03	0.99995
UDP-Gal	4.94E-03	0.99799
UDP-Glc	2.91E-03	0.99899
UDP-GlcA	2.82E-03	0.99477
UDP-GlcNAc	3.09E-03	0.99117
UTP	2.87E-03	0.99833

The value for the coefficient of determination is above 0.99 for all species. These values confirm that, within the range at which intracellular nucleotide and NSD pools are expected, quantification with the optimised method is linear and accurate. The calibration curves are presented in Appendix V.

Table 17. Optimised HPAEC sensitivity parameters

	52 pmol Injections			78 pmol Injections		
Compound	Mean Area (mV sec)	Std. Dev. (mV sec)	RSD (%)	Mean Area (mV sec)	Std. Dev. (mV sec)	RSD (%)
ATP	47,548	447	0.9%	72,791	626	0.9%
CMP-Neu5Ac	9,239	104	1.1%	13,994	124	0.9%
CTP	30,570	159	0.5%	46,577	211	0.5%
GDP-Fuc	32,938	1,572	4.8%	51,890	496	1.0%
GDP-Glc	34,349	3,205	**9.3%**	55,739	312	0.6%
GDP-Man	20,831	3,430	**16.5%**	33,768	140	0.4%
GTP	34,957	315	0.9%	56,388	201	0.4%
UDP-Gal	10,341	252	2.4%	15,487	93	0.6%
UDP-GalNAc	17,744	484	2.7%	26,378	140	0.5%
UDP-Glc	15,794	868	**5.5%**	22,428	467	2.1%
UDP-GlcA	**ND**	**ND**	**ND**	20,509	260	1.3%
UDP-GlcNAc	17,858	344	1.9%	26,404	154	0.6%
UTP	25,581	187	0.7%	38,892	132	0.3%

Method sensitivity was assessed by finding the minimum amount of sample injected that was quantified reliably. The lowest amount of sample injected for all analytes was 52pmol. These samples were injected in triplicate, and the standard deviation was calculated for each species across these injections. A threshold for relative standard deviation (RSD – i.e. the quotient between the mean value and the standard deviation) above 5% was defined as the limit for reproducible quantification. When the RSD exceeded 5%, the process was repeated for the next lowest injected quantity. With the 52pmol injections, all but three species had RSDs below 5%. The ones that exceeded 5% were GDP-Glc, GDP-Man, UDP-GlcA and UDP-Glc. The following lowest amount injected for these species was 78pmol. At this level, all species had an RSD below 5%. These results are summarised in Table 17.

From the data presented in Table 17, it can be concluded that the limit for reliable quantification with the optimised HPAEC method is 52pmol for all species except GDP-Glc, GDP-Man, UDP-GlcA and UDP-Glc. For these remaining species, the limit of quantification is 78pmol. When this amount is injected, the RSD for measured area in triplicate injections is below 5%, with the highest observed value being 2.1% for UDP-Glc. Overall, these results indicate that when a sample contains at least 78pmol of a given nucleotide or NSD, quantification is reliable and with deviations below 5%.

Considering that the lowest reported intracellular concentration for NSDs is UDP-Gal with 9.2µM (seen in Table 10 of this thesis), and that the maximum possible injection volume is 30µL on the HPLC, the quantity of injected UDP-Gal would be 276pmol, which is well above the 78pmol limit of detection of the current method. This suggests that the optimised method presented herein is capable of accurately quantifying the intracellular pools of all nucleotides and NSDs involved in mAb N-linked glycosylation.

5.2.2. Extraction of intracellular NSDs

Recently, several groups have discussed the impact of quenching, washing and extraction methods on the quantification of intracellular metabolites.[292, 293] The objective of these procedures is to ensure that i) the extraction method allows for reproducible retrieval of the desired metabolites, ii) that the metabolite pools to be extracted are representative of the metabolic state of the cells at the moment of extraction, and iii) that components from the cell culture medium do not contaminate the intracellular samples. In order to address these concerns, a series of experiments were performed so that a standard methodology for nucleotide and NSD extraction could be established.

Dietmair and collaborators recently published a comparison of 3 quenching methods and 12 extraction protocols for recovery of intracellular glucose, aminoacids, nucleotides and nucleotide sugars.[293] Dietmair et al. quenched using 60% aqueous methanol (MeOH) at -40°C, 60% aqueous MeOH buffered with ammonium bicarbonate (AMBIC) at -40°C and ice-cold 0.9% w/v sodium chloride (NaCl) and found that both methanol solutions damaged the cell membrane and caused leakage of intracellular metabolites. These findings contrast with those reported by Sellick et al., where it was found that AMBIC buffered MeOH yielded the highest metabolite recoveries.[292] However, the report by Sellick et al. focused on extraction efficiencies for metabolomics studies and did not assess extraction of nucleotides and NSDs. Based on the publication by Dietmair et al., 0.9% NaCl was selected as the quenching and washing solution in the development of this protocol.

Dietmair et al. also explored different extraction protocols, which included acetonitrile (ACN), MeOH freeze, cold 50% MeOH, MeOH/chloroform, hot 80% MeOH, cold 100% MeOH, hot EtOH, hot EtOH/HEPES, cold EtOH, hot H_2O, potassium hydroxide (KOH) and perchloric acid (PCA). From the data reported by this group, seven protocols achieved acceptable extraction of nucleotides and NSDs.[293] Based on this, two extraction methods (PCA and ACN) were selected for comparison. The protocol for PCA extraction was based on Dietmair et al.[293] and was the following:

1. Cell pellets were resuspended in 200μL of 0.5M ice-cold PCA (Sigma-Aldrich).
2. The above mixture was incubated on ice for 10 minutes.
3. The incubated mixture was centrifuged for 5min at 18,000g and 1°C in Hermle Z233 MK-2 centrifuge (Germany).
4. 200μL of the supernatant were removed and 40μL of 1.1M K_2HPO_4 (Sigma-Aldrich, UK) in 2.5M KOH (Sigma-Aldrich, UK) was added to the mixture so that the PCA was neutralized.
5. The neutralised mixture was centrifuged for 5min at 18,000g and 1°C.
6. 200μL of the supernatant was taken and stored at -80°C until HPAEC analysis.

The volume of added PCA was reduced from 1000μL in Dietmair et al. to 200μL and KOH from 200μL to 40μL so that the end extract could be injected directly into the HPLC without concentrating the sample.

The protocol for ACN was also based on Dietmair et al.:[293]

1. Cell pellets were resuspended in ice cold 50% aqueous acetonitrile (Sigma Aldrich). 1mL of 50% ACN was added per $5*10^6$ cells.
2. The cell suspension was incubated on ice for 15 minutes.
3. The incubated mixture was centrifuged for 5min at 18,000g and 1°C in a Hermle Z233 MK-2 centrifuge (Germany).
4. Samples were stored at -80°C until HPAEC analysis.
5. Immediately prior to HPAEC analysis, the samples were thoroughly dried using a SpeedVac (Savant SC210A, UK) and then resuspended in 240μL of deionised water.

For comparing ACN and PCA extractions, samples of protein-secreting cells were taken on day six of culture. Six equal volume samples were taken, three to be extracted with PCA and three to be extracted with ACN. The samples were then analysed with the HPAEC method described in the previous section, and the results are shown in Figure 24.

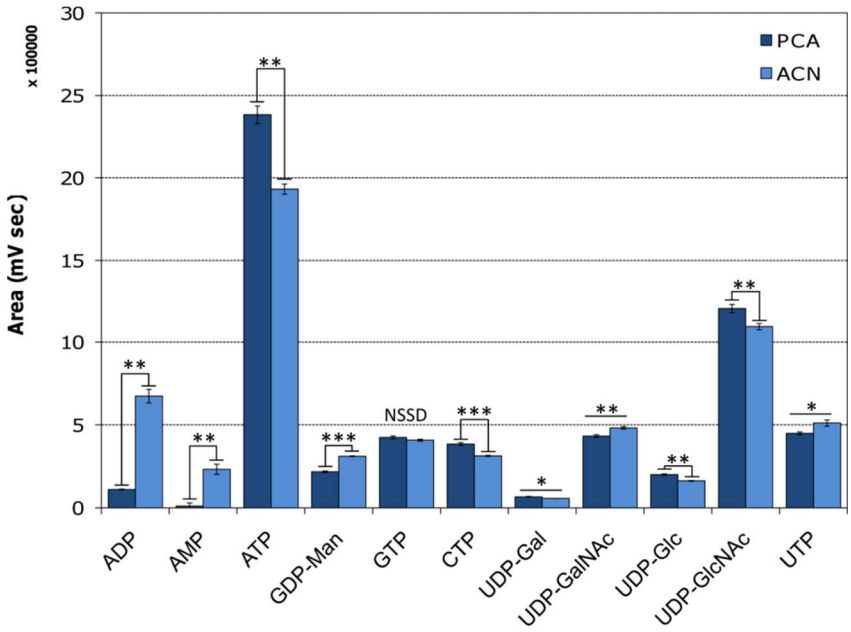


Figure 24. Comparison of PCA and ACN for nucleotide and NSD extraction
The dark bars show HPAEC results from PCA extraction, while the light bars show ACN extraction. Each extraction was performed in triplicate, and a t-test assuming unequal variances was performed to assess differences between the two treatments. Three stars represent a significant difference with a $p<0.001$; two stars correspond to $p<0.01$ and one star is for $p<0.05$. NSSD means there is no statistically significant difference between the two treatments.

Figure 24 shows that both extraction protocols yield different pools for all nucleotides and NSDs with the exception of GTP. In general, the extracted metabolites are lower for the ACN extraction than for PCA. However, there are three cases where ACN produces higher extraction, which are GDP-Man, UDP-GalNAc and UTP. PCA extracts 43%, 11% and 14% less of these metabolites, respectively. In contrast, PCA produces higher pools of ATP (19%), CTP (19%), UDP-Gal (14%), UDP-Glc (19%) and UDP-GlcNAc (10%).

An interesting case is that of ATP, ADP and AMP. Figure 24 shows that PCA produces more ATP (20%), while ACN generates 5 times more ADP and 20 times more AMP than PCA. These considerable differences may be attributed to degradation of ATP to form ADP and AMP during the concentration step using the SpeedVac, given that this process is lengthy (6-10 hours) and performed at room temperature.

Overall, it is possible that ACN has better extraction efficiency of some nucleotides and NSDs. However, it has also been shown that the concentrating step required for ACN extraction may lead to metabolite degradation. It is worth noting that the time required for concentrating the samples for the ACN protocol are not consistent and could lead to undesired sample-to-sample variation. For these reasons and for the sake of consistency, PCA was selected as the most appropriate nucleotide and NSD extraction protocol for the work presented herein.

5.2.3. Quenching and washing of cell samples

The final element to consider was whether quenching and washing with 0.9% NaCl influenced the measurement of nucleotide and NSD pools. In order to evaluate the effect of quenching and washing with 0.9% w/v NaCl, nine samples of equal volume were taken at day six from a culture of protein secreting cells. Three of the samples were left untreated, another three were quenched with 4x volumes of ice-cold 0.9% w/v NaCl, and the remaining three were quenched and washed with 0.9% NaCl. The overall procedure for quenching and washing was:

1. A 3mL sample from cell culture is taken and immediately added to 12mL of ice-cold 0.9% w/v NaCl (Sigma-Aldrich).
2. The quenched mixture is then centrifuged for 1 minute at 1000g on an IEC CL 30 centrifuge (Thermo Scientific, UK).
3. The supernatant is decanted and the pellet is resuspended with another 12mL of ice-cold 0.9% NaCl.
4. The washed mixture is centrifuged for 1 minute at 1000g.
5. The supernatant is decanted and the desired extraction protocol is performed on the cell pellet.

The untreated samples (NQ) were not subjected to any of the steps described above and their contents were extracted directly after having been centrifuged. The quenched samples (Q) were extracted after step two of the procedure described above. Finally, the quenched and washed samples (QW) were treated as described above. After the quenching and washing procedure, the pellets were extracted using the PCA method described earlier in this section and subsequently injected for the HPAEC analysis which was also developed in this work. The results for these experiments are shown in Figure 25.

In Figure 25, it can be seen that there is no significant difference between samples that were not quenched (NQ) and only quenched (Q). It is apparent, though, that there are significant differences between NQ and Q with samples that have been quenched and washed. More specifically, there is substantially lower recovery for all nucleotides when the samples were quenched and washed, when compared to the NQ and Q data (>30% lower recovery in all cases).

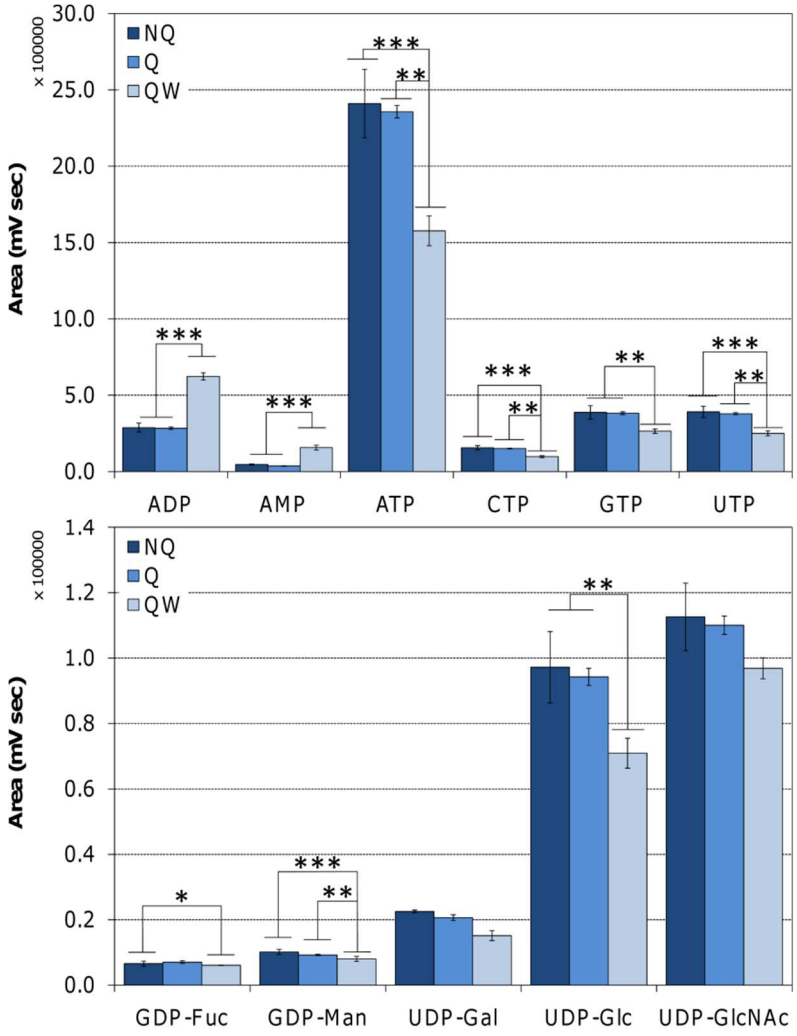


Figure 25. Comparison of quenching and quenching and washing with 0.9% NaCl
Data for non-treated (■), only quenched (■) and quenched and washed (□) samples are presented. Each treatment was done in triplicate, and an HSD Tukey's test was performed on the data. A significant difference with a p<0.001 is represented by (***), p<0.01 with (**) and p<0.05 with (*). No stars are presented for data that are not significantly different from each other.

A similar trend is observed with the NSDs, with the exception of UDP-Gal and UDP-GlcNAc, where no significant differences were measured between all three treatments. The reduction in recovery of NSDs when the cell pellet was quenched and washed was 7.5% for GDP-Fuc, 20% for GDP-Man and 27% for UDP-Glc. Also, and in a similar trend to the one observed for the extraction methods, a decrease in ATP recovery of 35% was observed in the QW treatment, when compared to the NQ treatment. Again, a corresponding increase in AMP and ADP was seen in the QW treatment. It is possible that this effect is due to degradation of ATP during the washing and second centrifugation step associated with the QW treatment. From this information, it is possible to conclude that washing is detrimental for adequate recovery of nucleotides and NSDs. Moreover, because there is no significant effect observed from

quenching, this step was also omitted from the overall protocol for extraction and quantification of nucleotides and NSDs developed in this work.

With the findings presented in this section, the final protocol that was defined for extracting nucleotides and NSDs from cells was the following:

1. A sample of $1x10^6$ cells is taken from the cell culture.
2. The sample is centrifuged at 1000g for 1 minute.
3. The supernatant is decanted and any excess supernatant is removed with a pipette.
4. The cell pellet is then treated with PCA as described in previous sections of this thesis.
5. The neutralised extract is stored at -80°C until HPAEC analysis.
6. Immediately prior to HPAEC, the frozen sample is quickly thawed and passed through a 0.2µm syringe-driven filter unit (Millipore, UK).
7. The sample is transferred to total recovery vials (Waters, UK) and placed in the temperature-controlled autosampler of the HPLC, where it remains at 5°C until it is injected for HPAEC analysis.

A representative chromatogram of intracellular nucleotides and NSDs that were extracted following the protocol described above is presented in Figure 26.

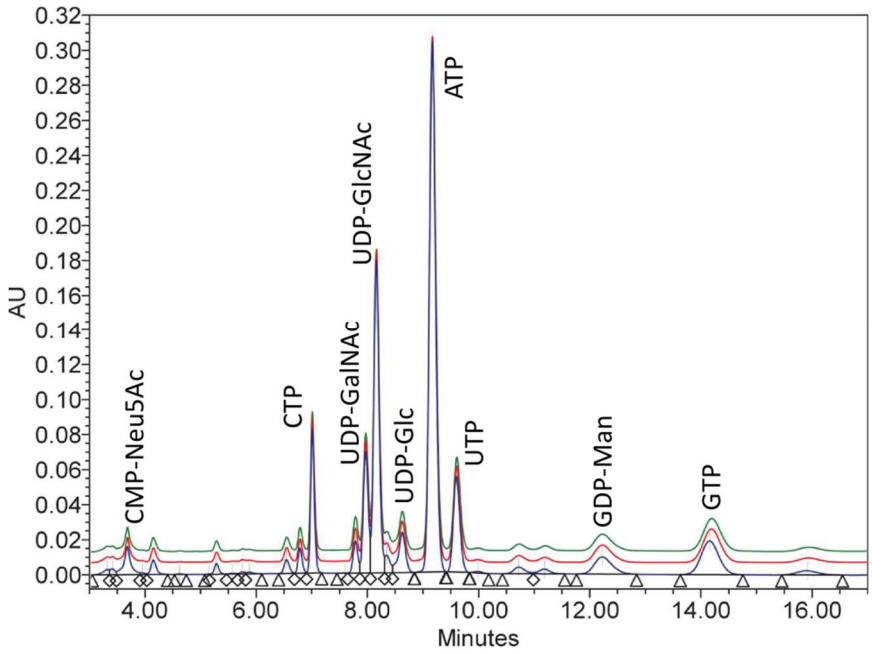


Figure 26. Typical HPAEC chromatogram of intracellular nucleotide and NSD pools
The chromatogram shows adequate separation of four nucleotides and six NSDs that have been extracted from cells using the methods described above. The overlay shows reproducibility, both of the analytical method, as well as the extraction protocols employed herein.

The methods described in this and preceding sections culminate with the chromatogram presented in Figure 26. In this figure, it can be seen that nucleotides and NSDs have been

successfully extracted from cells and have been separated adequately for quantification. Confidence in the accuracy of the measured nucleotide and NSD pools stems from the presented work regarding the optimised HPAEC technique for nucleotide and NSD quantification, the experiments that were performed to select the appropriate intracellular metabolite extraction protocol and the work shown on quenching and washing methodologies.

The work described throughout this chapter has produced an optimised HPAEC-based analytical method for accurately quantifying intracellular pools of nucleotides and NSDs in less than 30 minutes. In addition, perchloric acid (PCA) extraction of intracellular metabolites was selected over acetonitrile (ACN) because the latter methodology was observed to be prone to irreproducible degradation of intracellular nucleotides and NSDs. Finally, quenching and washing steps prior to metabolite extraction were omitted after it was observed that quenched samples yielded statistically indistinguishable nucleotide and NSD pools, and that the addition of a washing step had a detrimental effect on the recovery of these metabolites. Results yielded from these methods and their use towards the overall goal of this work will be presented in subsequent sections of this thesis.

5.3. Intracellular glucose analysis

For intracellular glucose analysis, cell samples were taken from culture in parallel to those taken for NSD analysis. However, these samples were quenched with four volumes of 0.9%w/v NaCl and further washed with four volumes of the same solution, as described in section 5.2.3. In contrast with NSDs, the quenching and washing procedure was performed prior to intracellular glucose analysis because large quantities of glucose are present in the cell culture medium (DMEM; GIBCO, UK). If the cell pellet was not quenched and washed, the observed intracellular glucose concentrations would be overestimated and inaccurate due to contamination from the medium. Extraction of intracellular glucose was performed using PCA, as described in section 5.2.2.

Once PCA extraction was performed, the intracellular glucose concentration was measured by the Amplex Red glucose assay kit from Invitrogen (UK). Amplex Red is an enzymatic assay where glucose is reacted with glucose oxidase to produce glucolactone and hydrogen peroxide (H_2O_2). H_2O_2 is then reacted with the Amplex Red reagent (on a 1:1 ratio) in presence of horseradish peroxidise (HRP) to produce resorufin. This is a fluorescent product, with an emission and excitation length of 571nm and 585nm, respectively. Because resorufin is reacted on a 1:1 ratio with glucose, detected fluorescence of this compound is proportional to the concentration of glucose in the sample.

The Amplex Red assay was performed according to the manufacturer's instructions in non-treated 96-well plates (NUNC, Denmark). Fluorescence of the Amplex treated samples was measured on a Tecan Infinite 200Pro fluorescence plate reader (Tecan, UK) with an excitation wavelength of 568nm and emission wavelength of 580nm.

5.4. Mass Spectroscopy for glycan analysis

The objective of this work is to produce an experimentally-validated mathematical model that relates mAb Fc N-glycosylation with nutrient availability during cell culture. Because mAb Fc glycosylation lies at the centre of this work, means to determine the glycan distributions of mAbs are intrinsically necessary. The following sections will describe the procedures that were followed to isolate, purify, derivatize and quantify the glycans bound to mAbs produced through cell culture.

5.4.1. Sample purification

Prior to mass spectroscopic analysis, mAb-containing samples that were drawn from cell culture were purified using the Proteus Protein G antibody purification kit (PROCHEM, USA). This purification kit relies on the affinity of Staphylococcal protein G for the Fc of antibodies. The Proteus purification kit is small-scale and designed to be used in microcentrifuges, thus avoiding use of more sophisticated equipment such as liquid chromatographers. The purification was done according to the manufacturer's instructions. Non-reducing sodium dodecyl sulphate polyacrylamide (SDS-PAGE) gels were run after purification to assess removal of FBS-associated protein contaminants.

