Karen Polizzi ORCID iD: 0000-0001-5435-2667

Ioscani Jimenez del Val ORCID iD: 0000-0001-9289-0191

Cleo Kontoravdi ORCID iD: 0000-0003-0213-4830

Model-based optimisation of antibody galactosylation in CHO cell culture

Short title: Model-based optimisation of antibody galactosylation

Pavlos Kotidis[a†], Philip Jedrzejewski[a,b,c†], Si Nga Sou[a,b,c,1], Christopher Sellick[d,2], Karen Polizzi[b,c], Ioscani Jimenez del Val[e], Cleo Kontoravdi[a*]

[a]Department of Chemical Engineering, Imperial College London, United Kingdom

[b]Department of Life Sciences, Imperial College London, United Kingdom

[c]Centre for Synthetic Biology and Innovation, Imperial College London, United Kingdom

[d]MedImmune, Granta Park, Cambridge, United Kingdom

[e]School of Chemical & Bioprocess Engineering, University College Dublin, Ireland

[†]Co-first author

*Corresponding author

Email: cleo.kontoravdi@imperial.ac.uk

[1] Present address: MedImmune, Granta Park, Cambridge, United Kingdom

[2] Present address: Kymab Ltd., Babraham Research Campus, Cambridge, United Kingdom

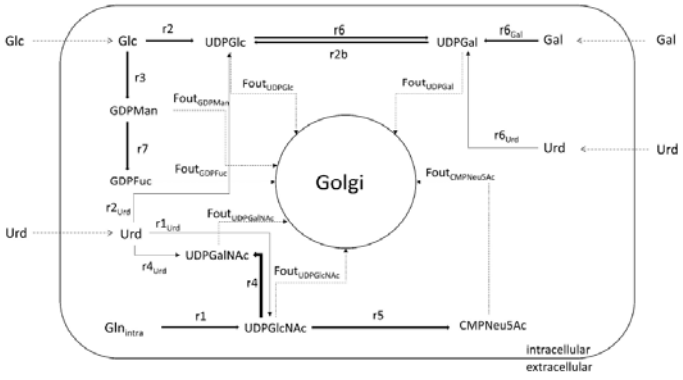
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bit.26960.

Abstract

Exerting control over the glycan moieties of antibody therapeutics is highly desirable from a product safety and batch-to-batch consistency perspective. Strategies to improve antibody productivity may compromise quality, while interventions for improving glycoform distribution can adversely affect cell growth and productivity. Process design therefore needs to consider the trade-off between preserving cellular health and productivity while enhancing antibody quality. In this work, we present a modelling platform that quantifies the impact of glycosylation precursor feeding – specifically, that of galactose and uridine – on cellular growth, metabolism as well as antibody productivity and glycoform distribution. The platform has been parameterised using an initial training data set yielding an accuracy of ±5% with respect to glycoform distribution. It was then used to design an optimised feeding strategy that enhances the final concentration of galactosylated antibody in the supernatant by over 90% compared to the control without compromising the integral of viable cell density or final antibody titre. This work supports the implementation of Quality by Design towards higher-performing bioprocesses.

Graphical Abstract

Exerting control over the glycan moieties of antibody therapeutics is highly desirable from a product safety and batch-to-batch consistency perspective. Strategies to improve antibody productivity may compromise quality, while interventions for improving glycoform distribution can adversely affect cell growth and productivity.



Keywords: antibody glycosylation, galactosylation, Chinese hamster ovary cells, mathematical modelling, nucleotide sugars, process optimisation

Introduction

Monoclonal antibodies (mAbs) are the most successful biopharmaceutical products of the last decade (Walsh 2014, Urquhart 2018). Chinese hamster ovary (CHO) cells are the workhorse of recombinant protein production in the biopharmaceutical industry, owing to their ability to perform complex post-translational modifications and secrete proteins that are compatible with and bioactive in humans (Lai, Yang et al. 2013, Yusufi, Lakshmanan et al. 2017). CHO cells are usually grown in fed-batch cultures to replenish the depleted nutrients, prolong the cell culture period and maximise protein productivity.

Most commercial mAbs are N-linked glycosylated in their constant fragment (Fc). Earlier studies have shown that N-linked glycosylation is strongly dependent on bioprocess conditions, such as temperature (Sou, Sellick et al. 2015), nutrient availability (Liu, Spearman et al. 2014) and catabolite accumulation (Yang and Butler 2000). Bioprocess conditions affect the monosaccharide composition and relative abundance of the oligosaccharides, which in turn, can alter the safety and potency of the mAb. The presence of terminal galactose residues on mAb Fc glycans is known to increase complement-dependent cytotoxicity (Hodoniczky, Zheng et al. 2005) and antibody-dependent cytotoxicity (Thomann, Reckermann et al. 2016) and is, therefore, a desirable attribute in oncological products. In addition, galactosylation is a major source of mAb structural variability during scale-up and technology transfer (Gramer, Eckblad et al. 2011). Manufacturing variability can therefore affect product structure consistency and potentially impact efficacy (Shinkawa, Nakamura et al. 2003, Goetze, Liu et al. 2011, Yu, Brown et al. 2012). Several efforts have been undertaken to increase galactosylation mainly by feeding galactose and uridine, the

metabolic precursors of uridine diphosphate galactose (UDPGal), to the cell culture (Wong, Wati et al. 2010, Gramer, Eckblad et al. 2011, Grainger and James 2013). UDP-Gal is the nucleotide sugar donor (NSD) required for galactosylation as the co-substrate of the β-1,4 galactosyltransferase (GalT). However, the benefits of precursor addition towards galactosylation can be offset by metabolic perturbations, compromised cell growth and reduced product titre (Grainger and James 2013).