Figure 27 shows a typical SDS-PAGE gel obtained after protein G purification. In this figure, we see that the unpurified supernatant sample (lane 2) shows a large band at approximately 60kDa, which is believed to be bovine serum albumin (BSA). Standard HFN 7.1 mAb purchased from Thermo Scientific (UK) has been added to lane 3, and also shows a band that corresponds to the molecular weight of BSA. This is consistent with the specification of the purchased antibody standard, given that it is said to contain 0.2% BSA (Thermo Fisher, UK). Protein G purified samples were added to lanes 4–7. There we see faint bands at approximately 150kDa which correspond to mAb that was not fully reduced. Also in these lanes are bands between 55kDa-58kDa which correspond to the mAb heavy chain and bands at approximately 28kDa which correspond to mAb light chains. Figure 27 also shows that no other visible bands are present in the purified samples, leading us to believe that the samples were purified adequately with the protein G columns. Sample purity for glycan analysis is of extreme importance,

especially for cultures that have been supplemented with BSA. Glycoproteins, such as transferrin, may be present in BSA and, if not removed, may skew the glycan profiles measured with MS.

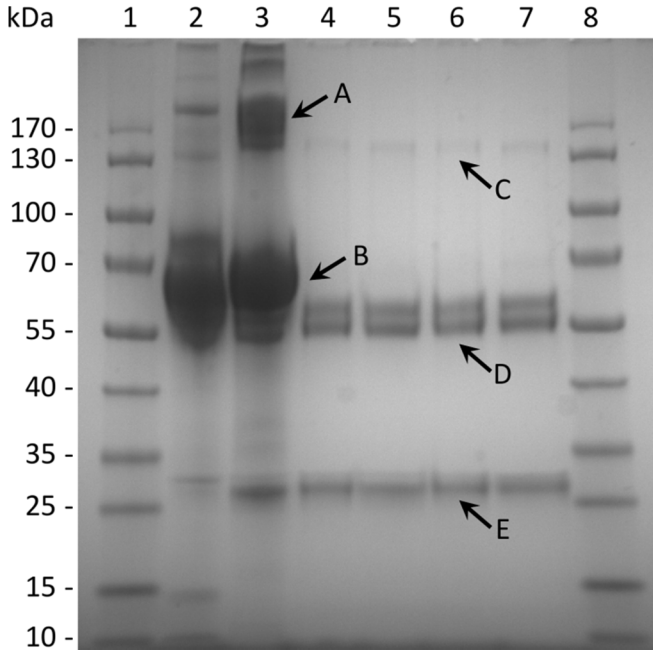


Figure 27. Reducing SDS-PAGE gel of purified mAbs
The lanes in the gel are: 170kDa ladder (1, 8), non-purified cell culture supernatant (2), HFN 7.1 mAb standard (3) and protein G purified cell culture supernatant (4–7). (A) shows the non-reduced HFN 7.1 mAb in the standard. (B) shows BSA that is present in both the non-purified supernatant and the standard. (C) shows residual intact mAb. (D) are bands corresponding to the heavy chain of the mAb. (E) are bands corresponding to the light chain of the mAb.

The mAb titre after purification was measured by absorbance at 280nm on a Nanodrop spectrophotometer (Thermo Scientific, UK), and ranged between 70mg/L and 120mg/L. Finally, the purified samples were freeze dried with a Thermo Savant ModulyoD freeze dryer (Thermo Scientific, USA). The purified and lyophilised samples were then prepared for MALDI mass spectrometry, the procedures for which are described below.

5.4.2. Buffers

The following solutions were made or purchased for MS sample preparation:

1. DTT: 2mg/mL dithioreitol (DTT, Sigma Aldrich) was prepared in 0.6M Tris (Amresco, USA), pH 8.5.
2. IAA: 12mg/mL iodoacetic acid (IAA, Sigma Aldrich) was prepared in 0.6M Tris, pH 8.5.
3. G-HCl: 8M guanidine HCl (Pierce, USA).
4. AMBIC: 50mM ammonium bicarbonate – (AMBIC, Fluka), pH 7.7.
5. AMBIC: 50mM ammonium bicarbonate, pH 8.4.

5.4.3. Reduction and carboxymethylation of mAbs

The purified and lyophilised sample was reduced and carboxymethylated to break all the disulphide bonds present in the mAb. The procedure was as follows:

1. 90µLof DTT and 10µL of G-HCl solution were added to the lyophilised sample.
2. The mixture was left to incubate for 1hr at 37°C.
3. 100µL of IAA solution was added
4. This mixture was incubated for 2hr at room temperature in absence of light.
5. The product of this reaction was dialysed against the AMBIC buffer for 24hr with regular buffer changes.
6. The samples were then lyophilised and stored.

5.4.4. Tryptic digestion of mAbs

In order to improve release of glycans from the reduced and carboxymethylated samples, they were treated with porcine sequencing-grade trypsin (Promega, USA). The procedure for tryptic digestion was as follows:

1. The samples from the step above were dissolved in 100µL of AMBIC buffer, pH 8.4.
2. 50µL of 0.02µg/µL trypsin solution were added.
3. This mixture was incubated for 14hr at 37°C.
4. The reaction products purified using Sep-Pak C_{18} cartridges (360mg 5-105µm) (Waters, UK).
5. The collected fractions from the Sep-Pak purification were concentrated in a Savant SPD122P Speed Vac (Thermo Scientific, UK).
6. All concentrated fractions were then pooled and lyophilised.

5.4.5. Peptide-N-glycosidase F (PNGase F) digestion

The purified and lyophilised tryptic digests were then treated with peptide-N-glycosydase F (PNGase-F – Roche Applied Science, UK) to release the N-linked glycans bound to them. The procedure for this step was the following:

1. The lyophilised samples from the previous step were resuspended in 200µL of AMBIC buffer, pH 8.4.
2. 5 units of PNGase-F were added to the resuspended tryptic digests.
3. The reaction mixture was incubated for 24hr at 37°C.
4. The samples were lyophilised and re-dissolved in 200µL of 5% acetic acid.

5. The digests were then purified using Sep-Pak C_{18} cartridges, using 1-propanol for elution.
6. The 1-propanol fractions were collected and concentrated with the Speed-Vac.
7. The concentrated samples were pooled and finally lyophilised.

5.4.6. Permethylation of released glycans

The released glycans were permethylated in order to ensure adequate ionisation during the MALDI analysis. Permethylation of the samples was done with the following protocol:

1. 5 NaOH pellets were ground with a mortar and pestle to produce a fine powder.
2. 3mL of anhydrous dimethylsulfoxide (DMSO – Romil, UK) were added to the NaOH powder to form a paste.
3. 1mL of the paste was added to the dry samples immediately before 0.5mL of methyl iodide (Alfa Aesar, UK) was added.
4. The mixture was vortexed and agitated for 15min at room temperature.
5. The reaction was quenched by dropwise addition of water.
6. The permethylated species were extracted using 1mL of chloroform (Romil, UK).
7. The chloroform was then dried with nitrogen.
8. The samples were resuspended in 50% aqueous methanol (MeOH – Romil, UK).
9. The re-dissolved permethylated glycans were purified using Sep-Pak C_{18} cartridges and eluting with aqueous acetonitrile (Romil, UK).
10. Fractions that eluted with 30% and 50% ACN were collected, concentrated with the Speed-Vac and subsequently lyophilised.

Prior to MALDI-MS, the lyophilised permethylated samples were co-crystallised with 2,5-dihydroxy benzoic acid (DHB, Sigma Aldrich). This was done by dissolving he dry, permethylated samples in 10μL of methanol (Romil, UK) and adding 1μL of 20mg/mL DHB in 70%v/v aqueous methanol. A 1μL sample was taken from the mixture and spotted onto a 100-well sample plate and dried under vacuum at room temperature.

5.4.7. MALDI MS data acquisition

MALDI-MS was performed on a Voyager-DE STR MALDI workstation (Perspective Biosystems, UK) equipped with delayed extraction technology. The machine was set in positive reflectron mode and data was acquired using the Voyager 5 Instrument Control Software. Data was processed using Data Explorer 4.9 software (Applied Biosystems, Germany). Using the software, data was baseline corrected and noise filtered (correction factor of 0.7). Obtained mass spectra

were then transferred to the GlycoWorkBench software[204] for processing. MALDI MS generally yields singly charged, sodiated molecular ions, $[M+Na]^+$. Knowing this, peaks on the spectra were assigned to glycan structures with the corresponding mass to charge ratio (m/z). GlycoWorkBench was also used to calculate the intensity of all assigned peaks and yield relative quantification of the observed glycans.

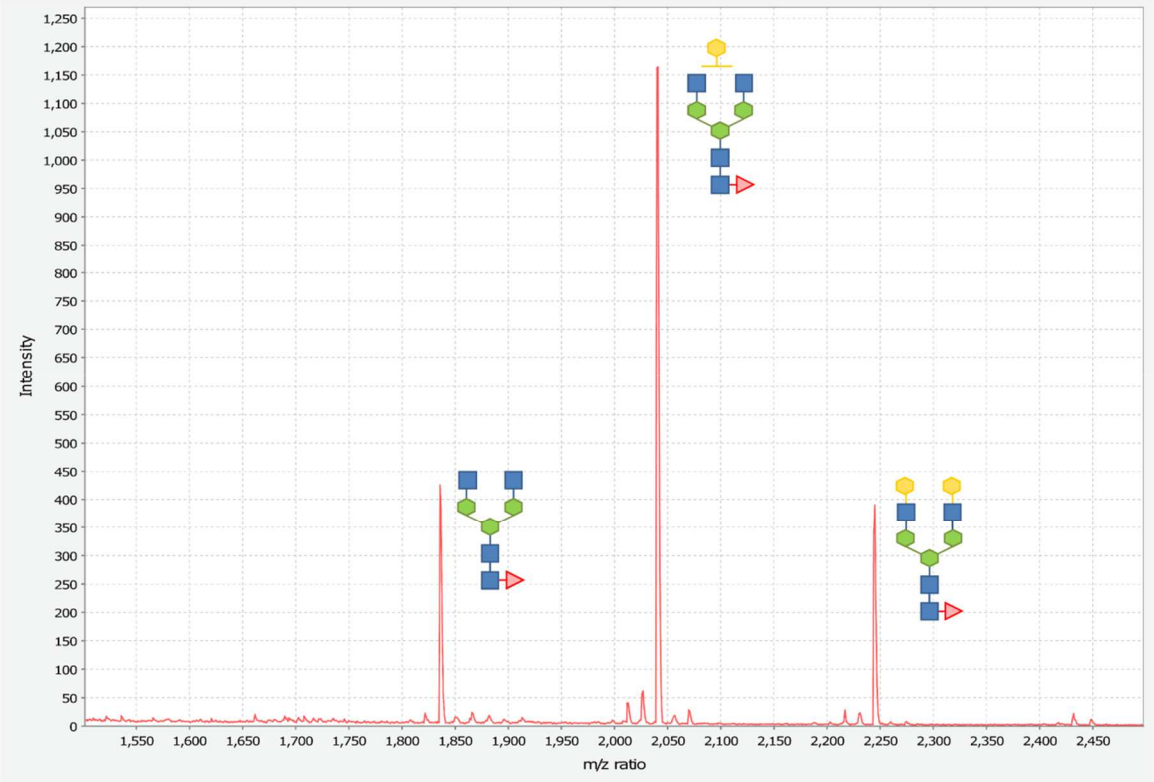


Figure 28. Typical mass spectrum of mAb-bound N-glycans
Three main peaks can be seen in the spectrum. Each peak corresponds to the glycan structure presented beside them.

Figure 28 shows a typical MALDI mass spectrum of permethylated N-linked glycans that have been obtained from mAbs that were produced in cell culture. Peaks have been assigned using GlycoWorkBench. The first peak has a m/z ratio of 1,835.77, corresponding to the molecular weight of a permethylated G0F structure as shown in the figure. The second peak has a m/z ratio of 2,039.80, which corresponds to a permethylated G1F glycan. The third peak has a m/z ratio of 2,243.90, which corresponds to a G2F glycan structure. As was reviewed in previous chapters of this thesis, these three glycan structures are the most commonly observed on the Fc of mAbs. These results, confirm that the methodologies used for MALDI-TOF MS of mAb Fc glycans is adequate and yields quantification of the glycan structures that are at the core of this thesis.

This chapter has described all the experimental methods that were developed and used to produce the data that will be presented in the following sections. These include methods that were used for mammalian cell culture, the experimental methods for determining the intracellular pool of NSDs throughout cell culture, and finally, the methods that were employed to determine the glycan profiles of mAbs that are produced from cell culture.

Chapter 6

Results and discussion

The previous chapters have defined the modelling and experimental framework that was developed to achieve our objective of producing a model-based tool that mechanistically describes mAb Fc glycosylation as a function of nutrient availability. This chapter will present a comparison of the experimental and modelling results and will also describe the different optimisation-based strategies that were employed to estimate unknown model parameters and discuss the obtained results.

6.1. Parameter estimation strategies

Both the glycosylation and cell culture dynamics models were developed independently. Because of this, different optimisation strategies have been defined to estimate their unknown parameters. The main experimental inputs for the Golgi glycosylation model are the specific mAb productivity and the Fc glycan profiles. Specific productivity of the mAb defines the Golgi resident protein concentration profiles through the optimisation strategies defined in sections 3.5.1 and 3.5.2 of this thesis. Once the resident protein concentration profiles are established, the glycan profiles are used to estimate the unknown dissociation constants of the glycosylation enzymes, as discussed in section 3.6. Similarly, estimation of the unknown parameters of the cell culture dynamics model has been split into two problems which must be solved sequentially. These two subsets of the cell culture dynamics and NSD model are cell culture dynamics (CCD) and intracellular processes (ICP).

CCD was defined in sections 4.1.1-4.1.3 and is the component that represents how cells grow and synthesise mAb as a function of nutrient availability and metabolite accumulation.

Estimation of the unknown parameters of the CCD component requires time-course data for *Xv*, *Xd*, [*Glcext*], [*Glnext*], [*Lact*], [*Amm*] and [*IgG*].

ICP includes the processes that occur intracellularly. The fundamental input of this subsystem is the transport of glucose and glutamine into the cell, as defined by equations 25-28. The unknown parameters of ICP are the transport rate coefficients for glucose and glutamine as well as the kinetic parameters contained in the reduced model for NSD biosynthesis. Estimation of these parameters requires intracellular data for glucose, glutamine and all NSDs. The flux leaving this system is consumption of NSDs for glycosylation reactions within the Golgi apparatus. Therefore, the flux through ICP is determined by mAb specific productivity as well.

6.2. Batch culture of HFN 7.1 cells

The parameter estimation strategies discussed in the previous section also indicate the experimental data required. In light of this, a series of batch overgrowth cultures were performed using the HFN 7.1 murine hybridoma cell line. Data for *Xv*, *Xd*, [*Glcext*], [*Glnext*], [*Lact*], [*Amm*], [*IgG*], intracellular glucose and NSDs was collected every 10 and 24 hours for the duration of the culture. The obtained data was then used to estimate the unknown parameters within the CCD, the ICP and the Golgi glycosylation models. In order to minimise error, four biological repeats of the batch experiment were done. The experimental data was collected and introduced into the parameter estimation function in gPROMS v. 3.4.0. gPROMS estimates parameters through the maximum likelihood formulation, which has been described in previous chapters of this thesis.

The estimated parameters for the CCD component of the culture dynamics model are presented in Table 18. The profiles calculated with the estimated parameters are shown in Figure 29. The experimental data are presented by points and the simulation results are shown with solid lines. Figure 29A shows that the simulation reproduces all phases of cell growth accurately. Dead cell density, on the other hand, is slightly underestimated towards the end of the culture. Considering that equation 13 defines accumulation of dead cells as a function of viable ones, it is expected for the decay of X_v to be symmetrical with the accumulation of X_d. Symmetry between both profiles is apparent in Figure 29A. The viability (Viab.) and the integral viable cell concentration (IVCC) are reported in Figure 29B. These CCD variables were not included in the parameter estimation. The figure shows that these variables are reproduced accurately by the model. Figure 29C shows mAb titre throughout culture. The model reproduces the data relatively well, although it must be mentioned that the standard deviations for this data are relatively large. This is likely due to the sensitivity of the ELISA protocol that was followed

to monitor mAb titre. However, though, the profile for mAb titre presented in Figure 29C is very close to previous reports for the same cell line.[294]

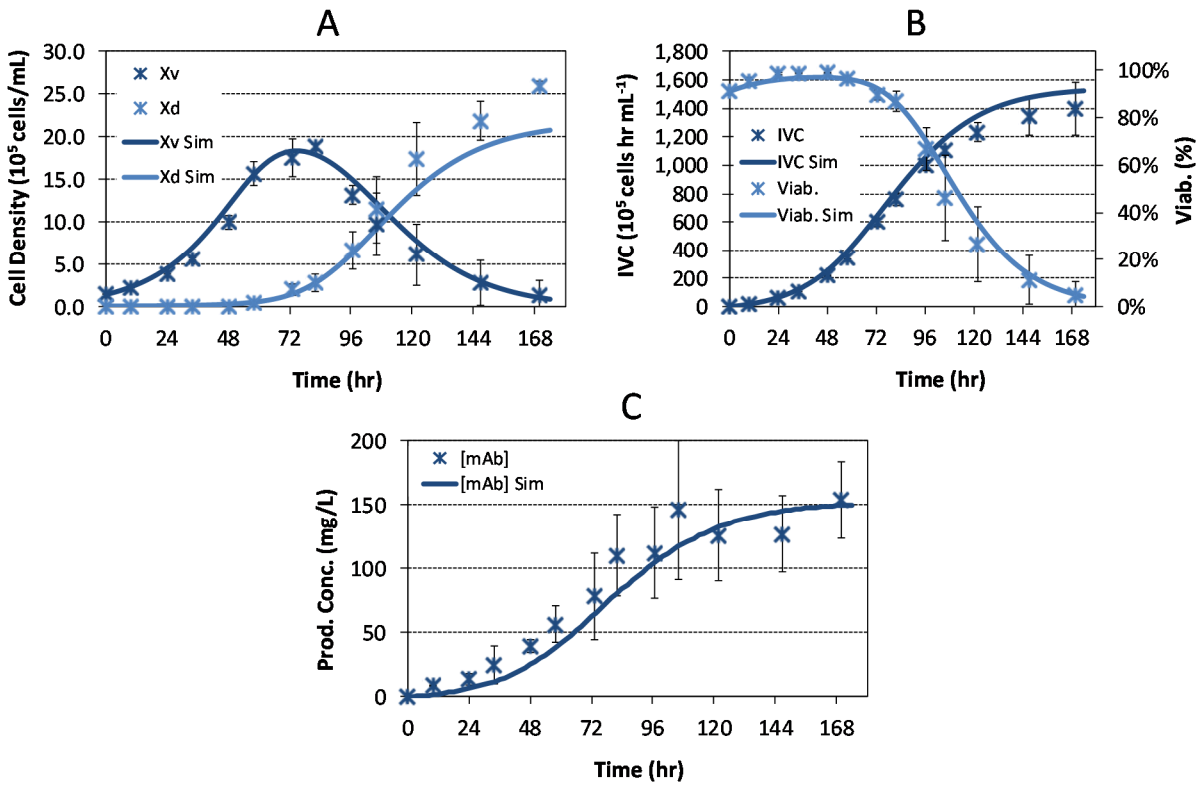


Figure 29. X_v, X_d, IVCC, viability and [mAb] profiles in batch culture
The points show the measured data and the lines show simulation results. Standard deviations for each point were calculated (n=4) for all data points as well.

Figure 30 shows the time profiles of nutrients and metabolites produced during the batch culture. The figure shows that the glucose and glutamine profiles are reproduced very well by the model. The behaviour of lactate is reproduced relatively well, although its accumulation is slightly underestimated during exponential growth. Towards the end of culture, the maximum lactate concentration is reproduced well. The behaviour for ammonia is the less well captured by the model. Figure 30D shows that the model produced a sigmoidal curve, whereas the experimental data shows saturation-type behaviour. This divergence leads to ammonia concentration being underestimated at early stages of the batch culture and slightly overestimated around 72hr. Glutamine has been reported to degrade to ammonia throughout cell culture.[295] This degradation would explain the profile of ammonia seen in Figure 30D. However, this would also impact the glutamine profile presented in panel C of the same figure.

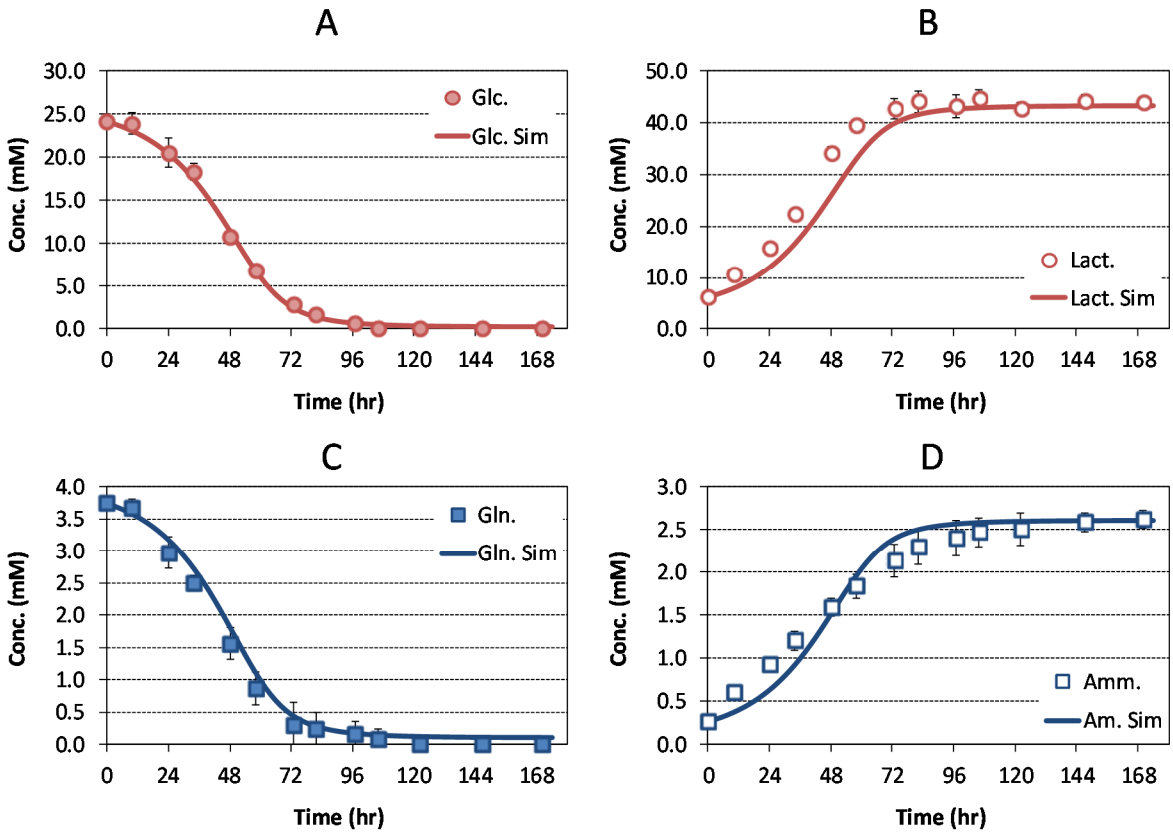


Figure 30. Extracellular glucose, lactate, glutamine and ammonia in batch overgrowth

Overall, it is possible to conclude that the experimental data has been reproduced accurately with the estimated parameters. Furthermore, the estimated parameters which are presented in Table 18 fall well within ranges that have been previously reported for murine hybridomas in culture.[272, 295, 296]

Table 18. Estimated parameters for culture dynamics model

Parameter	Final Value
μ_{max}	0.0758
$k_{d,max,glc}$	0.2341
$k_{d,max,gln}$	0.0190
$K_{M,glc}$	3.4043
$K_{M,gln}$	1.3459
$K_{d,glc}$	0.0615
$K_{d,gln}$	0.0011
$Y_{X/Glc}$	845.4
$Y_{X/Gln}$	5573
$Y_{Lact/Glc}$	1.5530
$Y_{Amm/Gln}$	0.6451
$Y_{mAb/Xv}$	98.538

Given the accurate representation of the dynamics of HFN 7.1 cells in culture, the model and its parameters were used to estimate the remaining unknown parameters associated with intracellular phenomena – namely, intracellular accumulation of NSDs.

6.2.1. Intracellular NSDs over batch culture

As was mentioned previously, samples for intracellular metabolites and were taken every 10 and 24 hours throughout culture. These samples were treated with the procedures described in sections 5.2 and 5.3 of this thesis. The intracellular extracts were then analysed with the HPAEC method presented in section 5.2.1 of this thesis. The intracellular NSD profiles that were obtained were then used for estimation of the unknown parameters present in equations 34-45. Again, parameter estimation was performed with gPROMS v. 3.4.0.

Table 19. Estimated parameters for NSD model

Parameter	Final value	Parameter	Final value
k_T^{Glc}	1.585E-06	K_M^{GDPMan}	68.76
k_T^{Gln}	2.544E-06	$K_M^{UDPGalNAc}$	2968
$k_{met,max}^{Glc}$	1248.95	$K_M^{CMPNeu5Ac}$	991.9
$k_{met,max}^{Gln}$	288.2	K_M^{UDPGal}	687.9
$K_{M,met}^{Glc}$	182.7	K_M^{GDPFuc}	794.5
$K_{M,met}^{Gln}$	9.485	$K_i^{CMPNeu5Ac}$	30.29
$k_f^{UDPGlcNAc}$	22.57	K_i^{GDPFuc}	29.43
k_f^{UDPGlc}	14.35	$K_{i,T}^{CMPNeu5Ac}$	125.2
k_f^{GDPMan}	26.75	$k_{Cons,max}^{UDPGalNAc}$	3022
$k_f^{UDPGalNAc}$	227.1	$K_{M,Cons}^{UDPGalNAc}$	1957
$k_f^{CMPNeu5Ac}$	59.17	$K_{UDPGlcNAc,glyc}$	90.82
k_f^{UDPGal}	17.65	$K_{UDPGlc,glyc}$	21.13
k_f^{GDPFuc}	2.864	$K_{GDPMan,glyc}$	159.4
$K_{M,glc}^{UDPGlcNAc}$	5.076	$K_{CMPNeu5Ac,glyc}$	332.0
$K_{M,gln}^{UDPGlcNAc}$	0.0288	$K_{UDPGal,glyc}$	39.84
K_M^{UDPGlc}	160.0	$K_{GDPFuc,glyc}$	416.9

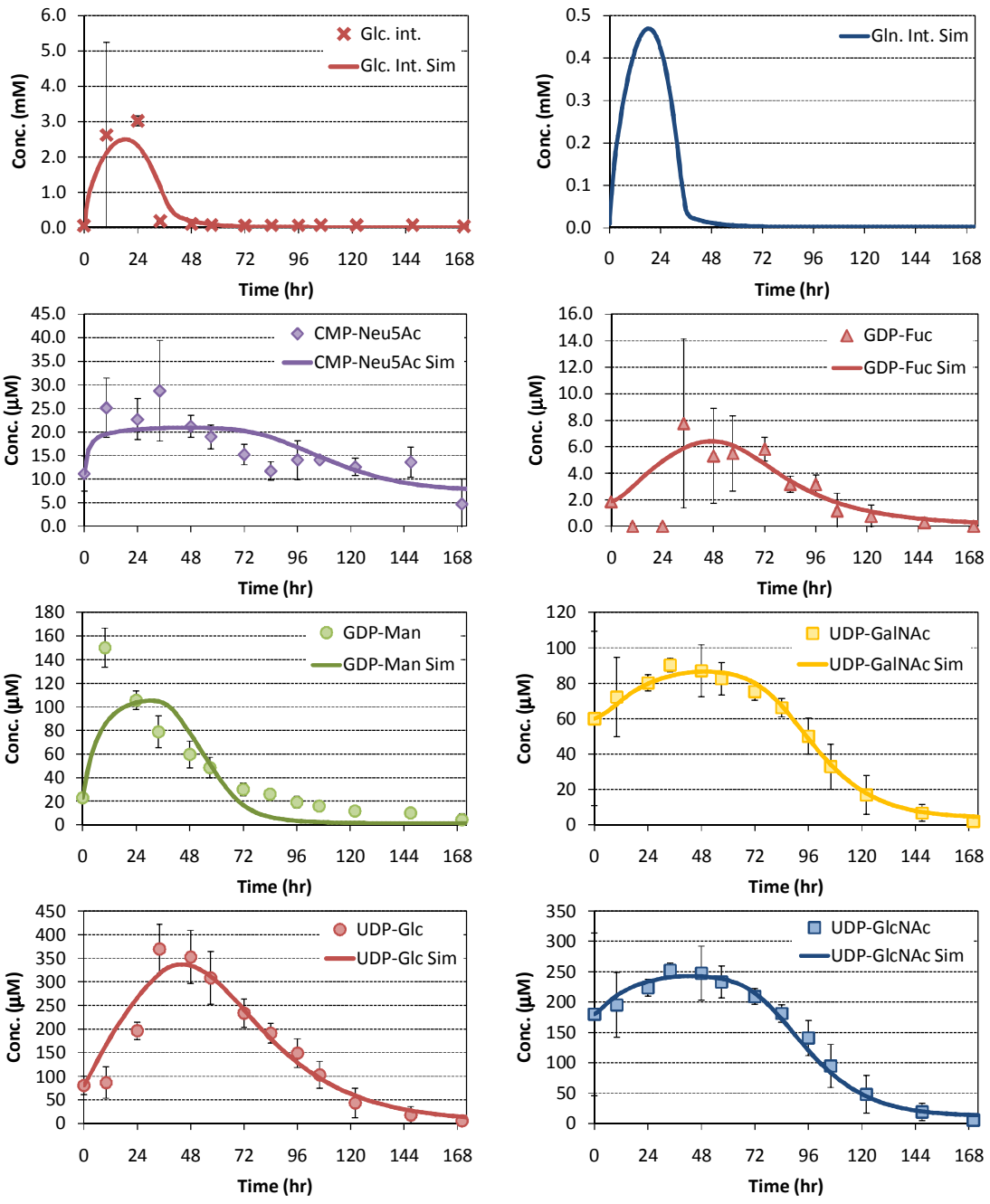


Figure 31. NSD profiles throughout batch culture
Experimental data is presented with points and model data is shown with solid lines. Again, the standard deviation for each point was calculated with n=4. Glucose is presented with red crosses, CMP-Neu5Ac is presented in purple, GDP-Fuc is shown in red triangles, GDP-Mannose is shown in green, UDP-GalNAc is shown in yellow boxes, UDP-Glc is shown with red circles and UDP-GlcNAc is described with blue squares.

Figure 31 shows all intracellular species involved in the parameter estimation. In this figure, we see the measured intracellular pool of glucose as measured with the Amplex Red assay kit. Very large measurement errors were observed at 10hr. All remaining time points, except the one for 24hrs, were below the limit of detection of the kit. Speculatively, intracellular glucose has a relatively high turnover because it is utilized as the main source of energy in mammalian

cells. The fact that it accumulates during initial stages of cell culture may be because its abundance with respect to number of cells in the culture is high. This could allow the relatively low number of cells present during the first 24hrs of culture to internalise it at rates that are higher than those at which it is consumed for metabolism. The intracellular glutamine concentration profile presented in Figure 31 was produced through simulations using estimated parameters. It can be seen that it follows the same trend as glucose. The reasons why this metabolite would peak at early stages of culture are similar to those behind initial glucose accumulation.

Figure 31 also shows that the NSDs whose synthesis is more closely related with essential nutrients tend to be the most prevalent ones (i.e. UDP-GlcNAc, UDP-Glc, and GDP-Man). Figure 31 shows that GDP-Mannose seems to spike early during the culture (first 10 hrs) and subsequently undergoes a fast decay. Another interesting feature of Figure 31 is that, despite the care that was taken with the NSD quantification protocol, GDP-Fuc is not detected at the first two time points, and also contains a very high standard deviation at 34hr. This is very likely due to two effects: the first is that GDP-Fuc has the lowest availability of all observed NSDs. This effect is compounded by the low cell densities observed at the first time points. Quantification of CMP-Neu5Ac also seems to yield fairly large variations, particularly between 10 and 48hr of culture. Again, this may be due to intrinsic low availability, but also due to the limited amount of viable cells in culture at those points.The solid lines in Figure 31 present the intracellular NSD concentrations calculated by the model. The estimated NSD parameters that were used for simulations are presented in Table 19.

Observing the modelling results presented in Figure 31, reproduction of the profiles for UDP-GlcNAc, UDP-GalNAc, UDP-Glc and GDP-Fuc are adequate. In contrast, the calculated profiles for CMP-Neu5Ac and GDP-Fuc show discrepancies. Instead of peaking at 10hrs, the calculated profile for GDP-Man accumulates more gradually during the first 24hrs. GDP-Man is also underestimated beyond 72hrs. The overall behaviour of CMP-Neu5Ac, seems to be reproduced relatively well. However, the model overestimates concentration of this NSD between 72 and 96hrs.

It is possible that adequate representation of GDP-Man is not achieved given certain assumptions underlying the lumped pathways. It is possible that the sharp accumulation of GDP-Man during the first 10hr of culture is associated with the mannose salvage pathway. The NSD model assumes that this NSD, along with the others are being transferred onto mAbs that are being secreted to the culture medium. However, the mannose salvage pathway has been reported to reduce net consumption of Mannose and its derivatives. It may be possible that at

early stages of culture, mannose is being recycled to form GDP-Man at rates that exceed consumption of this NSD in glycosylation reactions. REF: Direct utilization of mannose for mammalian glycoprotein biosynthesis, thus causing the sharp increase observed in the experimental data.

Additional data obtained from intracellular measurements is presented in Figure 32. This figure shows the measured intracellular availability of six nucleotides: ATP, ADP, AMP, CTP, GTP and UTP. In this figure we see that their profiles are quite consistent throughout cell culture and seem to be at concentrations that are comparable to the most abundant NSDs. ATP, however, has the highest concentration of all intracellular species quantified by HPAEC with a concentration in the range of mM.

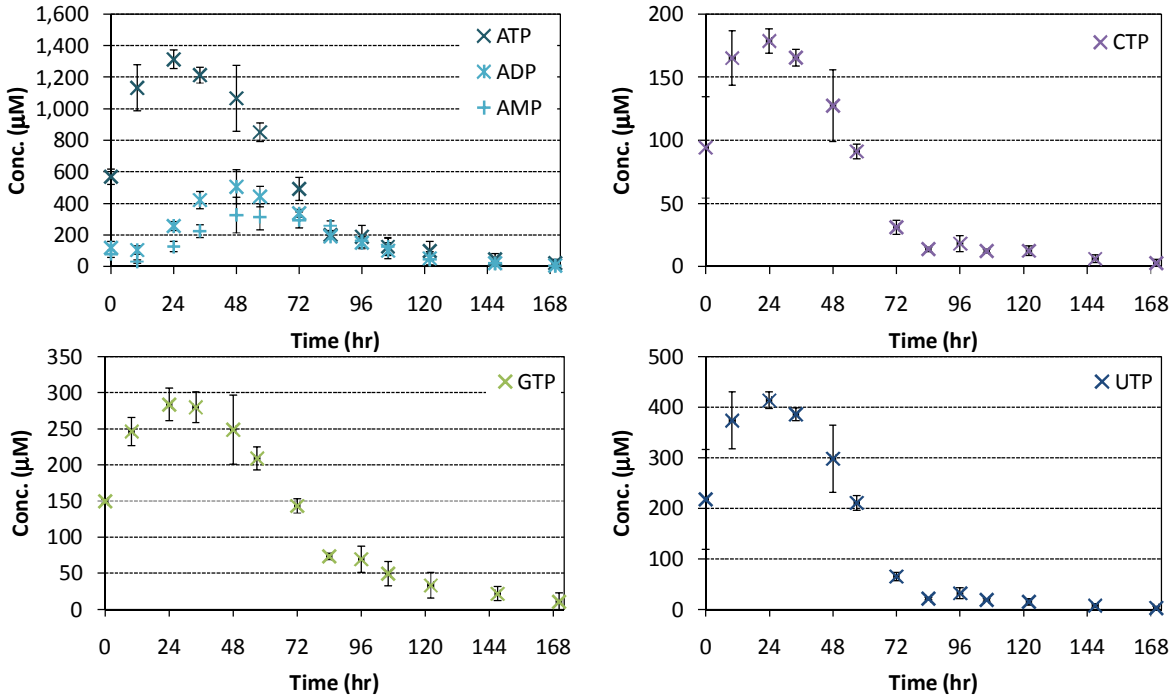


Figure 32. Intracellular nucleotide profiles throughout batch culture
Data for intracellular ATP, ADP, AMP, CTP, GTP and UTP are presented. The nucleotide data was produced with the HPAEC method described in previous chapters of this thesis.

Overall, and considering the number of simplifying assumptions that were made to describe NSD metabolism (e.g. reducing the reactions scheme from36 reactions to only 7), model performance is excellent. Apart from GDP-Man, the model has the ability of reproducing the experimental data rather well. Another conclusion that can be drawn from these results is that, with the exception of GDP-Fuc, the experimental methods to generate intracellular concentration profiles of NSDs are adequate.

6.2.2. mAb glycosylation profiles in batch culture

The remaining step to bridge the gap between extracellular nutrient availability is to estimate the unknown parameters of the Golgi glycosylation model using the obtained experimental data. Previously, it was discussed that the main experimental inputs into the Golgi glycosylation model are mAb specific productivity and the observed oligosaccharide profiles bound to the mAb. The specific productivity of mAb estimated from the CCD component is required to estimate resident protein profiles and the mAb Fc glycan profiles are used to estimate the glycosyltransferase dissociation constants so that the profile is matched.

Following this strategy, the specific mAb productivity estimated from previous steps in this procedure was used to calculate the Golgi enzyme profiles presented in Figure 33. As was described in chapter 3 of this thesis, the profiles presented in this figure represent the minimum amount of total enzyme required to fully process the glycan structures of a protein. In this case, these profiles are defined by the time it takes a glycoprotein to traverse the Golgi apparatus. This, in turn, is given by the specific mAb productivity.

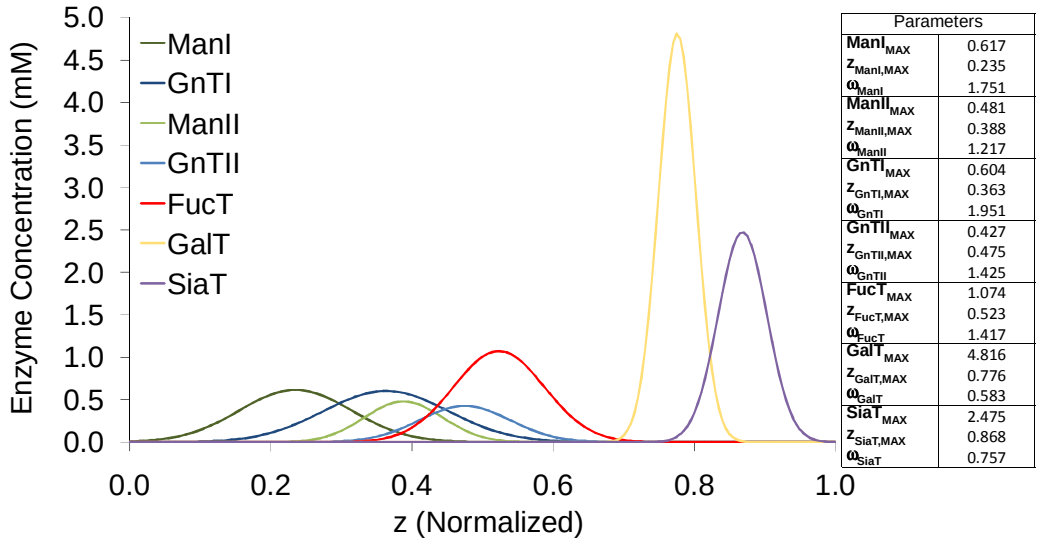


Figure 33. Estimated Golgi glycosylation enzyme profiles
The concentration profiles of ManI, GnTI, ManII, GnTII, FucT, GalT and SiaT along the length of te Golgi apparatus are presented. The parameters that were estimated and define these concentration profiles are presented on the right.

We must mention that when combining the NSD model and the Golgi glycosylation model, the NSD transport proteins are neglected. This assumption is consistent with the optimisation strategies discussed in section 3.5.2 of this thesis. The assumption underlying estimation of transport protein parameters is to minimise the difference between the rate of transport and the rate of consumption of NSDs in the Golgi glycosylation reactions. This may be interpreted as

saying that the rate-limiting step of NSD consumption in Golgi is not transport, and is mainly controlled by the glycosylation reactions. For this reason, the transport protein profiles were not estimated.

Once the enzyme profiles are defined, the measured glycan profiles are used to estimate the unknown dissociation constants of the glycosyltransferases, as discussed in chapter 3 of this thesis.

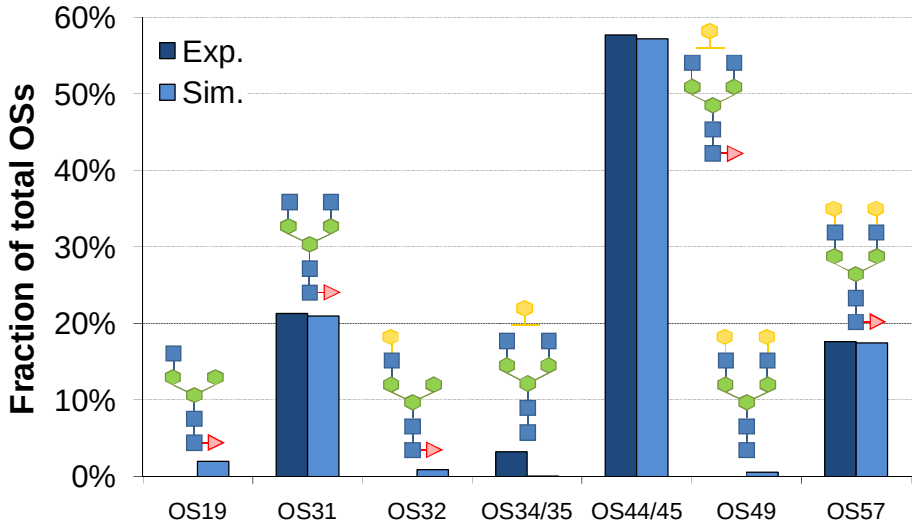


Figure 34. Comparison of measured mAb glycan profiles vs. simulation results
The dark blue bars represent the mAb glycan profiles that were obtained through MALDI MS. The light blue bars show the simulation results for the mAb glycan profile.

Figure 34. shows that the glycosylation model indeed reproduces the experimentally observed glycan profiles for the HFN 7.1 mAb. Table 20 presents the parameters that were estimated to produce the results shown in Figure 34.

Table 20. Estimated dissociation constants with measured glycan profiles

Dissociation Constant (μM)	HFN 7.1
$K_{d,i}^{GnTII}$	384.6
$K_{d,i}^{FucTA}$	7.37
$K_{d,i}^{FucTB}$	2.12
$K_{d,i}^{GalTa1A}$	168.0
$K_{d,i}^{GalTa1B}$	9.11
$K_{d,i}^{GalTa2A}$	52.0
Deviation from experimental data (%)	7.7

The bottom row of the table shows percent deviation of the simulation from the experimental data. The 7.7% deviation is due to underestimation of the OS34/35 species and overestimation of OS19 and OS32. It is likely the model underestimates synthesis of OS34/35 due to the activity of FucT. It can clearly be seen that the three main glycans present on the mAb are core fucosylated. Because these species are in higher abundance, the estimation yields parameters that favour their formation. In addition, the two species that are overestimated are ones where GlcNAc has not been added to the α-1,6 arm of the biantennary glycan. This is also likely related to the estimated FucT activity and to possible differential affinity of GnTII to fucosylated and non-fucosylated glycans. With additional experimental data, the cause of these divergences could be determined more thoroughly.

Through these results, it can be seen that all models defined in this work have the ability to accurately reproduce experimentally-observed cell culture dynamics, intracellular NSD concentration profiles and also the glycan profiles of a mAb. This section has also shown the high-quality experimental results that were obtained through the methodologies developed in this thesis.

6.3. Case study: glutamine-limited culture

Up to this point, the models have been presented as separate entities that have the ability of replicating experimental data. However, the overall objective of this project was to generate a single mathematical model that mechanistically describes the glycan profiles of mAbs as a function of nutrient availability. This section will present simulation results for a case where the initial concentration of glutamine is limited.

Two simulations were run with the coupled model. The parameters used for both simulations were the ones estimated in the previous section. The first simulation, which is noted as 'Base Case' was performed with an initial glutamine concentration of ~3.75mM. The second simulation was run with ~1.8mM initial glutamine concentration and is marked as 'Low Gln'.

Results for both simulations are compared in Figure 35 through 37. Figure 35 presents representative results for the cell culture dynamics element of the coupled model. In this figure, we see that viable cell density is clearly affected by the lack of glutamine. In the base case, X_v reaches close to 2e6cells/mL. Under glutamine starvation, X_v reaches approximately 1e6 cells/mL. Interestingly, though, while cells decline quickly after 120hrs in the base case, the decline under glutamine starvation is less pronounced. This is possibly due to the additive death rate expression that was defined in equation 15. In the base case, both glucose and glutamine

are depleted in 96hr yielding a more pronounced death phase. In the glutamine deprivation simulation, growth ceases when glutamine is depleted. When this occurs, glucose is no longer consumed for growth and remains at a constant concentration after 96hr. Because of this, glucose does not contribute to cell death through equation 15, and hence X_v decline is slower in the glutamine depletion case.

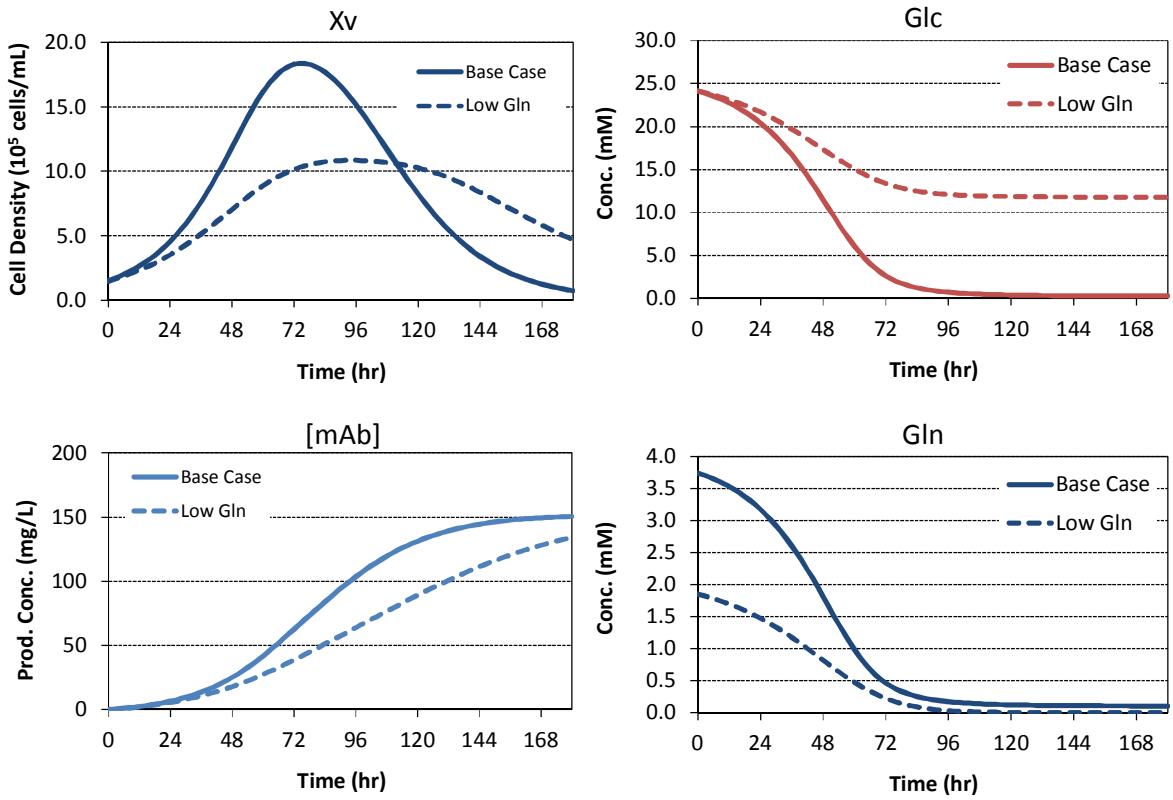


Figure 35. Cell culture dynamics under glutamine deprivation
The base case is represented by solid lines whereas the glutamine deficiency simulation is presented by broken lines.

Figure 36 compares the concentration profiles of NSDs under the base case and under glutamine deficiency. The effect of glutamine depletion is noticeable in the nitrogen containing NSDs. A clear depletion is observed for CMP-Neu5Ac, UDP-GalNAc and UDP-GlcNAc. The inverse occurs with the NSDs whose biosynthesis only depends on glucose. Given that glucose is not consumed for cell growth after 96 hours, more of it available for transport into cells where it is converted to UDP-Glc, GDP-Man and GDP-Fuc. Accordingly, these NSDs reach higher concentrations than those observed in the base case.

The final element to be observed is how glutamine depletion impacts mAb Fc glycosylation. Glycan profiles calculated from the base case are compared with those computed under

glutamine deprivation in Figure 37. On the left of this figure, we see that there is a clear decrease in biantennary oligosaccharide species in under glutamine starvation.

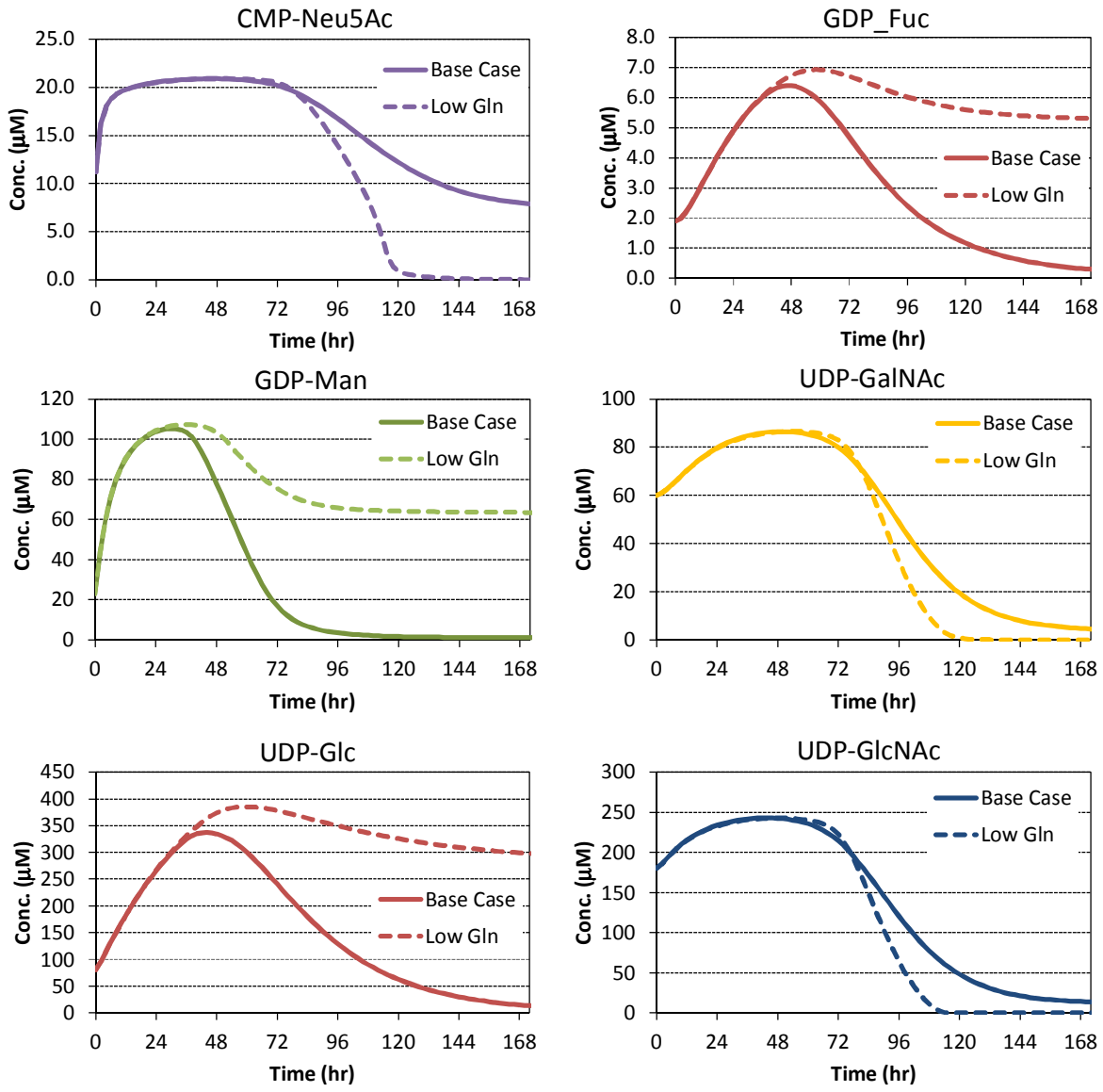


Figure 36. NSDs under glutamine deprivation
NSD concentration profiles simulated under normal conditions are shown with solid lines. Simulation results under glutamine depletion are presented with broken lines.

The right hand side of Figure 37 shows a very interesting result. In the base case, the fraction of G1F remains relatively unchanged throughout the cell culture simulation. In contrast, under gutamine deprivation, the fraction of this oligosaccharide drops sharply after 120hr, and while this occurs, there is a clear increase of Man5 oligosaccharide present on the mAb.

Accumulation of this particular glycan is of significant in the QbD context. mAbs with Man5 gycans have been reported to be immunogenic and have a decreased serum half-life[25, 92]

Observing Figure 36, the reason why Man5 accumulates is straight forward. The moment when Man5 begins to accumulate within the system corresponds directly with intracellular depletion of UDP-GlcNAc.

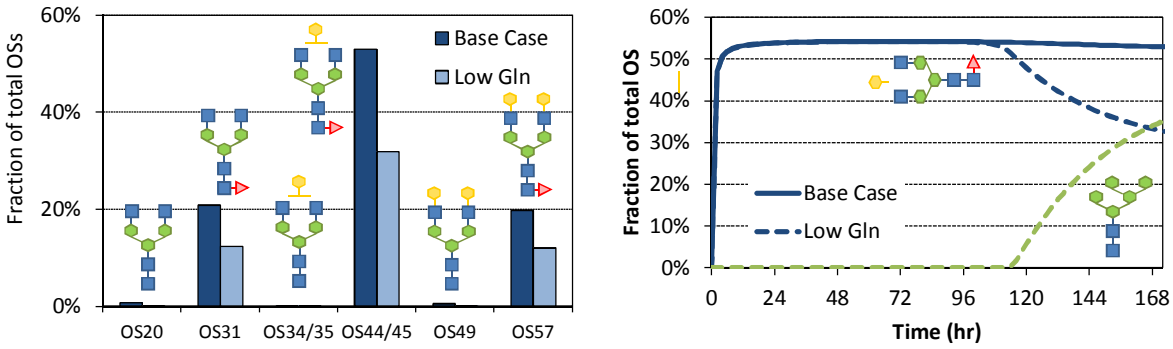


Figure 37. mAb glycan profiles under glutamine deprivation
On the left, the glycan profile calculated for the base case is shown with the dark bars. The light bars show the glycans calculated under glutamine depletion. On the right, a graph showing the accumulation of G1F throughout culture under the base case is shown in the solid blue line. Accumulation of the same glycan is shown with the broken blue line under glutamine deprivation. The green broken line shows accumulation of the Man5 glycan under glutamine starvation.

When no UDP-GlcNAc is available, the N-acetylgucosamine residue cannot be added to the α-1,3 arm of the oligosaccharide, and if this step does not occur, ManII cannot trim the excess mannose residues form the glycan.

With this particular example, the impact of nutrient limitation on glycosylation associated quality of mAbs is described in a fully mechanistic and quantitative manner. Moreover, this case-study highlights the possible applications of quantitative, mechanistic models for further bioprocess understanding under the QbD scope.

Chapter 7

Conclusions and future work

7.1. Summary of results

This thesis has presented the development of a combined modelling and experimental framework that aims to mechanistically and quantitatively describe the relationship between nutrient availability during cell culture and the N-linked glycosylation process.

The thesis first presented the current situation of the pharmaceutical and biopharmaceutical industry. It continued by describing the Quality by Design paradigm, along with its convenience for development and manufacture of biopharmaceutical products. This was followed by a review of the relevant literature concerning N-linked glycosylation, the impact bioprocess conditions have on this post translation modification, as well as the impact it has on the therapeutic properties of biological drugs. A brief review was also made on the analytical techniques that are used to characterise the N-glycan profiles of protein therapeutics.

Subsequently, the formulation of a dynamic mathematical model that describes the intracellular process of N-linked glycosylation was presented. Details surrounding the formulation of this model along with additional optimisation-based strategies that were developed to estimate its unknown parameters were presented. The validity of the Golgi glycosylation model was further demonstrated by comparing its performance with previous models of this phenomenon. This Golgi glycosylation model was also shown to adequately reproduce the impact a complex perturbation, such as gene silencing, has on N-linked glycosylation.

This was followed by the description of a second mathematical model that was developed to represent cell culture dynamics and the biosynthetic pathway of nucleotide sugars. This basis of

this model was to capture as many mechanistic features of the NSD pathway as possible, while reducing its size and unknown parameters.

Subsequent sections detailed the experimental methods that were employed throughout this work, including a chromatographic technique that was optimised to separate and quantify intracellular nucleotide sugar pools in under 30 minutes. Details regarding the mass spectrometric techniques that were employed to analyse the N-glycosylation of a mAb were also presented.

The following sections presented the experimental and modelling results that were obtained throughout this work. This section mainly detailed the parameter estimation strategies that coupled the modelling work with the experimental results. Finally, a case study describing how a combined model which incorporates cell culture dynamics, intracellular NSD metabolism and mAb N-linked glycosylation was presented. The results shown by this case study highlight the purpose of this work by mechanistically and quantitatively describing how nutrient depletion during cell culture impacts the N-glycosylation quality of mAbs.

7.2. Main conclusions

The following conclusions can be drawn from this work:

1. Synergy between pharmaceutical companies and regulatory agencies is essential to ensure that the highest quality drugs are produced without additional expense for healthcare providers and patients. Quality by Design is a platform that ensures end product quality for patients while incentivising large pharmaceutical companies to introduce innovation in their products and manufacturing processes.

2. Under the QbD scope, N-linked glycosylation of therapeutic proteins is key for their safety and therapeutic efficacy. Many of the mechanisms by which the N-glycosylation process is affected by process inputs are still unknown. Much work has to be done to understand the fundamental mechanisms that impact this process.

3. Quantitative mechanistic models will play a critical role in the implementation of QbD to development, manufacture and approval of biopharmaceuticals.

4. The framework for modelling Golgi N-linked glycosylation that was presented in this work is robust and has immediate applicability for guiding genetic engineering strategies to knock out or overexpress N-glycosylation genes. This is highlighted by the accurate representation of FucT silencing it achieved.

5. Strategies for model reduction are critical in the field of systems biology. However, it is important to obtain a balanced model that does not compromise mechanistic fidelity for simplicity. Models which are too complex include several unknown mechanisms and pathways, which if left disregarded, will negatively impact model accuracy.

6. Combination of mathematical modelling with experimentation is crucial for accurate representation of complex phenomena. This was clear from all the estimation strategies that were implemented to extract kinetic parameters from extracellular and intracellular data.

7. The final product of this thesis is a mathematical model that mechanistically links nutrient availability during cell culture with N-linked glycosylation. When further validated, this model can be used for process development, design and optimisation. Because it attempts to capture as many mechanistic features of the underlying physical phenomena it describes, it is congruent with QbD principles. Again, with further validation, models like this will play a critical role in biopharmaceutical process development, control and optimisation.

7.3. Future work

Based on the results and conclusions of this thesis, the following are suggestions for future work that can be done to as a continuation of this work:

7.3.1. Golgi glycosylation model improvement

The model for Golgi glycosylation that was described in this thesis requires experimental validation to some of its assumptions. As was discussed previously, the concentration profiles of resident proteins within the Golgi are estimates, and would need to be confirmed or complemented by experimental work. In addition, these resident protein profiles are fitted to three parameter normal functions. Ideally, the mechanisms by which these proteins are recycled by COPI coated vesicles would be included in the model. That way, deformation of the profiles due to excessive secretory traffic through Golgi would be represented. Mechanistic descriptions of this nature would certainly contribute to model fidelity, and allow for an improved description of the impact of process conditions on the Golgi glycosylation process. AN example of this would be temperature or pH, as was discussed in the body of this thesis. If the mechanisms for Golgi enzyme recycling are well established, the impact of temperature on glycosylation could be described mechanistically.

7.3.2. NSD model improvement

The model for NSD metabolism that was presented in this thesis is rather simplified. It is possible that certain intermediate species along the pathway have heavy regulatory effects. The simplest extension to the model would be to consider availability of nucleotides within the kinetics. The HPAEC method that was optimised in this work allows for accurate quantification of these compounds. Furthermore, their intracellular concentrations seem to be in the same order of magnitude as the most prevalent NSDs. It is possible that nucleotide availability indeed impacts the biosynthetic pathway. Given that they can be readily monitored through HPAEC, including them in the model would be relatively simple.

7.3.3. NSD quantification

It is possible that extraction with PCA is not ideal for NSD quantification. However, the caveats of ACN extraction were also discussed. Further work on methods for quenching, washing and extracting NSDs would lead to improved datasets. This, in turn, would lead to higher confidence in estimated parameters within the models.

7.3.4. Intracellular glucose and glutamine determination

The methods that were used in this thesis to monitor intracellular glucose and glutamine were not adequate. Although the Amplex Red assay kit should be sensitive enough to produce full time profiles of intracellular glucose, very few points were measured. The case of glutamine is even more interesting. Because glutamine serves as the nitrogen source for the NSD pathway, understanding how its intracellular concentration varies with respect to time is desirable. Furthermore, analytical methods to measure both of these metabolites are preferable over kits because the risk of human error is reduced. Full data sets for intracellular glucose and glutamine would contribute to further understanding of the dynamics behind NSD biosynthesis.

7.3.5. High throughput glycan analysis

The mathematical models defined in this thesis are dynamic. As such, time courses for glycan profiles would contribute to improved model accuracy. An example would be case study. If the calculations were to be confirmed experimentally, the time course data for glycan profiles would be necessary.

7.3.6. Model validation

Experimental data was used to estimate unknown parameters of the system. These parameters, though, have not yet been validated by an independent set of data. Validation experiments should lead to improved confidence in the parameter values themselves but would also shed light into further mechanistic effects that occur within the system.

List of publications

1. Jedrzejewski PM, del Val IJ, Polizzi KM, and Kontoravdi C, Applying Quality by Design (QbD) to Glycoprotein Therapeutics: Experimental and Computational Efforts towards Glycoform Control. *Pharmaceutical Bioprocessing* 2013 *accepted*.

2. del Val IJ, Jedrzejewski PM, Exley K, Sou SN, Kyriakopoulos S, Polizzi KM, and Kontoravdi C: Application of Quality by Design Paradigm to the Manufacture of Protein Therapeutics, in *Glycosylation*, S. Petrescu, (Ed.), InTech, 2012.

3. Chen N, del Val IJ, Kyriakopoulos S, Polizzi KM, Kontoravdi C: Metabolic network reconstruction: advances in *in silico* interpretation of analytical information. *Current Opinion in Biotechnology* 2012, 23(1):77-82.

4. del Val IJ, Nagy JM, Kontoravdi C: A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnology Progress* 2011, 27(6):1730-1743.

5. del Val IJ, Kontoravdi C, Nagy JM: Towards the implementation of quality by design to the production of therapeutic monoclonal antibodies with desired glycosylation patterns. *Biotechnology Progress* 2010, 26:1505-1527.

References

1. Elvin, J.G., Couston, R.G. & van der Walle, C.F. Therapeutic antibodies: Market considerations, disease targets and bioprocessing. *International journal of pharmaceutics* (2011).

2. World Preview 2018: Embracing the Patent Cliff. *EvaluatePharma Report* (2012).

3. Reichert, J.M. Which are the antibodies to watch in 2012? *Mabs-Austin* **4**, 1-3 (2012).

4. Beck, A., Wurch, T., Bailly, C. & Corvaia, N. Strategies and challenges for the next generation of therapeutic antibodies. *Nat Rev Immunol* **10**, 345-352 (2010).

5. Walsh, G. Biopharmaceutical benchmarks 2010. *Nat Biotechnol* **28**, 917-924 (2010).

6. Reichert, J.M. & Dhimolea, E. The future of antibodies as cancer drugs. *Drug discovery today* (2012).

7. Pollock, A. Transforming the critique of Big Pharma. *Biosocieties* **6**, 106-118 (2011).

8. Paul, S.M. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature reviews. Drug discovery* **9**, 203-214 (2010).

9. Kaitin, K.I. Deconstructing the drug development process: the new face of innovation. *Clinical Pharmacology and Therapeutics* **87**, 356-361 (2010).

10. DiMasi, J.A., Hansen, R.W. & Grabowski, H.G. The price of innovation: new estimates of drug development costs. *J Health Econ* **22**, 151-185 (2003).

11. Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C. & Greyson, D. The cost of drug development: A systematic review. *Health Policy* **100**, 4-17 (2011).

12. Adams, C.P. & Brantner, V.V. Spending on New Drug Development. *Health Econ* **19**, 130-141 (2010).

13. Light, D.W. & Warburton, R. Demythologizing the high costs of pharmaceutical research. *Biosocieties* **6**, 34-50 (2011).

14. DiMasi, J.A., Grabowski, H.G. & Vernon, J. R&D costs and returns by therapeutic category. *Drug Inf J* **38**, 211-223 (2004).

15. Adams, C.P. & Brantner, V.V. Estimating the cost of new drug development: Is it really $802 million? *Health Affair* **25**, 420-428 (2006).

16. Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products. *USFDA Report* (2004).

17. Huang, J. et al. Quality by design case study: An integrated multivariate approach to drug product and process development. *International journal of pharmaceutics* **382**, 23-32 (2009).

18. Rathore, A.S., Saleki-Gerhardt, A., Montgomery, S.H. & Tyler, S.M. Quality by Design: Industrial Case Studies on Defining and Implementing Design Space for Pharmaceutical Processes Part 1. *Biopharm Int* **21**, 37-+ (2008).

19. Rathore, A.S., Saleki-Gerhardt, A., Montgomery, S.H. & Tyler, S.M. Quality by Design: Industrial Case Studies on Defining and Implementing Design Space for Pharmaceutical Processes Part 2. *Biopharm Int* **22**, 36+ (2009).

20. Xu, X.M., Khan, M.A. & Burgess, D.J. A quality by design (QbD) case study on liposomes containing hydrophilic API: I. Formulation, processing design and risk assessment. *International journal of pharmaceutics* **419**, 52-59 (2011).

21. Xu, X.M., Khan, M.A. & Burgess, D.J. A quality by design (QbD) case study on liposomes containing hydrophilic API: II. Screening of critical variables, and establishment of design space at laboratory scale. *International journal of pharmaceutics* **423**, 543-553 (2012).

22. Harris, L.J., Skaletsky, E. & McPherson, A. Crystallographic structure of an intact IgG1 monoclonal antibody. *J Mol Biol* **275**, 861-872 (1998).

23. Jimenez del Val, I. et al. in Application of Quality by Design Paradigm to the Manufacture of Protein Therapeutics. (ed. S. Petrescu) (InTech, 2012).

24. Walsh, G. Post-translational modifications of protein biopharmaceuticals. *Drug discovery today* **15**, 773-780 (2010).

25. Wright, A. & Morrison, S.L. Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol* **15**, 26-32 (1997).

26. Krapp, S., Mimura, Y., Jefferis, R., Huber, R. & Sondermann, P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* **325**, 979-989 (2003).

27. Raju, T.S. Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Curr Opin Immunol* **20**, 471-478 (2008).

28. Hossler, P., Khattak, S.F. & Li, Z.J. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* **19**, 936-949 (2009).

29. del Val, I.J., Kontoravdi, C. & Nagy, J.M. Towards the Implementation of Quality by Design to the Production of Therapeutic Monoclonal Antibodies with Desired Glycosylation Patterns. *Biotechnol Progr* **26**, 1505-1527 (2010).

30. Quality by Design in Biomanufacturing. *Bioproduction Group Whitepaper* (2011).

31. Umaña, P. & Bailey, J.E. A mathematical model of N-linked glycoform biosynthesis. *Biotechnol Bioeng* **55**, 890-908 (1997).

32. Krambeck, F.J. & Betenbaugh, M.J. A mathematical model of N-linked glycosylation. *Biotechnol Bioeng* **92**, 711-728 (2005).

33. Hossler, P., Mulukutla, B.C. & Hu, W.-S. Systems analysis of N-glycan processing in mammalian cells. *PLoS ONE* **2**, e713 (2007).

34. Yu, L.X. Pharmaceutical quality by design: Product and process development, understanding, and control. *Pharmaceut Res* **25**, 781-791 (2008).

35. Hinz, D.C. Process analytical technologies in the pharmaceutical industry: the FDA's PAT initiative. *Anal Bioanal Chem* **384**, 1036-1042 (2006).

36. Kenett, R.S. & Kenett, D.A. Quality by Design applications in biosimilar pharmaceutical products. *Accredit Qual Assur* **13**, 681-690 (2008).

37. ICH (International Conference on Harmonisation, 2009).

38. Rathore, A.S. & Winkle, H. Quality by design for biopharmaceuticals. *Nat Biotechnol* **27**, 26-34 (2009).

39. Reichert, J.M. Trends in development and approval times for new therapeutics in the United States. *Nature reviews. Drug discovery* **2**, 695-702 (2003).

40. DiMasi, J. & Grabowski, H. The Cost of Biopharmaceutical R&D: Is Biotech Different? *Manage Decis Econ* **28** (2007).

41. Kenett, R. & Kenett, D. Quality by Design applications in biosimilar pharmaceutical products. *Accreditation and Quality Assurance* **13**, 681-690 (2008).

42. USFDA (US Food and Drug Administration, Rockville, MD; 2004).

43. ICH (International Conference on Harmonisation, 2005).

44. ICH (International Conference on Harmonisation, 2008).

45. Stadlmann, J., Pabst, M., Kolarich, D., Kunert, R. & Altmann, F. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* **8**, 2858-2871 (2008).

46. Vogt, F.G. & Kord, A.S. Development of Quality-By-Design Analytical Methods. *J Pharm Sci-Us* **100**, 797-812 (2011).

47. Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-497 (1975).

48. Mullard, A. Can next-generation antibodies offset biosimilar competition? *Nature reviews. Drug discovery* **11**, 426-428 (2012).

49. Stern, M. & Herrmann, R. Overview of monoclonal antibodies in cancer therapy: present and promise. *Crit Rev Oncol Hematol* **54**, 11-29 (2005).

50. Raju, T.S. Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Current Opinion in Immunology* **20**, 471-478 (2008).

51. Jefferis, R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nature reviews. Drug discovery* **8**, 226-234 (2009).

52. Gaetano, N.D. et al. Complement activation determines the therapeutic activity of rituximab in vivo. *J Immunol* **171**, 1581-1587 (2003).

53. Gelderman, K.A., Tomlinson, S., Ross, G.D. & Gorter, A. Complement function in mAb-mediated cancer immunotherapy. *Trends Immunol* **25**, 158-164 (2004).

54. Kaneko, Y., Nimmerjahn, F. & Ravetch, J.V. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313**, 670-673 (2006).

55. Anthony, R.M. et al. Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science* **320**, 373-376 (2008).

56. Nimmerjahn, F. & Ravetch, J.V. Anti-inflammatory actions of intravenous immunoglobulin. *Annu Rev Immunol* **26**, 513-533 (2008).

57. Janeway, C.A., Travers, P., Walport, M. & J., S.M. Immunobiology: The Immune System in Health and Disease, Edn. 6th ed. (Garland Science, New York; 2005).

58. Hirayama, K. et al. Complete and rapid peptide and glycopeptide mapping of mouse monoclonal antibody by LC/MS/MS using ion trap mass spectrometry. *Anal Chem* **70**, 2718-2725 (1998).

59. Leibiger, H., Wüstner, D., Stigler, R.D. & Marx, U. Variable domain-linked oligosaccharides of a human monoclonal IgG: structure and influence on antigen binding. *Biochem J* **338 (Pt 2)**, 529-538 (1999).

60. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* **54**, 631-664 (1985).

61. Cumming, D.A. Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology* **1**, 115-130 (1991).

62. Rabouille, C. et al. Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. *J Cell Sci* **108 (Pt 4)**, 1617-1627 (1995).

63. de Graffenried, C.L. & Bertozzi, C.R. The roles of enzyme localisation and complex formation in glycan assembly within the Golgi apparatus. *Curr Opin Cell Biol* **16**, 356-363 (2004).

64. Varki, A. Essentials of glycobiology, Edn. 2nd. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; 2009).

65. Taylor, M.E. & Drickamer, K. Introduction to glycobiology, Edn. 3rd. (Oxford University Press, Oxford; 2011).

66. Murrell, M.F., Yarema, K.J. & Levchenko, A. The systems biology of glycosylation. *Chembiochem* **5**, 1334-1347 (2004).

67. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**, D109-D114 (2012).

68. Wallick, S.C., Kabat, E.A. & Morrison, S.L. Glycosylation of a VH residue of a monoclonal antibody against alpha (1----6) dextran increases its affinity for antigen. *J Exp Med* **168**, 1099-1109 (1988).

69. Tandai, M. et al. Structural study of the sugar moieties of monoclonal antibodies secreted by human-mouse hybridoma. *Arch Biochem Biophys* **291**, 339-348 (1991).

70. Wright, A., Tao, M.H., Kabat, E.A. & Morrison, S.L. Antibody variable region glycosylation: position effects on antigen binding and carbohydrate structure. *EMBO J* **10**, 2717-2723 (1991).

71. Endo, T., Wright, A., Morrison, S.L. & Kobata, A. Glycosylation of the variable region of immunoglobulin G--site specific maturation of the sugar chains. *Mol Immunol* **32**, 931-940 (1995).

72. Jefferis, R. Glycosylation of recombinant antibody therapeutics. *Biotechnol Prog* **21**, 11-16 (2005).

73. Holland, M. et al. Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim Biophys Acta* **1760**, 669-677 (2006).

74. Jefferis, R. et al. A comparative study of the N-linked oligosaccharide structures of human IgG subclass proteins. *Biochem J* **268**, 529-537 (1990).

75. Mimura, Y., Ashton, P.R., Takahashi, N., Harvey, D.J. & Jefferis, R. Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. *J Immunol Methods* **326**, 116-126 (2007).

76. Kobata, A. The N-linked sugar chains of human immunoglobulin G: their unique pattern, and their functional roles. *Biochim Biophys Acta* **1780**, 472-478 (2008).

77. Wormald, M.R. et al. Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* **36**, 1370-1380 (1997).

78. Hills, A.E., Patel, A., Boyd, P. & James, D.C. Metabolic control of recombinant monoclonal antibody N-glycosylation in GS-NS0 cells. *Biotechnol Bioeng* **75**, 239-251 (2001).

79. Lifely, M.R., Hale, C., Boyce, S., Keen, M.J. & Phillips, J. Glycosylation and biological activity of CAMPATH-1H expressed in different cell lines and grown under different culture conditions. *Glycobiology* **5**, 813-822 (1995).

80. Shields, R.L. et al. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc gamma RIII and antibody-dependent cellular toxicity. *J Biol Chem* **277**, 26733-26740 (2002).

81. Shinkawa, T. et al. 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. *J Biol Chem* **278**, 3466-3473 (2003).

82. Matsumiya, S. et al. Structural Comparison of Fucosylated and Nonfucosylated Fc Fragments of Human Immunoglobulin G1. *J Mol Biol* **368**, 767-779 (2007).

83. Umaña, P., Jean-Mairet, J., Moudry, R., Amstutz, H. & Bailey, J.E. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat Biotechnol* **17**, 176-180 (1999).

84. Ferrara, C. et al. Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. *Biotechnol Bioeng* **93**, 851-861 (2006).

85. Mori, K. et al. Engineering Chinese hamster ovary cells to maximize effector function of produced antibodies using FUT8 siRNA. *Biotechnol Bioeng* **88**, 901-908 (2004).

86. Imai-Nishiya, H. et al. Double knockdown of alpha1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC. *BMC Biotechnol* **7**, 84 (2007).

87. Yamane-Ohnuki, N. et al. Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity. *Biotechnol Bioeng* **87**, 614-622 (2004).

88. Kanda, Y. et al. Comparison of cell lines for stable production of fucose-negative antibodies with enhanced ADCC. *Biotechnol Bioeng* **94**, 680-688 (2006).

89. Alley, S.C. et al. in Proceedings of the American Association for Cancer Research Annual Meeting2010).

90. Siadak, T. et al. in Cell Culture Engineering XI2008).

91. Wright, A. & Morrison, S.L. Effect of altered CH2-associated carbohydrate structure on the functional properties and in vivo fate of chimeric mouse-human immunoglobulin G1. *J Exp Med* **180**, 1087-1096 (1994).

92. Dwek, R.A., Lellouch, A.C. & Wormald, M.R. Glycobiology: 'the function of sugar in the IgG molecule'. *J Anat* **187 (Pt 2)**, 279-292 (1995).

93. Alessandri, L. et al. Increased serum clearance of oligomannose species present on a human IgG1 molecule. *Mabs-Austin* **4**, 509-520 (2012).

94. Mimura, Y. et al. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* **37**, 697-706 (2000).

95. Boyd, P.N., Lines, A.C. & Patel, A.K. The effect of the removal of sialic acid, galactose and total carbohydrate on the functional activity of Campath-1H. *Molecular Immunology* **32**, 1311-1318 (1995).

96. Malhotra, R. et al. Glycosylation Changes of Igg Associated with Rheumatoid-Arthritis Can Activate Complement Via the Mannose-Binding Protein. *Nat Med* **1**, 237-243 (1995).

97. Arnold, J.N., Dwek, R.A., Rudd, P.M. & Sim, R.B. Mannan binding lectin and its interaction with immunoglobulins in health and in disease. *Immunol Lett* **106**, 103-110 (2006).

98. Hodoniczky, J., Zheng, Y.Z. & James, D.C. Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. *Biotechnol Prog* **21**, 1644-1652 (2005).

99. Weikert, S. et al. Engineering Chinese hamster ovary cells to maximize sialic acid content of recombinant glycoproteins. *Nat Biotechnol* **17**, 1116-1121 (1999).

100. Bragonzi, A. et al. A new Chinese hamster ovary cell line expressing alpha2,6-sialyltransferase used as universal host for the production of human-like sialylated recombinant glycoproteins. *Biochim Biophys Acta* **1474**, 273-282 (2000).

101. Hamilton, S.R. et al. Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313**, 1441-1443 (2006).

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

103. Goochee, C.F. & Monica, T. Environmental effects on protein glycosylation. *Bio/Technology* **8**, 421-427 (1990).

104. Moellering, B.J. et al. Electrophoretic differences in a mAb expressed in three media. *Biopharm Int* **3**, 30-38 (1989).

105. Tachibana, H. et al. Changes of monosaccharide availability of human hybridoma lead to alteration of biological properties of human monoclonal antibody. *Cytotechnology* **16**, 151-157 (1994).

106. Hayter, P.M. et al. Glucose-limited chemostat culture of chinese hamster ovary cells producing recombinant human interferon-gamma. *Biotechnol Bioeng* **39**, 327-335 (1992).

107. Nyberg, G.B., Balcarcel, R.R., Follstad, B.D., Stephanopoulos, G. & Wang, D.I.C. Metabolic effects on recombinant interferon-gamma glycosylation in continuous culture of Chinese hamster ovary cells. *Biotechnol Bioeng* **62**, 336-347 (1999).

108. Wong, D.C.F., Wong, K.T.K., Goh, L.T., Heng, C.K. & Yap, M.G.S. Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures. *Biotechnol Bioeng* **89**, 164-177 (2005).

109. Wright, A. et al. In vivo trafficking and catabolism of IgG1 antibodies with Fc associated carbohydrates of differing structure. *Glycobiology* **10**, 1347-1355 (2000).

110. Gu, X. & Wang, D.I. Improvement of interferon-gamma sialylation in Chinese hamster ovary cell culture by feeding of N-acetylmannosamine. *Biotechnol Bioeng* **58**, 642-648 (1998).

111. Rijcken, W.R.P., Overdijk, B., den Eijnden, D.H.V. & Ferwerda, W. The effect of increasing nucleotide-sugar concentrations on the incorporation of sugars into glycoconjugates in rat hepatocytes. *Biochem J* **305 (Pt 3)**, 865-870 (1995).

112. Trummer, E. et al. Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors. *Biotechnol Bioeng* **94**, 1033-1044 (2006).

113. Chotigeat, W., Watanapokasin, Y., Mahler, S. & Gray, P.P. Role of environmental conditions on the expression levels, glycoform pattern and levels of sialyltransferase for hFSH produced by recombinant CHO cells. *Cytotechnology* **15**, 217-221 (1994).

114. Gawlitzek, M., Conradt, H.S. & Wagner, R. Effect of different cell culture conditions on the polypeptide integrity and N-glycosylation of a recombinant model glycoprotein. *Biotechnol Bioeng* **46**, 536-544 (1995).

115. Kunkel, J.P., Jan, D.C.H., Jamieson, J.C. & Butler, M. Dissolved oxygen concentration in serum-free continuous culture affects N-linked glycosylation of a monoclonal antibody. *J Biotechnol* **62**, 55-71 (1998).

116. Regoeczi, E., Kay, J.M., Chindemi, P.A., Zaimi, O. & Suyama, K.L. Transferrin glycosylation in hypoxia. *Biochem Cell Biol* **69**, 239-244 (1991).

117. Borys, M.C., Linzer, D.I. & Papoutsakis, E.T. Culture pH affects expression rates and glycosylation of recombinant mouse placental lactogen proteins by Chinese hamster ovary (CHO) cells. *Bio/Technology* **11**, 720-724 (1993).

118. Gawlitzek, M., Ryll, T., Lofgren, J. & Sliwkowski, M.B. Ammonium alters N-glycan structures of recombinant TNFR-IgG: degradative versus biosynthetic mechanisms. *Biotechnol Bioeng* **68**, 637-646 (2000).

119. Axelsson, M.A. et al. Neutralization of pH in the Golgi apparatus causes redistribution of glycosyltransferases and changes in the O-glycosylation of mucins. *Glycobiology* **11**, 633-644 (2001).

120. Rivinoja, A., Hassinen, A., Kokkonen, N., Kauppila, A. & Kellokumpu, S. Elevated Golgi pH impairs terminal N-glycosylation by inducing mislocalization of Golgi glycosyltransferases. *J Cell Physiol* **220**, 144-154 (2009).

121. Gawlitzek, M., Valley, U. & Wagner, R. Ammonium ion and glucosamine dependent increases of oligosaccharide complexity in recombinant glycoproteins secreted from cultivated BHK-21 cells. *Biotechnol Bioeng* **57**, 518-528 (1998).

122. Yoon, S.K., Choi, S.L., Song, J.Y. & Lee, G.M. Effect of culture pH on erythropoietin production by Chinese hamster ovary cells grown in suspension at 32.5 and 37.0 degrees C. *Biotechnol Bioeng* **89**, 345-356 (2005).

123. Müthing, J. et al. Effects of buffering conditions and culture pH on production rates and glycosylation of clinical phase I anti-melanoma mouse IgG3 monoclonal antibody R24. *Biotechnol Bioeng* **83**, 321-334 (2003).

124. Andersen, D.C., Bridges, T., Gawlitzek, M. & Hoy, C. Multiple cell culture factors can affect the glycosylation of Asn-184 in CHO-produced tissue-type plasminogen activator. *Biotechnol Bioeng* **70**, 25-31 (2000).

125. Yoon, S.K., Kim, S.H. & Lee, G.M. Effect of low culture temperature on specific productivity and transcription level of anti-4-1BB antibody in recombinant Chinese hamster ovary cells. *Biotechnol Prog* **19**, 1383-1386 (2003).

126. Ahn, W.S., Jeon, J.-J., Jeong, Y.-R., Lee, S.J. & Yoon, S.K. Effect of culture temperature on erythropoietin production and glycosylation in a perfusion culture of recombinant CHO cells. *Biotechnol Bioeng* **101**, 1234-1244 (2008).

127. Kukuruzinska, M.A. & Lennon, K. Growth-related coordinate regulation of the early N-glycosylation genes in yeast. *Glycobiology* **4**, 437-443 (1994).

128. Barnabé, N. & Butler, M. Effect of temperature on nucleotide pools and monoclonal antibody production in a mouse hybridoma. *Biotechnol Bioeng* **44**, 1235-1245 (1994).

129. Jenkins, N., Castro, P., Menon, S., Ison, A. & Bull, A. Effect of lipid supplements on the production and glycosylation of recombinant interferon-gamma expressed in CHO cells. *Cytotechnology* **15**, 209-215 (1994).

130. Serrato, J.A., Hernández, V., Estrada-Mondaca, S., Palomares, L.A. & Ramírez, O.T. Differences in the glycosylation profile of a monoclonal antibody produced by hybridomas cultured in serum-supplemented, serum-free or chemically defined media. *Biotechnol Appl Biochem* **47**, 113-124 (2007).

131. Schmelzer, A.E. & Miller, W.M. Hyperosmotic stress and elevated pCO2 alter monoclonal antibody charge distribution and monosaccharide content. *Biotechnol Prog* **18**, 346-353 (2002).

132. Yang, M. & Butler, M. Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. *Biotechnol Bioeng* **68**, 370-380 (2000).

133. Senger, R.S. & Karim, M.N. Effect of shear stress on intrinsic CHO culture state and glycosylation of recombinant tissue-type plasminogen activator protein. *Biotechnol Prog* **19**, 1199-1209 (2003).

134. Lund, J. et al. Control of IgG/Fc glycosylation: a comparison of oligosaccharides from chimeric human/mouse and mouse subclass immunoglobulin Gs. *Mol Immunol* **30**, 741-748 (1993).

135. Goldman, M.H. et al. Monitoring recombinant human interferon-gamma n-glycosylation during perfused fluidized-bed and stirred-tank batch culture of CHO cells. *Biotechnol Bioeng* **60**, 596-607 (1998).

136. Majid, F.A.A., Butler, M. & Al-Rubeai, M. Glycosylation of an immunoglobulin produced from a murine hybridoma cell line: the effect of culture mode and the anti-apoptotic gene, bcl-2. *Biotechnol Bioeng* **97**, 156-169 (2007).

137. Yoon, S.K., Ahn, Y.-H. & Jeong, M.H. Effect of culture temperature on follicle-stimulating hormone production by Chinese hamster ovary cells in a perfusion bioreactor. *Appl Microbiol Biotechnol* **76**, 83-89 (2007).

138. Robinson, D.K. et al. Characterization of a recombinant antibody produced in the course of a high yield fed-batch process. *Biotechnol Bioeng* **44**, 727-735 (1994).

139. Trummer, E. et al. Process parameter shifting: Part II. Biphasic cultivation-A tool for enhancing the volumetric productivity of batch processes using Epo-Fc expressing CHO cells. *Biotechnol Bioeng* **94**, 1045-1052 (2006).

140. Shuler, M.L. & Kargi, F. Bioprocess Engineering: Basic Concepts, Edn. 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ 07458; 2002).

141. Raju, T., Briggs, J., Borge, S. & Jones, A. Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* **10**, 477-486 (2000).

142. Sheeley, D.M., Merrill, B.M. & Taylor, L.C. Characterization of monoclonal antibody glycosylation: comparison of expression systems and identification of terminal alpha-linked galactose. *Anal Biochem* **247**, 102-110 (1997).

143. Beck, A. et al. Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NS0 cells. *J Chromatogr B Analyt Technol Biomed Life Sci* **819**, 203-218 (2005).

144. Wagner-Rousset, E. et al. The way forward, enhanced characterization of therapeutic antibody glycosylation: comparison of three level mass spectrometry-based strategies. *J Chromatogr B Analyt Technol Biomed Life Sci* **872**, 23-37 (2008).

145. Chung, C.H. et al. Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N Engl J Med* **358**, 1109-1117 (2008).

146. Wacker, M. et al. N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. *Science* **298**, 1790-1793 (2002).

147. Kowarik, M. et al. N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314**, 1148-1150 (2006).

148. Kowarik, M. et al. Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* **25**, 1957-1966 (2006).

149. Linton, D. et al. Functional analysis of the Campylobacter jejuni N-linked protein glycosylation pathway. *Mol Microbiol* **55**, 1695-1703 (2005).

150. Gerngross, T.U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat Biotechnol* **22**, 1409-1414 (2004).

151. Li, H. et al. Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nat Biotechnol* **24**, 210-215 (2006).

152. Tekoah, Y. et al. Controlled glycosylation of therapeutic antibodies in plants. *Arch Biochem Biophys* **426**, 266-278 (2004).

153. Hooker, A.D. et al. Constraints on the transport and glycosylation of recombinant IFN-gamma in Chinese hamster ovary and insect cells. *Biotechnol Bioeng* **63**, 559-572 (1999).

154. Betenbaugh, M.J., Tomiya, N., Narang, S., Hsu, J.T. & Lee, Y.C. Biosynthesis of human-type N-glycans in heterologous systems. *Curr Opin Struct Biol* **14**, 601-606 (2004).

155. Tomiya, N. Humanization of recombinant glycoproteins expressed in insect cells. *Trends Glycosci Glyc* **21**, 71-86 (2009).

156. Segawa, H., Kawakita, M. & Ishida, N. Human and Drosophila UDP-galactose transporters transport UDP-N-acetylgalactosamine in addition to UDP-galactose. *Eur J Biochem* **269**, 128-138 (2002).

157. Altmann, F., Kornfeld, G., Dalik, T., Staudacher, E. & Glössl, J. Processing of asparagine-linked oligosaccharides in insect cells. N-acetylglucosaminyltransferase I and II activities in cultured lepidopteran cells. *Glycobiology* **3**, 619-625 (1993).

158. Léonard, R. et al. The Drosophila fused lobes gene encodes an N-acetylglucosaminidase involved in N-glycan processing. *J Biol Chem* **281**, 4867-4875 (2006).

159. Hamilton, S.R. et al. Production of complex human glycoproteins in yeast. *Science* **301**, 1244-1246 (2003).

160. Bobrowicz, P. et al. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast Pichia pastoris: production of complex humanized glycoproteins with terminal galactose. *Glycobiology* **14**, 757-766 (2004).

161. Jarvis, D.L. & Finn, E.E. Modifying the insect cell N-glycosylation pathway with immediate early baculovirus expression vectors. *Nat Biotechnol* **14**, 1288-1292 (1996).

162. Bakker, H. et al. Galactose-extended glycans of antibodies produced by transgenic plants. *Proc Natl Acad Sci U S A* **98**, 2899-2904 (2001).

163. Cox, K.M. et al. Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat Biotechnol* **24**, 1591-1597 (2006).

164. Shelikoff, M., Sinskey, A.J. & Stephanopoulos, G. A modeling framework for the study of protein glycosylation. *Biotechnology and Bioengineering* **50**, 73-90 (1996).

165. Losev, E. et al. Golgi maturation visualized in living yeast. *Nature* **441**, 1002-1006 (2006).

166. Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K. & Nakano, A. Live imaging of yeast Golgi cisternal maturation. *Nature* **441**, 1007-1010 (2006).

167. PAT -- A framework for innovative pharmaceutical development, manufacturing, and quality assurance. *Guidance for industry* (2004).

168. Marino, K., Bones, J., Kattla, J.J. & Rudd, P.M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* **6**, 713-723 (2010).

169. Huhn, C., Selman, M.H.J., Ruhaak, L.R., Deelder, A.M. & Wuhrer, M. IgG glycosylation analysis. *Proteomics* **9**, 882-913 (2009).

170. Olivova, P., Chen, W., Chakraborty, A.B. & Gebler, J.C. Determination of N-glycosylation sites and site heterogeneity in a monoclonal antibody by electrospray quadrupole ion-mobility time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **22**, 29-40 (2008).

171. Lim, A., Reed-Bogan, A. & Harmon, B.J. Glycosylation profiling of a therapeutic recombinant monoclonal antibody with two N-linked glycosylation sites using liquid chromatography coupled to a hybrid quadrupole time-of-flight mass spectrometer. *Anal Biochem* **375**, 163-172 (2008).

172. Adamo, M. et al. Evaluation of oligosaccharide methods for carbohydrate analysis in a fully human monoclonal antibody and comparison of the results to the monosaccharide composition determination by a novel calculation. *J Pharm Biomed Anal* **49**, 181-192 (2009).

173. Serrato, J.A., Palomares, L.A., Meneses-Acosta, A. & Ramírez, O.T. Heterogeneous conditions in dissolved oxygen affect N-glycosylation but not productivity of a monoclonal antibody in hybridoma cultures. *Biotechnol Bioeng* **88**, 176-188 (2004).

174. Matamoros Fernandez, L.E. et al. Characterization of a recombinant monoclonal antibody by mass spectrometry combined with liquid chromatography. *J Chromatogr B Biomed Sci Appl* **752**, 247-261 (2001).

175. Saba, J.A. et al. A study of immunoglobulin G glycosylation in monoclonal and polyclonal species by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Anal Biochem* **305**, 16-31 (2002).

176. Kamoda, S., Ishikawa, R. & Kakehi, K. Capillary electrophoresis with laser-induced fluorescence detection for detailed studies on N-linked oligosaccharide profile of therapeutic recombinant monoclonal antibodies. *J Chromatogr A* **1133**, 332-339 (2006).

177. Morelle, W. & Michalski, J.-C. Analysis of protein glycosylation by mass spectrometry. *Nat Protoc* **2**, 1585-1602 (2007).

178. Mechref, Y., Muzikar, J. & Novotny, M.V. Comprehensive assessment of N-glycans derived from a murine monoclonal antibody: a case for multimethodological approach. *Electrophoresis* **26**, 2034-2046 (2005).

179. Qian, J. et al. Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole-quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. *Anal Biochem* **364**, 8-18 (2007).

180. Siemiatkoski, J., Lyubarskaya, Y., Houde, D., Tep, S. & Mhatre, R. A comparison of three techniques for quantitative carbohydrate analysis used in characterization of therapeutic antibodies. *Carbohydr Res* **341**, 410-419 (2006).

181. O'Shea, T., Lunte, S. & LaCourse, W. Detection of carbohydrates by capillary electrophoresis with pulsed amperometric detection. *Anal Chem* **65**, 948-951 (1993).

182. Lee, Y. Carbohydrate analyses with high-performance anion-exchange chromatography. *J Chromatogr A* **720**, 137-149 (1996).

183. Clarke, A., Harmon, B. & DeFelippis, M.R. Analysis of 3-(acetylamino)-6-aminoacridine-derivatized oligosaccharides from recombinant monoclonal antibodies by liquid chromatography-mass spectrometry. *Anal Biochem* **390**, 209-211 (2009).

184. Prater, B.D., Connelly, H.M., Qin, Q. & Cockrill, S.L. High-throughput immunoglobulin G N-glycan characterization using rapid resolution reverse-phase chromatography tandem mass spectrometry. *Anal Biochem* **385**, 69-79 (2009).

185. Lamari, F.N., Kuhn, R. & Karamanos, N.K. Derivatization of carbohydrates for chromatographic, electrophoretic and mass spectrometric structure analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* **793**, 15-36 (2003).

186. Pabst, M. et al. Comparison of fluorescent labels for oligosaccharides and introduction of a new postlabeling purification method. *Anal Biochem* **384**, 263-273 (2009).

187. Kinoshita, A. & Sugahara, K. Microanalysis of glycosaminoglycan-derived oligosaccharides labeled with a fluorophore 2-aminobenzamide by high-performance liquid chromatography: application to disaccharide composition analysis and exosequencing of oligosaccharides. *Anal Biochem* **269**, 367-378 (1999).

188. Domann, P.J. et al. Separation-based glycoprofiling approaches using fluorescent labels. *Proteomics* **7 Suppl 1**, 70-76 (2007).

189. Gennaro, L.A. & Salas-Solano, O. On-line CE-LIF-MS technology for the direct characterization of N-linked glycans from therapeutic antibodies. *Anal Chem* **80**, 3838-3845 (2008).

190. O'Shea, M., Samuel, M., Konik, C. & Morell, M. Fluorophore-assisted carbohydrate electrophoresis (FACE) of oligosaccharides: efficiency of labelling and high-resolution separation. *Carbohyd Res* **307**, 1-12 (1998).

191. Nakano, M. et al. Capillary electrophoresis-electrospray ionization mass spectrometry for rapid and sensitive N-glycan analysis of glycoproteins as 9-fluorenylmethyl derivatives. *Glycobiology* **19**, 135-143 (2009).

192. Chen, X. & Flynn, G.C. Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry. *Anal Biochem* **370**, 147-161 (2007).

193. Wada, Y. et al. Comparison of the methods for profiling glycoprotein glycans - HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* **17**, 411-422 (2007).

194. Ma, S. & Nashabeh, W. Carbohydrate analysis of a chimeric recombinant monoclonal antibody by capillary electrophoresis with laser-induced fluorescence detection. *Anal Chem* **71**, 5185-5192 (1999).

195. Kunkel, J.P., Jan, D.C., Butler, M. & Jamieson, J.C. Comparisons of the glycosylation of a monoclonal antibody produced under nominally identical cell culture conditions in two different bioreactors. *Biotechnol Prog* **16**, 462-470 (2000).

196. Elbers, I.J.W. et al. Influence of growth conditions and developmental stage on N-glycan heterogeneity of transgenic immunoglobulin G and endogenous proteins in tobacco leaves. *Plant Physiology* **126**, 1314-1322 (2001).

197. Damen, C.W.N. et al. Electrospray Ionization Quadrupole Ion-Mobility Time-of-Flight Mass Spectrometry as a Tool to Distinguish the Lot-to-Lot Heterogeneity in N-Glycosylation Profile of the Therapeutic Monoclonal Antibody Trastuzumab. *J Am Soc Mass Spectr* **20**, 2021-2033 (2009).

198. Loos, A. et al. Production of monoclonal antibodies with a controlled N-glycosylation pattern in seeds of Arabidopsis thaliana. *Plant Biotechnol J* **9**, 179-192 (2011).

199. Mauko, L. et al. Glycan profiling of monoclonal antibodies using zwitterionic-type hydrophilic interaction chromatography coupled with electrospray ionization mass spectrometry detection. *Analytical Biochemistry* **408**, 235-241 (2011).

200. Wacker, C., Berger, C.N., Girard, P. & Meier, R. Glycosylation profiles of therapeutic antibody pharmaceuticals. *Eur J Pharm Biopharm* **79**, 503-507 (2011).

201. Montesino, R. et al. Structural characterization of N-linked oligosaccharides on monoclonal antibody Nimotuzumab through process development. *Biologicals* **40**, 288-298 (2012).

202. Ito, H. et al. Strategy for Glycoproteomics: Identification of Glyco-Alteration Using Multiple Glycan Profiling Tools. *J Proteome Res* **8**, 1358-1367 (2009).

203. Mittermayr, S., Bones, J., Doherty, M., Guttman, A. & Rudd, P.M. Multiplexed Analytical Glycomics: Rapid and Confident IgG N-Glycan Structural Elucidation. *J Proteome Res* **10**, 3820-3829 (2011).

204. Ceroni, A. et al. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of Glycans. *J Proteome Res* **7**, 1650-1659 (2008).

205. Royle, L., Radcliffe, C.M., Dwek, R.A. & Rudd, P.M. Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods in molecular biology* **347**, 125-143 (2006).

206. Royle, L., Dwek, R.A. & Rudd, P.M. Determining the structure of oligosaccharides N- and O-linked to glycoproteins. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **Chapter 12**, Unit 12 16 (2006).

207. Artemenko, N.V., McDonald, A.G., Davey, G.P. & Rudd, P.M. Databases and tools in glycobiology. *Methods in molecular biology* **899**, 325-350 (2012).

208. Yu, C.Y., Mayampurath, A. & Tang, H. Software tools for glycan profiling. *Methods in molecular biology* **951**, 269-276 (2013).

209. Doubet, S. & Albersheim, P. Carbbank. *Glycobiology* **2**, 505-505 (1992).

210. Ranzinger, R., Herget, S., Wetter, T. & von der Lieth, C.W. GlycomeDB - integration of open-access carbohydrate structure databases. *Bmc Bioinformatics* **9** (2008).

211. Ranzinger, R., Herget, S., von der Lieth, C.W. & Frank, M. GlycomeDB-a unified database for carbohydrate structures. *Nucleic Acids Res* **39**, D373-D376 (2011).

212. Hashimoto, K. et al. A global representation of the carbohydrate structures: a tool for the analysis of glycan. *Genome informatics. International Conference on Genome Informatics* **16**, 214-222 (2005).

213. Hashimoto, K. et al. KEGG as a glycome informatics resource. *Glycobiology* **16**, 63r-70r (2006).

214. Campbell, M.P., Royle, L., Radcliffe, C.M., Dwek, R.A. & Rudd, P.M. GlycoBase and autoGU: tools for HPLC-based glycan analysis. *Bioinformatics* **24**, 1214-1216 (2008).

215. Damerell, D. et al. The GlycanBuilder and GlycoWorkbench glycoinformatics tools: updates and new developments. *Biological chemistry* **393**, 1357-1362 (2012).

216. Lutteke, T. et al. GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research. *Glycobiology* **16**, 71R-81R (2006).

217. Toukach, P., Joshi, H.J., Ranzinger, R., Knirel, Y. & von der Lieth, C.W. Sharing of worldwide distributed carbohydrate-related digital resources: online connection of the Bacterial Carbohydrate Structure DataBase and GLYCOSCIENCES.de. *Nucleic Acids Res* **35**, D280-286 (2007).

218. Wu, H.Q. & Khan, M.A. Quality-by-Design (QbD): An Integrated Process Analytical Technology (PAT) Approach for Real-Time Monitoring and Mapping the State of a Pharmaceutical Coprecipitation Process. *J Pharm Sci-Us* **99**, 1516-1534 (2010).

219. Huang, J. et al. A PAT Approach to Improve Process Understanding of High Shear Wet Granulation Through In-Line Particle Measurement Using FBRM C35. *J Pharm Sci-Us* **99**, 3205-3212 (2010).

220. Wu, H.Q. & Khan, M.A. Quality-By-Design (QbD): An Integrated Approach for Evaluation of Powder Blending Process Kinetics and Determination of Powder Blending End-point. *J Pharm Sci-Us* **98**, 2784-2798 (2009).

221. Brooks, S.A. Strategies for Analysis of the Glycosylation of Proteins: Current Status and Future Perspectives. *Mol Biotechnol* **43**, 76-88 (2009).

222. Doherty, M., McManus, C.A., Duke, R. & Rudd, P.M. High-throughput quantitative N-glycan analysis of glycoproteins. *Methods in molecular biology* **899**, 293-313 (2012).

223. Royle, L. et al. HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. *Analytical Biochemistry* **376**, 1-12 (2008).

224. Flynn, G.C., Chen, X., Liu, Y.D., Shah, B. & Zhang, Z. Naturally occurring glycan forms of human immunoglobulins G1 and G2. *Molecular Immunology* **47**, 2074-2082 (2010).

225. del Val, I.J., Nagy, J.M. & Kontoravdi, C. A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnol Progr* **27**, 1730-1743 (2011).

226. Glick, B.S. & Nakano, A. Membrane traffic within the Golgi apparatus. *Annu Rev Cell Dev Biol* **25**, 113-132 (2009).

227. Nilsson, T., Au, C.E. & Bergeron, J.J.M. Sorting out glycosylation enzymes in the Golgi apparatus. *Febs Letters* **583**, 3764-3769 (2009).

228. Glick, B.S. & Malhotra, V. The curious status of the Golgi apparatus. *Cell* **95**, 883-889 (1998).

229. Pelham, H.R. Getting through the Golgi complex. *Trends Cell Biol* **8**, 45-49 (1998).

230. Ihara, H. et al. Crystal structure of mammalian alpha1,6-fucosyltransferase, FUT8. *Glycobiology* **17**, 455-466 (2007).

231. Ramakrishnan, B., Boeggeman, E., Ramasamy, V. & Qasba, P.K. Structure and catalytic cycle of beta-1,4-galactosyltransferase. *Curr Opin Struct Biol* **14**, 593-600 (2004).

232. Ikeda, Y. et al. Kinetic basis for the donor nucleotide-sugar specificity of beta1, 4-N-acetylglucosaminyltransferase III. *J Biochem* **128**, 609-619 (2000).

233. Tabas, I. & Kornfeld, S. Purification and characterization of a rat liver Golgi alpha-mannosidase capable of processing asparagine-linked oligosaccharides. *J Biol Chem* **254**, 11655-11663 (1979).

234. Moremen, K.W., Touster, O. & Robbins, P.W. Novel purification of the catalytic domain of Golgi alpha-mannosidase II. Characterization and comparison with the intact enzyme. *J Biol Chem* **266**, 16876-16885 (1991).

235. Velasco, A. et al. Cell type-dependent variations in the subcellular distribution of alpha-mannosidase I and II. *J Cell Biol* **122**, 39-51 (1993).

236. Unligil, U.M. et al. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* **19**, 5269-5280 (2000).

237. Bendiak, B. & Schachter, H. Control of glycoprotein synthesis. Kinetic mechanism, substrate specificity, and inhibition characteristics of UDP-N-acetylglucosamine:alpha-D-mannoside beta 1-2 N-acetylglucosaminyltransferase II from rat liver. *J Biol Chem* **262**, 5784-5790 (1987).

238. Bendiak, B. & Cook, G.M. Kinetic parameters of a beta-D-galactoside alpha 2 leads to 6 sialyltransferase from embryonic chicken liver. *Eur J Biochem* **128**, 355-362 (1982).

239. Hirschberg, C.B., Robbins, P.W. & Abeijon, C. Transporters of nucleotide sugars, ATP, and nucleotide sulfate in the endoplasmic reticulum and Golgi apparatus. *Annu Rev Biochem* **67**, 49-69 (1998).

240. Waldman, B.C. & Rudnick, G. UDP-GlcNAc transport across the Golgi membrane: electroneutral exchange for dianionic UMP. *Biochemistry* **29**, 44-52 (1990).

241. Lepers, A. et al. Transport of CMP-N-glycoloylneuraminic acid into mouse liver Golgi vesicles. *Febs Letters* **250**, 245-250 (1989).

242. Milla, M.E., Clairmont, C.A. & Hirschberg, C.B. Reconstitution into proteoliposomes and partial purification of the Golgi apparatus membrane UDP-galactose, UDP-xylose, and UDP-glucuronic acid transport activities. *J Biol Chem* **267**, 103-107 (1992).

243. Sommers, L.W. & Hirschberg, C.B. Transport of sugar nucleotides into rat liver Golgi. A new Golgi marker activity. *J Biol Chem* **257**, 10811-10817 (1982).

244. Frame, K.K. & Hu, W.S. Cell volume measurement as an estimation of mammalian cell biomass. *Biotechnol Bioeng* **36**, 191-197 (1990).

245. Ladinsky, M.S., Mastronarde, D.N., McIntosh, J.R., Howell, K.E. & Staehelin, L.A. Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *J Cell Biol* **144**, 1135-1149 (1999).

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

247. Bibila, T.A. & Robinson, D.K. In pursuit of the optimal fed-batch process for monoclonal antibody production. *Biotechnol Prog* **11**, 1-13 (1995).

248. Hirschberg, K. & Lippincott-Schwartz, J. Secretory pathway kinetics and in vivo analysis of protein traffic from the Golgi complex to the cell surface. *FASEB J* **13 Suppl 2**, S251-S256 (1999).

249. Karaveg, K. & Moremen, K.W. Energetics of substrate binding and catalysis by class 1 (glycosylhydrolase family 47) alpha-mannosidases involved in N-glycan processing and endoplasmic reticulum quality control. *J Biol Chem* **280**, 29837-29848 (2005).

250. Tulsiani, D.R. & Touster, O. Mannosidase IA from rat liver Golgi membranes. *Methods Enzymol* **179**, 446-451 (1989).

251. Nishikawa, Y., Pegg, W., Paulsen, H. & Schachter, H. Control of glycoprotein synthesis. Purification and characterization of rabbit liver UDP-N-acetylglucosamine:alpha-3-D-mannoside beta-1,2-N-acetylglucosaminyltransferase I. *J Biol Chem* **263**, 8270-8281 (1988).

252. Fukada, T., Iida, K., Kioka, N., Sakai, H. & Komano, T. Cloning of a cDNA encoding N-acetylglucosaminyltransferase I from rat liver and analysis of its expression in rat tissues. *Biosci Biotechnol Biochem* **58**, 200-201 (1994).

253. Nishikawa, A., Ihara, Y., Hatakeyama, M., Kangawa, K. & Taniguchi, N. Purification, cDNA cloning, and expression of UDP-N-acetylglucosamine: beta-D-mannoside beta-1,4N-acetylglucosaminyltransferase III from rat kidney. *J Biol Chem* **267**, 18199-18204 (1992).

254. Uozumi, N. et al. Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha1-->6fucosyltransferase. *J Biol Chem* **271**, 27810-27817 (1996).

255. Bendiak, B., Ward, L.D. & Simpson, R.J. Proteins of the Golgi apparatus. Purification to homogeneity, N-terminal sequence, and unusually large Stokes radius of the membrane-bound form of UDP-galactose:N-acetylglucosamine beta 1-4galactosyltransferase from rat liver. *Eur J Biochem* **216**, 405-417 (1993).

256. Gerhard, D.S. et al. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res* **14**, 2121-2127 (2004).

257. Weinstein, J., de Souza-e-Silva, U. & Paulson, J.C. Purification of a Gal beta 1 to 4GlcNAc alpha 2 to 6 sialyltransferase and a Gal beta 1 to 3(4)GlcNAc alpha 2 to 3 sialyltransferase to homogeneity from rat liver. *J Biol Chem* **257**, 13835-13844 (1982).

258. Harduin-Lepers, A., Mollicone, R., Delannoy, P. & Oriol, R. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology* **15**, 805-817 (2005).

259. Chandrasekaran, E.V., Davila, M., Nixon, D. & Mendicino, J. Purification and properties of alpha-D-mannose:beta-1,2-N-acetylglucosaminyl-transferases and alpha-D-mannosidases from human adenocarcinoma. *Cancer Research* **44**, 4059-4068 (1984).

260. Chen, W., Unligil, U.M., Rini, J.M. & Stanley, P. Independent Lec1A CHO glycosylation mutants arise from point mutations in N-acetylglucosaminyltransferase I that reduce affinity for both substrates. Molecular consequences based on the crystal structure of GlcNAc-TI. *Biochemistry* **40**, 8765-8772 (2001).

261. Pâquet, M.R., Narasimhan, S., Schachter, H. & Moscarello, M.A. Branch specificity of purified rat liver Golgi UDP-galactose: N-acetylglucosamine beta-1,4-galactosyltransferase. Preferential transfer of of galactose on the GlcNAc beta 1,2-Man alpha 1,3-branch of a complex biantennary Asn-linked oligosaccharide. *J Biol Chem* **259**, 4716-4721 (1984).

262. Datta, A.K. & Paulson, J.C. The sialyltransferase "sialylmotif" participates in binding the donor substrate CMP-NeuAc. *J Biol Chem* **270**, 1497-1500 (1995).

263. Kochanowski, N. et al. Intracellular nucleotide and nucleotide sugar contents of cultured CHO cells determined by a fast, sensitive, and high-resolution ion-pair RP-HPLC. *Anal Biochem* **348**, 243-251 (2006).

264. Yurchenco, P.D. & Atkinson, P.H. Fucosyl-glycoprotein and precursor polls in HeLa cells. *Biochemistry* **14**, 3107-3114 (1975).

265. Process Systems Enterprise, L. 2009).

266. Quinn, P., Griffiths, G. & Warren, G. Density of newly synthesized plasma membrane proteins in intracellular membranes II. Biochemical studies. *J Cell Biol* **98**, 2142-2147 (1984).

267. Gilchrist, A. et al. Quantitative proteomics analysis of the secretory pathway. *Cell* **127**, 1265-1281 (2006).

268. Baker, K.N. et al. Metabolic control of recombinant protein N-glycan processing in NS0 and CHO cells. *Biotechnology and Bioengineering* **73**, 188-202 (2001).

269. Ishida, N., Yoshioka, S., Chiba, Y., Takeuchi, M. & Kawakita, M. Molecular cloning and functional expression of the human Golgi UDP-N-acetylglucosamine transporter. *J Biochem* **126**, 68-77 (1999).

270. Yoshioka, S., Sun-Wada, G.H., Ishida, N. & Kawakita, M. Expression of the human UDP-galactose transporter in the Golgi membranes of murine Had-1 cells that lack the endogenous transporter. *J Biochem* **122**, 691-695 (1997).

271. Zhao, W., Chen, T.-L.L., Vertel, B.M. & Colley, K.J. The CMP-sialic acid transporter is localized in the medial-trans Golgi and possesses two specific endoplasmic reticulum export motifs in its carboxyl-terminal cytoplasmic tail. *J Biol Chem* **281**, 31106-31118 (2006).

272. Kontoravdi, C., Asprey, S.P., Pistikopoulos, E.N. & Mantalaris, A. Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring. *Comput Chem Eng* **31**, 392-400 (2007).

273. Jang, D.J. & Barford, J.P. An unstructured kinetic model of macromolecular metabolism in batch and fed-batch cultures of hybridoma cells producing monoclonal antibody. *Biochem Eng J* **4**, 153-168 (2000).

274. Ljunggren, J. & Haggstrom, L. Catabolic Control of Hybridoma Cells by Glucose and Glutamine Limited Fed-Batch Cultures. *Biotechnology and Bioengineering* **44**, 808-818 (1994).

275. Fitzpatrick, L., Jenkins, H.A. & Butler, M. Glucose and Glutamine-Metabolism of a Murine B-Lymphocyte Hybridoma Grown in Batch Culture. *Appl Biochem Biotech* **43**, 93-116 (1993).

276. Bibila, T.A. & Flickinger, M.C. Use of a Structured Kinetic-Model of Antibody-Synthesis and Secretion for Optimization of Antibody-Production Systems .1. Steady-State Analysis (Biotechnology and Bioengineering, Vol 39, Pg 251, 1992). *Biotechnology and Bioengineering* **41**, 682-682 (1993).

277. Hediger, M.A. et al. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction. *Pflugers Archiv : European journal of physiology* **447**, 465-468 (2004).

278. Wong, N.S.C. et al. An investigation of intracellular glycosylation activities in CHO cells: effects of nucleotide sugar precursor feeding. *Biotechnol Bioeng* **107**, 321-336 (2010).

279. Liebermeister, W. & Klipp, E. Biochemical networks with uncertain parameters. *Iee P Syst Biol* **152**, 97-107 (2005).

280. Dokoumetzidis, A. & Aarons, L. Proper lumping in systems biology models. *Iet Syst Biol* **3**, 40-51 (2009).

281. Nolan, R.P. & Lee, K. Dynamic model of CHO cell metabolism. *Metab Eng* **13**, 108-124 (2011).

282. Surovtsova, I., Simus, N., Hubner, K., Sahle, S. & Kummer, U. Simplification of biochemical models: a general approach based on the analysis of the impact of individual species and reactions on the systems dynamics. *Bmc Syst Biol* **6** (2012).

283. Voet, D. & Voet, J.G. Biochemistry, Edn. 4th. (John Wiley & Sons, Hoboken, NJ; 2011).

284. Scheer, M. et al. BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* **39**, D670-D676 (2011).

285. Cornish-Bowden, A. Fundamentals of enzyme kinetics, Edn. 4th, completely rev. and greatly enl. (Wiley-Blackwell, Weinheim; 2012).

286. Hinderlich, S., Stasche, R., Zeitler, R. & Reutter, W. A bifunctional enzyme catalyzes the first two steps in N-acetylneuraminic acid biosynthesis of rat liver - Purification and characterization of UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase. *Journal of Biological Chemistry* **272**, 24313-24318 (1997).

287. Bisso, A., Sturla, L., Zanardi, D., De Flora, A. & Tonetti, M. Structural and enzymatic characterization of human recombinant GDP-D-mannose-4,6-dehydratase. *Febs Letters* **456**, 370-374 (1999).

288. Tomiya, N., Ailor, E., Lawrence, S.M., Betenbaugh, M.J. & Lee, Y.C. Determination of nucleotides and sugar nucleotides involved in protein glycosylation by high-performance anion-exchange chromatography: Sugar nucleotide contents in cultured insect cells and mammalian cells. *Analytical Biochemistry* **293**, 129-137 (2001).

289. Kochanowski, N. et al. Influence of intracellular nucleotide and nucleotide sugar contents on recombinant interferon-gamma glycosylation during batch and fed-batch cultures of CHO cells. *Biotechnology and Bioengineering* **100**, 721-733 (2008).

290. Feng, H.T., Wong, N., Wee, S. & Lee, M.M. Simultaneous determination of 19 intracellular nucleotides and nucleotide sugars in Chinese Hamster ovary cells by capillary electrophoresis. *J Chromatogr B* **870**, 131-134 (2008).

291. Nakajima, K. et al. Simultaneous determination of nucleotide sugars with ion-pair reversed-phase HPLC. *Glycobiology* **20**, 865-871 (2010).

292. Sellick, C.A. et al. Effective Quenching Processes for Physiologically Valid Metabolite Profiling of Suspension Cultured Mammalian Cells. *Analytical Chemistry* **81**, 174-183 (2009).

293. Dietmair, S., Timmins, N.E., Gray, P.P., Nielsen, L.K. & Kromer, J.O. Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol. *Analytical Biochemistry* **404**, 155-164 (2010).

294. Peppers, S.C., Allison, D.W., Johnson, T.K., Talley, D.L. & Caple, M.V. Replacing the animal component in serum - Evaluating raw materials for inclusion in optimized hybridoma media. *Biopharm-Appl T Bio* **14**, 22-+ (2001).

295. Ozturk, S.S. & Palsson, B.O. Effects of Ammonia and Lactate on Cell-Growth, Metabolism and Monoclonal-Antibody Production. *Abstr Pap Am Chem S* **200**, 6-Biot (1990).

296. Ozturk, S.S. & Palsson, B.O. Physiological-Changes during the Adaptation of Hybridoma Cells to Low Serum and Serum-Free Media. *Biotechnology and Bioengineering* **37**, 35-46 (1991).

297. Schachter, H. The joys of HexNAc. The synthesis and function of N-and O-glycan branches. *Glycoconjugate J* **17**, 465-483 (2000).

298. Berg, J.M., Tymoczko, J.L. & Stryer, L. Biochemistry, Edn. 6th. (W.H. Freeman, New York; 2007).

Appendix I

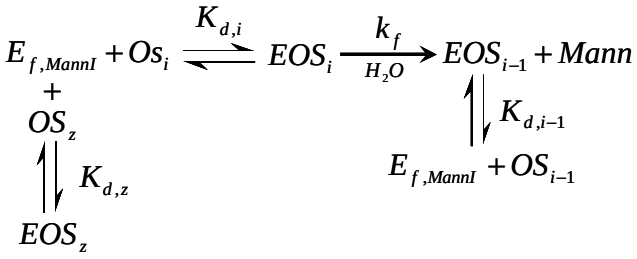
Rate expression derivation

Michaelis-Menten kinetics

This rate expression derivation applies to the ManI and ManII enzymes.[233-235, 251] In a Michaelis-Menten kinetic mechanism, the substrate binds to the enzyme to form an enzyme-substrate complex. Once the complex is formed, the rate-limiting step of the catalytic cycle occurs, thus yielding the product. In the case of the general reaction scheme shown in Figure 8, it can be seen that ManI and ManII act on different substrates. Because of this, a competitive inhibition term will be added to the catalytic cycle. Additional underlying assumptions for this derivation are the following:

a) Equilibrium is reached rapidly for all reactions.
b) There is competitive inhibition through substrate analogues.
c) There is product inhibition because the product of these reactions has high binding affinity to the enzyme because it is a subsequent reactant itself.
d) The enzyme needs a water molecule to catalyse the reaction. Since the reaction occurs within an aqueous environment (lumen of the Golgi compartments), it is possible to assume that the enzymes are saturated with water, and hence, the availability of this component does not affect the reaction in any way.
e) The rate-limiting step is assumed to be the chemical removal of the mannose residue from the oligosaccharide chain.
f) The rate-limiting step is assumed to be irreversible given that the products from the reaction will participate in further reactions within the same Golgi compartment. This will maintain their concentrations low and greatly reduces the possibility of the reaction occurring in the opposite sense. Another factor that facilitates this assumption is that the enzymes are saturated with water. This also contributes to reducing the likelihood of a reverse reaction. Finally, the co-products of these reactions are mannose residues. These mannose molecules have little or no interaction with the enzyme, so it is quite unlikely for them to react in the opposite sense.

These assumptions lead to the following kinetic mechanism:



From this kinetic reaction mechanism we can derive the following mathematical expressions:

Dissociation constants:

$$K_{d,i} = \frac{[E_f]\cdot[OS_i]}{[EOS_i]} \quad \rightarrow \quad [EOS_i] = \frac{[E_f]\cdot[OS_i]}{K_{d,i}} \qquad \textbf{...Eq. 45}$$

$$K_{d,i-1} = \frac{[E_f]\cdot[OS_{i-1}]}{[EOS_{i-1}]} \quad \rightarrow \quad [EOS_{i-1}] = \frac{[E_f]\cdot[OS_{i-1}]}{K_{d,i-1}} \qquad \textbf{...Eq. 46}$$

$$K_{d,z} = \frac{[E_f]\cdot[OS_z]}{[EOS_z]} \quad \rightarrow \quad [EOS_z] = \frac{[E_f]\cdot[OS_z]}{K_{d,z}} \qquad \textbf{...Eq. 47}$$

Total (or initial) enzyme balance:

$$[E_0] = [E_f] + [EOS_i] + \sum_{z=1}^{NC}[EOS_z] + [EOS_{i-1}] \qquad \textbf{...Eq. 48}$$

Combining the dissociation constant expressions with the enzyme balance:

$$[E_0] = [E_f] + [E_f]\cdot\frac{[OS_i]}{K_{d,i}} + [E_f]\cdot\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + [E_f]\cdot\frac{[OS_{i-1}]}{K_{d,i-1}}$$

Solving for the free enzyme term:

$$[E_f] = \frac{[E_0]}{\left(1 + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + \frac{[OS_{i-1}]}{K_{d,i-1}}\right)} \qquad \textbf{...Eq. 49}$$

The reaction rate expression for the rate-limiting reaction is defined as:

$$r_j = k_{f,j}\cdot[EOS_i] \qquad \textbf{...Eq. 50}$$

Combining Eq. 45 with Eq. 49:

$$[EOS_i]\cdot\frac{K_{d,i}}{[OS_i]} = \frac{[E_0]}{\left(1 + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + \frac{[OS_{i-1}]}{K_{d,i-1}}\right)}$$ and solving for $[EOS_i]$:

$$[EOS_i] = \frac{[E_0] \cdot \frac{[OS_i]}{K_{d,i}}}{\left(1 + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[OS_{i-1}]}{K_{d,i-1}}\right)} \quad \textbf{…Eq. 51}$$

Combining Eq. 50 with Eq. 51:

$$r_j = \frac{k_{f,j} \cdot [E_0] \cdot \frac{[OS_i]}{K_{d,i}}}{\left(1 + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[OS_{i-1}]}{K_{d,i-1}}\right)}$$ Finally, rearranging:

$$r_j = \frac{r_{MAX,j} \cdot [OS_i]}{K_{d,i} \cdot \left(1 + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[OS_{i-1}]}{K_{d,i-1}}\right)}, \text{ where } r_{MAX,j} = k_{f,j} \cdot [E_0] \quad \textbf{…Eq. 52}$$

Sequential-order Bi-Bi kinetics

This rate expression derivation applies to the GnTI, GnTII and GalT enzymes.[231, 236, 237] Although the underlying concepts of this derivation are similar to Michaelis-Menten, Bi-Bi kinetics consider binding of two substrates onto the enzyme for the reaction to occur. In sequential Bi-Bi kinetics, the order in which the substrates bind to the enzyme is important. For all three enzymes that catalyse reactions with this mechanism, the nucleotide sugar donor first binds to the enzyme and triggers a conformational change which creates the binding pocket for the oligosaccharide to bind. Thus the NSD binds first, and the oligosaccharide binds second. Because the enzyme undergoes a conformational change, the unbinding of the products from the enzyme occur in the inverse order: the oligosaccharide product dissociates from the enzyme, inducing the enzyme to recover its original configuration from which the by-product of the NSD (which for all of these reactions is UDP) can dissociate from the enzyme. Additional assumptions for the derivation of these kinetic expressions are the following:

a) Equilibrium is reached rapidly for all reactions.[237]
b) The kinetics of this reaction will be assumed as an ordered, sequential mechanism. In this case, the sugar donor must bind to the enzyme, once this occurs, the acceptor binds, and through the three species complex reacts to form product and by-product. Since the reaction is constrained to this order, the products complex must also dissociate in the corresponding

order. The product oligosaccharide unbinds leaving a by-product enzyme complex which will dissociate to complete the catalytic cycle.[237]

c) These enzymes may act on several different substrates along the N-glycosylation pathway. Hence, competitive inhibition from analogue acceptor oligosaccharide (OS_z) chains will be assumed.

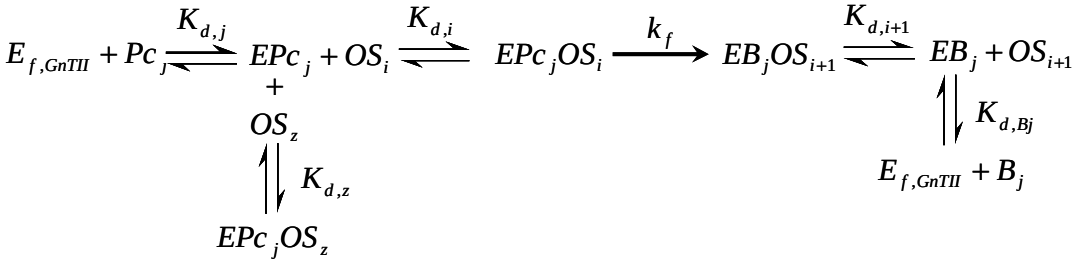
d) No excess of either of the substrates will be assumed which implies that the reaction rate will depend on the concentration of both substrates. (Enzyme is <u>not</u> assumed to be saturated with NSD).

e) Dual product inhibition will be considered for this reaction given that many reports indicate that these enzymes have affinity for substrate analogues, which in some cases include reaction product analogues.[297]

f) The rate-limiting step is assumed to be the transfer of the monosaccharide (GlcNAc or Gal) residue to the growing oligosaccharide chain.

g) The rate-limiting step is assumed to be irreversible due to the consumption of energy during this step.[64, 65]

These considerations lead to the following reaction mechanism:



From this kinetic reaction mechanism we can derive the following mathematical expressions:

Dissociation constants:

$$K_{d,k} = \frac{[E_f]\cdot[Pc_k]}{[EPc_k]} \longrightarrow [EPc_k] = \frac{[E_f]\cdot[Pc_k]}{K_{d,k}} \quad \textbf{...Eq. 53}$$

$$K_{d,i} = \frac{[EPc_k]\cdot[OS_i]}{[EPc_kOS_i]} \longrightarrow [EPc_kOS_i] = \frac{[EPc_k]\cdot[OS_i]}{K_{d,i}} \quad \textbf{...Eq. 54}$$

$$K_{d,z} = \frac{[EPc_k]\cdot[OS_z]}{[EPc_kOS_z]} \longrightarrow [EPc_kOS_z] = \frac{[EPc_k]\cdot[OS_z]}{K_{d,z}} \quad \textbf{...Eq. 55}$$

$$K_{d,i+1} = \frac{[EPc_k]\cdot[OS_{i+1}]}{[EPc_kOS_{i+1}]} \longrightarrow [EPc_kOS_{i+1}] = \frac{[EPc_k]\cdot[OS_{i+1}]}{K_{d,i+1}} \quad \textbf{...Eq. 56}$$

$$K_{d,Bk} = \frac{[E_f]\cdot[B_k]}{[EB_k]} \quad \rightarrow \quad [EB_k] = \frac{[E_f]\cdot[B_k]}{K_{d,Bk}} \qquad \textbf{…Eq. 57}$$

$$K_{d,k}\cdot K'_{d,i} = \frac{[E_f]\cdot[Pc_k]}{[EPc_k]}\cdot\frac{[EPc_k]\cdot[OS_i]}{[EPc_kOS_i]} = [E_f]\cdot\frac{[Pc_k]\cdot[OS_i]}{[EPc_kOS_i]} \qquad \textbf{…Eq. 58}$$

$$K_{d,k}\cdot K'_{d,z} = \frac{[E_f]\cdot[Pc_k]}{[EPc_k]}\cdot\frac{[EPc_k]\cdot[OS_z]}{[EPc_kOS_z]} = [E_f]\cdot\frac{[Pc_k]\cdot[OS_z]}{[EPc_kOS_z]} \qquad \textbf{…Eq. 59}$$

$$K_{d,i+1}\cdot K_{d,Bk} = \frac{[EB_k]\cdot[OS_{i+1}]}{[EB_kOS_{i+1}]}\cdot\frac{[E_f]\cdot[B_k]}{[EB_k]} = [E_f]\cdot\frac{[B_k]\cdot[OS_{i+1}]}{[EB_kOS_{i+1}]} \qquad \textbf{…Eq. 60}$$

Total (or initial) enzyme balance:

$$[E_0] = [E_f] + [EPc_k] + [EPc_kOS_i] + \sum_{z=1}^{NC}[EPc_kOS_z] + [EB_k] + [EB_kOS_{i+1}] \qquad \textbf{…Eq. 61}$$

Combining the dissociation constant expressions with the enzyme balance:

$$[E_0] = [E_f] + [E_f]\cdot\frac{[Pc_k]}{K_{d,k}} + [E_f]\cdot\frac{[Pc_k]}{K_{d,k}}\cdot\frac{[OS_i]}{K_{d,i}} + [E_f]\cdot\frac{[Pc_k]}{K_{d,k}}\cdot\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + [E_f]\cdot\frac{[B_k]}{K_{d,Bk}}\cdot\frac{[OS_{i+1}]}{K_{d,i+1}} + [E_f]\cdot\frac{[B_k]}{K_{d,Bk}}$$

…Eq. 62

Solving for the free enzyme term:

$$[E_f] = \frac{[E_0]}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[Pc_k]}{K_{d,k}}\cdot\frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K_{d,k}}\cdot\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}}\cdot\frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}}\right)} \qquad \textbf{…Eq. 63}$$

The reaction rate expression for the rate-limiting reaction is defined as:

$$r_j = k_{f,j}\cdot[EPc_kOS_i] \qquad \textbf{…Eq. 64}$$

Combining Eq. 63 with Eq. 58:

$$[EPc_kOS_i]\cdot\frac{K_{d,k}}{[Pc_k]}\cdot\frac{K_{d,i}}{[OS_i]} = \frac{[E_0]}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[Pc_k]}{K_{d,k}}\cdot\frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K_{d,k}}\cdot\sum_{z=1}^{NC}\frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}}\cdot\frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}}\right)}$$

and solving for $[EPc_jOS_i]$:

$$[EPc_kOS_i] = \frac{[E_0] \cdot \frac{[Pc_k]}{K_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}}}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[Pc_k]}{K_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}}\right)} \quad \textbf{...Eq. 65}$$

Substituting Eq. 65 into Eq. 64:

$$r_j = \frac{k_{f,j} \cdot [E_0] \cdot \frac{[Pc_k]}{K_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}}}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[Pc_k]}{K_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}}\right)}$$

Rearranging:

$$r_j = \frac{r_{MAX,j} \cdot [Pc_k] \cdot [OS_i]}{K_{d,k} \cdot K_{d,i} \cdot \left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[Pc_k]}{K_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}}\right)} \quad \textbf{...Eq. 66}$$

Where $r_{MAX,j} = k_{f,j} \cdot [E_0]$

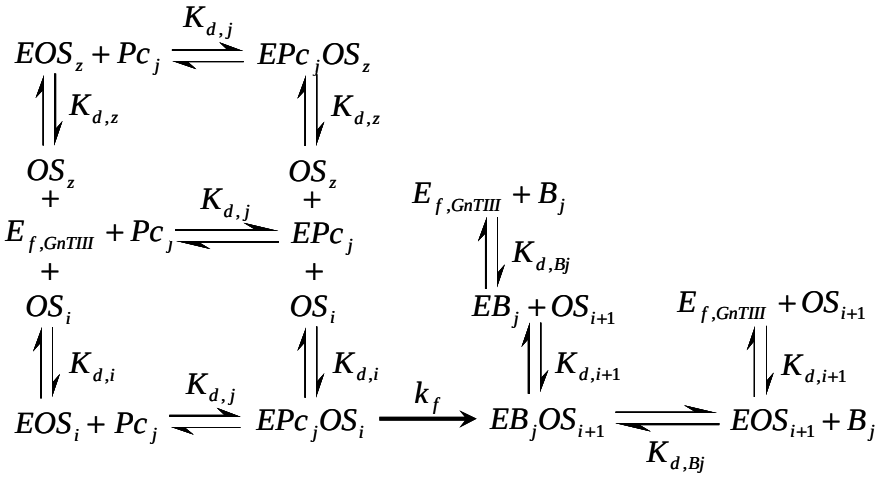
Random-order Bi-Bi kinetics

This rate equation derivation applies to the GnTIII, FucT and SiaT enzymes.[230, 232, 238] In this case, there is no conformation change in the enzyme, which means that both substrates can bind simultaneously to the enzyme. As with the previous cases, competitive inhibition will also be considered in this mechanism due to several substrates competing for the binding sites of the enzyme. This entails competitive inhibition, and when considering the random binding order, the resulting reaction mechanism is more complex than the previous two. Additional assumptions for the derivation of these kinetic expressions are the following:

a) Equilibrium is reached rapidly for all reactions.[232]
b) These enzymes act on several substrates along the N-glycosylation pathway. Therefore, competitive inhibition from analogue acceptor oligosaccharide (OS_z) chains will be assumed.
c) No disproportionate excess of one substrate with respect to the other is assumed within the Golgi compartments; hence, the reaction rate will depend on the concentrations of both species. (Enzyme is <u>not</u> assumed to be saturated with NSD).
d) Dual product inhibition will be considered for this reaction given that both the products have affinity for the enzyme.[232]
e) The rate-limiting step is assumed to be the transfer of the monosaccharide residue to the growing oligosaccharide chain.

f) Again, the rate-limiting step is assumed to be irreversible.[64, 65]

These assumptions lead to the following kinetic mechanism:



From this kinetic reaction mechanism we can derive the following mathematical expressions:

Dissociation constants:

$$K_{d,k} = \frac{[E_f]\cdot[Pc_k]}{[EPc_k]} \longrightarrow [EPc_k] = \frac{[E_f]\cdot[Pc_k]}{K_{d,k}} \quad \textbf{…Eq. 67}$$

$$K'_{d,k} = \frac{[EOS_i]\cdot[Pc_k]}{[EPc_kOS_i]} \longrightarrow [EPc_kOS_i] = \frac{[EOS_i]\cdot[Pc_k]}{K'_{d,k}} \quad \textbf{…Eq. 68}$$

$$K''_{d,k} = \frac{[EOS_z]\cdot[Pc_k]}{[EPc_kOS_z]} \longrightarrow [EPc_kOS_z] = \frac{[EOS_z]\cdot[Pc_k]}{K''_{d,k}} \quad \textbf{…Eq. 69}$$

$$K_{d,i} = \frac{[E_f]\cdot[OS_i]}{[EOS_i]} \longrightarrow [EOS_i] = \frac{[E_f]\cdot[OS_i]}{K_{d,i}} \quad \textbf{…Eq. 70}$$

$$K'_{d,i} = \frac{[EPc_k]\cdot[OS_i]}{[EPc_kOS_i]} \longrightarrow [EPc_kOS_i] = \frac{[EPc_k]\cdot[OS_i]}{K'_{d,i}} \quad \textbf{…Eq. 71}$$

$$K_{d,z} = \frac{[E_f]\cdot[OS_z]}{[EOS_z]} \longrightarrow [EOS_z] = \frac{[E_f]\cdot[OS_z]}{K_{d,z}} \quad \textbf{…Eq. 72}$$

$$K_{d,i+1} = \frac{[E_f]\cdot[OS_{i+1}]}{[EOS_{i+1}]} \longrightarrow [EOS_{i+1}] = \frac{[E_f]\cdot[OS_{i+1}]}{K_{d,i+1}} \quad \textbf{…Eq. 73}$$

$$K_{d,Bk} = \frac{[E_f]\cdot[B_k]}{[EB_k]} \longrightarrow [EB_k] = \frac{[E_f]\cdot[B_k]}{K_{d,Bk}} \quad \textbf{…Eq. 74}$$

$$K_{d,i} \cdot K'_{d,k} = \frac{[E_f] \cdot [OS_i]}{[EOS_i]} \cdot \frac{[EOS_i] \cdot [Pc_k]}{[EPc_k OS_i]} = \frac{[E_f] \cdot [Pc_k] \cdot [OS_i]}{[EPc_k OS_i]} \qquad \textbf{…Eq. 75}$$

$$K''_{d,k} \cdot K_{d,z} = \frac{[EOS_z] \cdot [Pc_k]}{[EPc_k OS_z]} \cdot \frac{[E_f] \cdot [OS_z]}{[EOS_z]} = \frac{[E_f] \cdot [Pc_k] \cdot [OS_z]}{[EPc_k OS_z]} \qquad \textbf{…Eq. 76}$$

$$K_{d,i+1} \cdot K_{d,Bk} = \frac{[EB_k] \cdot [OS_{i+1}]}{[EB_k OS_{i+1}]} \cdot \frac{[E_f] \cdot [B_k]}{[EB_k]} = \frac{[E_f] \cdot [B_k] \cdot [OS_{i+1}]}{[EB_k OS_{i+1}]} \qquad \textbf{…Eq. 77}$$

Total (or initial) enzyme balance:

$$[E_0] = [E_f] + [EPc_k] + [EOS_i] + \sum_{z=1}^{NC} [EOS_z] + [EPc_k OS_i] + \sum_{z=1}^{NC} [EPc_k OS_z] + [EB_k] + [EOS_{i+1}] + [EB_k OS_{i+1}]$$

…Eq. 78

Combining the dissociation constant expressions with the enzyme balance:

$$[E_0] = [E_f] + [E_f] \cdot \frac{[Pc_k]}{K_{d,k}} + [E_f] \cdot \frac{[OS_i]}{K_{d,i}} + [E_f] \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + [E_f] \cdot \frac{[Pc_k]}{K'_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + [E_f] \cdot \frac{[Pc_k]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} +$$
$$+ [E_f] \cdot \frac{[B_k]}{K_{d,Bk}} + [E_f] \cdot \frac{[OS_{i+1}]}{K_{d,i+1}} + [E_f] \cdot \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}}$$

…Eq. 79

Solving for the free enzyme term:

$$[E_f] = \frac{[E_0]}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[Pc_k]}{K'_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}} + \frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}}\right)}$$

…Eq. 80

The reaction rate expression for the rate-limiting reaction is defined as:

$$r_j = k_{f,j} \cdot [EPc_k OS_i] \qquad \textbf{…Eq. 81}$$

Combining Eq. 75 with Eq. 80:

$$[EPc_k OS_i] \cdot \frac{K'_{d,k}}{[Pc_k]} \cdot \frac{K_{d,i}}{[OS_i]} = \frac{[E_0]}{\left(1 + \frac{[Pc_k]}{K_{d,k}} + \frac{[OS_i]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[Pc_k]}{K'_{d,k}} \cdot \frac{[OS_i]}{K_{d,i}} + \frac{[Pc_k]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{[OS_z]}{K_{d,z}} + \frac{[B_k]}{K_{d,Bk}} + \frac{[OS_{i+1}]}{K_{d,i+1}} + \frac{[B_k]}{K_{d,Bk}} \cdot \frac{[OS_{i+1}]}{K_{d,i+1}}\right)}$$

and solving for $[EPc_j OS_i]$:

$$\left[EPc_j OS_i\right] = \frac{\left[E_0\right] \cdot \frac{\left[Pc_k\right]}{K'_{d,k}} \cdot \frac{\left[OS_i\right]}{K_{d,i}}}{\left(1 + \frac{\left[Pc_k\right]}{K_{d,k}} + \frac{\left[OS_i\right]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[Pc_k\right]}{K'_{d,k}} \cdot \frac{\left[OS_i\right]}{K_{d,i}} + \frac{\left[Pc_k\right]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[B_k\right]}{K_{d,Bk}} + \frac{\left[OS_{i+1}\right]}{K_{d,i+1}} + \frac{\left[B_k\right]}{K_{d,Bk}} \cdot \frac{\left[OS_{i+1}\right]}{K_{d,i+1}}\right)}$$

…Eq. 82

Substituting Eq. 82 into Eq. 81:

$$r_j = \frac{k_{f,j} \cdot \left[E_0\right] \cdot \frac{\left[Pc_k\right]}{K'_{d,k}} \cdot \frac{\left[OS_i\right]}{K_{d,i}}}{\left(1 + \frac{\left[Pc_k\right]}{K_{d,k}} + \frac{\left[OS_i\right]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[Pc_k\right]}{K'_{d,k}} \cdot \frac{\left[OS_i\right]}{K_{d,i}} + \frac{\left[Pc_k\right]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[B_k\right]}{K_{d,Bk}} + \frac{\left[OS_{i+1}\right]}{K_{d,i+1}} + \frac{\left[B_k\right]}{K_{d,Bk}} \cdot \frac{\left[OS_{i+1}\right]}{K_{d,i+1}}\right)}$$

Rearranging:

$$r_j = \frac{r_{MAX,j} \cdot \left[Pc_k\right] \cdot \left[OS_i\right]}{K'_{d,k} \cdot K_{d,i} \cdot \left(1 + \frac{\left[Pc_k\right]}{K_{d,k}} + \frac{\left[OS_i\right]}{K_{d,i}} + \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[Pc_k\right]}{K'_{d,k}} \cdot \frac{\left[OS_i\right]}{K_{d,i}} + \frac{\left[Pc_k\right]}{K''_{d,k}} \cdot \sum_{z=1}^{NC} \frac{\left[OS_z\right]}{K_{d,z}} + \frac{\left[B_k\right]}{K_{d,Bk}} + \frac{\left[OS_{i+1}\right]}{K_{d,i+1}} + \frac{\left[B_k\right]}{K_{d,Bk}} \cdot \frac{\left[OS_{i+1}\right]}{K_{d,i+1}}\right)}$$

…Eq. 83

Where $r_{MAX,j} = k_{f,j} \cdot \left[E_0\right]$

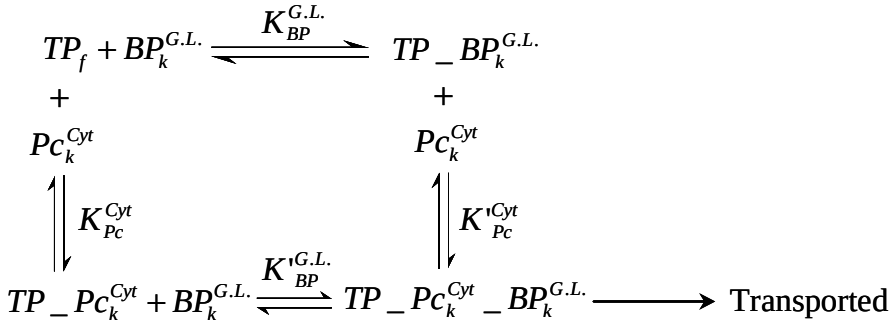
Appendix II

Transport equations derivation

In order for transport to occur, the NSD must bind to the transport protein on the cytosolic side of the membrane and the dephosphorylated by-product must bind to the transport protein on the luminal side of the Golgi membrane. Additionally, it has been reported that the process can also occur in the opposite sense.[239, 240] In this case, the NSD must bind to the transport protein on the luminal side while a dephosphorylated by-product must bind to the cytosolic side of the Golgi. Given this information, we have proposed a gradient that determines the transport as:

$$r_{Transp,k} = k_{T,k} \cdot \left(\left[TP_BP_k^{G.L.}_Pc_k^{Cyt} \right] - \left[TP_BP_k^{Cyt}_Pc_k^{G.L.} \right] \right) \qquad \textbf{...Eq. 84}$$

If binding of the NSD and the dephosphorylated by-product are considered to be analogous to substrate binding in enzymes, we obtain the following kinetic mechanism:



Using procedures analogous to enzyme derivations:

$$K_{Pc}^{Cyt} = \frac{\left[TP_f \right] \cdot \left[Pc_k^{Cyt} \right]}{\left[TP_Pc_k^{Cyt} \right]} \longrightarrow \left[TP_Pc_k^{Cyt} \right] = \frac{\left[TP_f \right] \cdot \left[Pc_k^{Cyt} \right]}{K_{Pc}^{Cyt}} \qquad \textbf{…Eq. 85}$$

$$K_{BP}^{G.L.} = \frac{\left[TP_f \right] \cdot \left[BP_k^{G.L.} \right]}{\left[TP_BP_k^{G.L.} \right]} \longrightarrow \left[TP_BP_k^{G.L.} \right] = \frac{\left[TP_f \right] \cdot \left[BP_k^{G.L.} \right]}{K_{BP}^{G.L.}} \qquad \textbf{…Eq. 86}$$

$$K'^{Cyt}_{Pc} = \frac{\left[TP_BP_k^{G.L.} \right] \cdot \left[Pc_k^{Cyt} \right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.} \right]} \longrightarrow \left[TP_Pc_k^{Cyt}_BP_k^{G.L.} \right] = \frac{\left[TP_BP_k^{G.L.} \right] \cdot \left[Pc_k^{Cyt} \right]}{K'^{Cyt}_{Pc}}$$

…Eq. 87

$$K'^{G.L.}_{BP} = \frac{\left[TP_Pc_k^{Cyt} \right] \cdot \left[BP_k^{G.L.} \right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.} \right]} \longrightarrow \left[TP_Pc_k^{Cyt}_BP_k^{G.L.} \right] = \frac{\left[TP_Pc_k^{Cyt} \right] \cdot \left[BP_k^{G.L.} \right]}{K'^{G.L.}_{BP}} \qquad \textbf{…Eq. 88}$$

$$K_{Pc}^{Cyt} \cdot K'^{G.L.}_{BP} = \frac{\left[TP_f\right]\cdot\left[Pc_k^{Cyt}\right]}{\left[TP_Pc_k^{Cyt}\right]} \cdot \frac{\left[TP_Pc_k^{Cyt}\right]\cdot\left[BP_k^{G.L.}\right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right]} = \frac{\left[TP_f\right]\cdot\left[Pc_k^{Cyt}\right]\cdot\left[BP_k^{G.L.}\right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right]} \quad \textbf{...Eq. 89}$$

$$K'^{Cyt}_{Pc} \cdot K_{BP}^{G.L.} = \frac{\left[TP_f\right]\cdot\left[BP_k^{G.L.}\right]}{\left[TP_BP_k^{G.L.}\right]} \cdot \frac{\left[TP_BP_k^{G.L.}\right]\cdot\left[Pc_k^{Cyt}\right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right]} = \frac{\left[TP_f\right]\cdot\left[Pc_k^{Cyt}\right]\cdot\left[BP_k^{G.L.}\right]}{\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right]} \quad \textbf{...Eq. 90}$$

Total (or initial) transport protein balance:

$$\left[TP_0\right] = \left[TP_f\right] + \left[TP_Pc_k^{Cyt}\right] + \left[TP_BP_k^{G.L.}\right] + \left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right] \quad \textbf{...Eq. 91}$$

Combining the dissociation constant expressions with the transport protein balance:

$$\left[TP_0\right] = \left[TP_f\right] + \left[TP_f\right]\cdot\frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}} + \left[TP_f\right]\cdot\frac{\left[BP_k^{G.L.}\right]}{K_{BP}^{G.L.}} + \left[TP_f\right]\cdot\frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}}\cdot\frac{\left[BP_k^{G.L.}\right]}{K'^{G.L.}_{BP}} \quad \textbf{...Eq. 92}$$

Solving for the free transport protein term:

$$\left[TP_f\right] = \frac{\left[TP_0\right]}{\left(1 + \frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}} + \frac{\left[BP_k^{G.L.}\right]}{K_{BP}^{G.L.}} + \frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}}\cdot\frac{\left[BP_k^{G.L.}\right]}{K'^{G.L.}_{BP}}\right)} \quad \textbf{...Eq. 93}$$

Combining equations 93 and 90 and solving for $\left[TP_Pc_j^{Cyt}_BP_j^{G.L.}\right]$:

$$\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right] = \frac{\left[TP_0\right]\cdot\left[Pc_k^{Cyt}\right]\cdot\left[BP_k^{G.L.}\right]}{K'^{Cyt}_{Pc}\cdot K_{BP}^{G.L.}\left(1 + \frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}} + \frac{\left[BP_k^{G.L.}\right]}{K_{BP}^{G.L.}} + \frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt}}\cdot\frac{\left[BP_k^{G.L.}\right]}{K'^{G.L.}_{BP}}\right)}$$

Simplifying:

$$\left[TP_Pc_k^{Cyt}_BP_k^{G.L.}\right] = \left[TP_0\right]\cdot\left(\frac{\left[Pc_k^{Cyt}\right]}{K_{Pc}^{Cyt} + \left[Pc_k^{Cyt}\right]}\right)\cdot\left(\frac{\left[BP_k^{G.L.}\right]}{K_{BP}^{G.L.} + \left[BP_k^{G.L.}\right]}\right) \quad \textbf{...Eq. 94}$$

Considering that the same derivation can be followed for transport in the opposite direction:

$$\left[TP_Pc_k^{G.L.}_BP_k^{Cyt}\right] = \left[TP_0\right]\cdot\left(\frac{\left[Pc_k^{G.L.}\right]}{K_{Pc}^{G.L.} + \left[Pc_k^{G.L.}\right]}\right)\cdot\left(\frac{\left[BP_k^{Cyt}\right]}{K_{BP}^{Cyt} + \left[BP_k^{Cyt}\right]}\right) \quad \textbf{...Eq. 95}$$

Substituting equations 94 and 95 into equation 84:

$$r_{Transp,k} = k_{T,k} \cdot [TP_0] \cdot \left[\left(\frac{\left[Pc_k^{Cyt} \right]}{K_{Pc}^{Cyt} + \left[Pc_k^{Cyt} \right]} \right) \cdot \left(\frac{\left[BP_k^{G.L.} \right]}{K_{BP}^{G.L.} + \left[BP_k^{G.L.} \right]} \right) - \left(\frac{\left[Pc_k^{G.L.} \right]}{K_{Pc}^{G.L.} + \left[Pc_k^{G.L.} \right]} \right) \cdot \left(\frac{\left[BP_k^{Cyt} \right]}{K_{BP}^{Cyt} + \left[BP_k^{Cyt} \right]} \right) \right]$$

…Eq. 96

Finally, if we consider that by-product accumulation in the cytosol is limited given the dilution effect of going from a limited volume (Golgi cisternae) to a larger one (cytosol) and that the uptake of dephosphorylated nucleotides occurs readily in the cytosol through the nucleotide salvage pathways,[298] the second term of equation 96 can be neglected so the final expression obtained is:

$$r_{Transp,k} = k_{T,k} \cdot [TP_0] \cdot \left[\left(\frac{\left[Pc_k^{Cyt} \right]}{K_{Pc}^{Cyt} + \left[Pc_k^{Cyt} \right]} \right) \cdot \left(\frac{\left[BP_k^{G.L.} \right]}{K_{BP}^{G.L.} + \left[BP_k^{G.L.} \right]} \right) \right] \qquad \textbf{...Eq. 97}$$

Appendix III

Glycosylation model parameter derivation

Enzyme turnover rates

The turnover rates have been estimated from the literature similarly to previous modelling studies. The maximum rate has usually been expressed in units of μmol mg_{prot}^{-1} min^{-1}. In order to obtain the turnover rate in units of μmol_{prod} μmol_{enz}^{-1} min^{-1}, the reported V_{max} has been multiplied by the molecular weight of the enzyme. This is valid only if the kinetic analysis has been performed with purified enzyme.

Transport protein turnover rates

Gilchrist et al.[267] performed a quantitative proteomic analysis of the secretory pathway in rat hepatocytes. In their study, the only NSD transport protein detected in the Golgi fraction is the UDP-GlcNAc transporter and was reported to account for 0.005116 % of all Golgi resident proteins. With this value, an average molecular weight of 60,600 g mol^{-1} (also estimated from [267]) and the maximum transport rate for TP_1 shown in Table 8Table 9, a maximum turnover rate for this transport protein was estimated to be 1,084 min^{-1}. The remaining maximum transport turnover rates were derived from this value by assuming that the transport turnover rates are proportional to the reaction turnover rates.

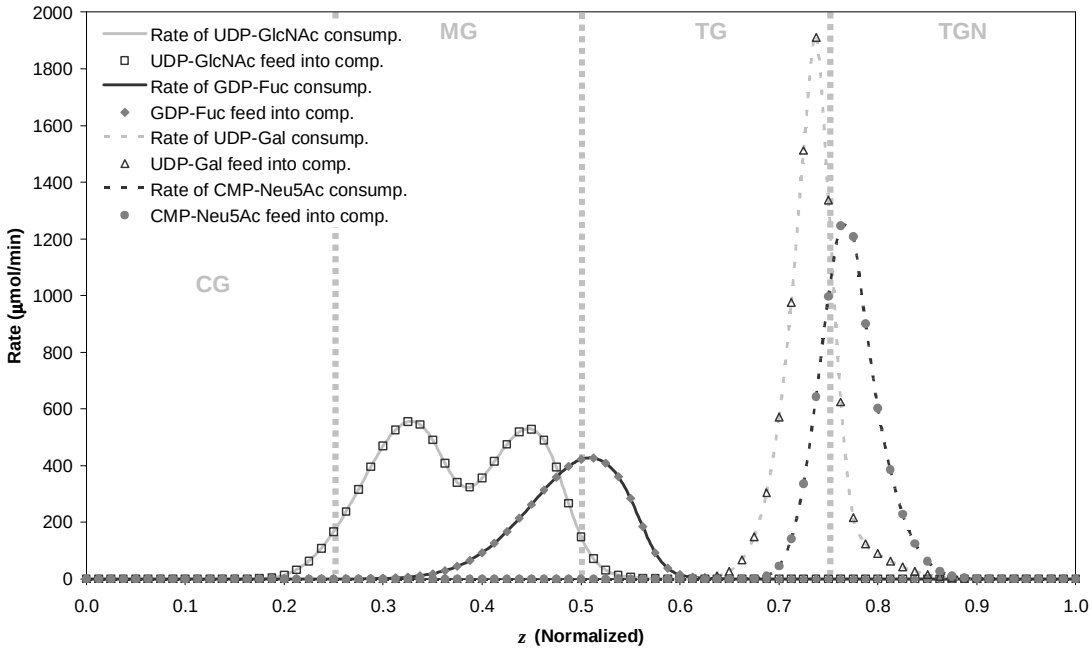
NSD concentration in cytosol and Golgi cisternae

The concentration of NSDs has been reported for overall cell volume and thus, their distribution between the Golgi apparatus and the cytosol is unknown. It has previously been reported that TPs have the ability to concentrate them up to 50-fold with respect to extraluminal concentrations.[239] A 40-fold increase in NSD concentration was assumed (similar to what Krambeck and Betenbaugh[32] assumed) and a material balance was performed to find the distribution of these components on the luminal and cytosolic side of Golgi cisternae.

Appendix IV

Rate of NSD consumption vs. rate of NSD transport into Golgi

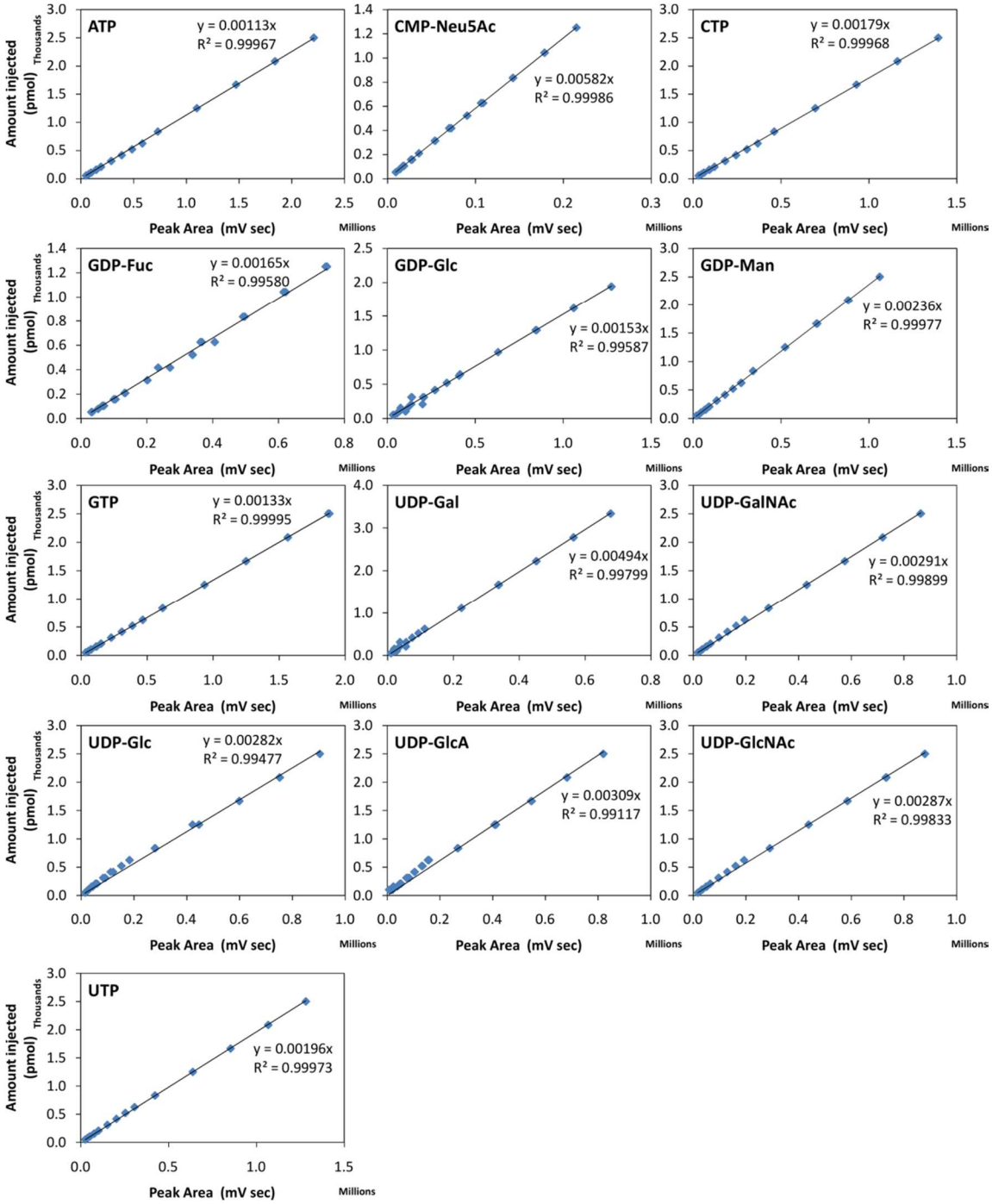
The NSD transport parameters were estimated with the methodology presented in the *parameter estimation* section of the main paper. With the estimated parameters, simulations were performed in order to assess the performance of the estimated parameters. In the figure presented below, the profiles of NSD consumption and transport over the length of Golgi are shown. The lines represent the rate of NSD consumption and the symbols represent the rate of transport for each NSD into the Golgi along the length of the Golgi apparatus. The vertical dotted lines mark the zones where the maturing compartments are defined as either cis, medial, trans or trans Golgi network.



In the figure presented above, it is clear that the rate of NSD consumption overlaps with the NSD transport rate indicating that the obtained parameters achieve the conditions defined in the optimization methodology. The implication behind both sets of data overlapping is that the transport of NSDs is not limiting. However, when the concentration of NSDs on the cytosolic side of the Golgi membranes is limited due to metabolic effects, transport is hindered (see equation 97 for NSD transport) and, only then would NSD availability inside Golgi limit the glycosylation reactions.

Appendix V

HPAEC Supporting information



Calibration curves for 4 nucleotides and 9 NSDs
These curves were generated by plotting the measured area below the peak vs. the known amount of analyte injected. The equation for the linear regression and the corresponding R^2 are also presented in the figure.