Mathematical models are a useful tool that has been widely used to describe cell metabolism (Kontoravdi, N. Pistikopoulos et al. 2010, Nolan and Lee 2011) and N-linked glycosylation (Krambeck and Betenbaugh 2005, del Val, Nagy et al. 2011, Spahn, Hansen et al. 2016, Kontoravdi and del Val 2018). Metabolic flux analysis (MFA) and flux balance analysis (FBA) are usually conducted to unveil the most active metabolic pathways in cells during culture. Glycosylation flux analysis (GFA) has also been proposed to calculate the intracellular fluxes and distributions of glycans (Hutter, Villiger et al. 2017, Kremkow and Lee 2018). In an effort to provide a holistic model for mAb production, Jedrzejewski, del Val et al. (2014) linked the extracellular concentrations of nutrients to glycosylation, using a model framework that included the nucleotide and NSD synthesis pathways as the channel between cell metabolism and glycosylation. The proposed concept opened a new path for monitoring mAb production and glycan distribution through mathematical models (Kaveh, Hengameh et al. 2013).

Following the Quality by Design (QbD) paradigm, where the end-product quality is rationally built into the manufacturing process, a mathematical model that could describe the perturbations affected by different feeding strategies on cell metabolism and mAb glycosylation would be a valuable tool for process and design space

optimization. In order to bridge this gap, herein we propose a combined experimental and computational strategy that captures the effects of feeding strategies on cell metabolism, intracellular NSD concentrations and mAb glycan distribution. The proposed mechanistic model is able to capture the metabolic alterations caused by the addition of uridine and galactose at different concentrations. The model has been trained with a set of experimental data, and its predictive capabilities have been validated against an independent feeding experiment that was designed using model-based optimisation. The proposed model sets the groundwork for *in silico* glycan optimisation by means of metabolic control, speeding up process development and reducing glycoform variability.

2. Materials and Methods

2.1 Cell culture

A Chinese hamster ovary CHO cell line (kindly donated by MedImmune, Cambridge, UK) producing an IgG antibody was used in this study. The cells were maintained in suspension culture in CD CHO medium (Life Technologies, Paisley, UK) at 36.5 °C and 5% CO_2 on an orbital shaking platform rotating at 150 rpm and were passaged every 3 days. 50 μM methionine sulfoximine (MSX) was added for the first two passages post-revival. Experiments were conducted in 500 mL vented Erlenmeyer flasks with a working volume of 100 mL at a seeding density of 2×10^5 cells·mL^{-1}. Cell concentration was determined using a Neubauer ruling haemocytometer, and viability was estimated by the trypan blue dye exclusion method using light microscopy. All cultures were supplemented with 10% v/v CD EfficientFeedC™ C AGT™ Nutrient supplement (Life Technologies, Paisley, UK) on all even culture days starting on day 2. Cell cultures were supplemented with 1 μM manganese(II)

chloride solution at seeding as well as varying amounts of D-(+)-galactose (6.5 mM to 50 mM in total culture volume) and uridine (0.76 mM to 20 mM in total culture volume) (all Sigma-Aldrich, Dorset, UK). Cultures were terminated at day 12 or when cell viability dropped below 60%. In total, six different feeding strategies were followed: (i) no galactose and uridine addition (control experiment), (ii) addition of 10mM galactose on days 4 and 8 (10G), (iii) addition of 10mM galactose and 5mM uridine on days 4 and 8 (10G5U), (iv) addition of 10mM galactose and 20mM uridine on days 4 and 8 (10G20U), (v) addition of 50mM galactose and 5mM uridine on days 4 and 8 (50G5U) and (vi) addition of 6.5mM galactose and 0.76mM uridine on day 4, addition of 9.3mM galactose and 1.34mM uridine on day 6, addition of 9mM galactose and 2.8mM uridine on day 8 and addition of 8.7mM galactose and 10mM uridine on day 10 (independent experiment).

2.2 Metabolite and antibody quantification

Antibody concentration in cell culture supernatant was determined using the BLItz® system and the Dip and Read™ Protein A (ProA) Biosensors (all Pall ForteBio, Portsmouth, UK). Extracellular glucose, glutamine, glutamate, ammonia and lactate concentrations were measured using the BioProfile 400 (NOVA Biomedical, MA, USA). Extracellular galactose concentration was determined using the Amplex® Red Galactose/Galactose Oxidase Assay kit (Life Technologies, Paisley, UK) according to the manufacturer's instructions. Intracellular glutamine concentration was determined using the Glutamine EnzyChrom™ Assay Kit (Universal Biologicals, Cambridge, UK). Extracellular uridine concentrations were measured using an optimized high-performance anion-exchange chromatography (HPAEC) method as described in del Val, Kyriakopoulos et al. (2013).

2.3 Nucleotide and NSD quantification

Cell culture samples containing 2×10^6 viable cells were centrifuged at 800 rpm. In a washing step the resulting cell pellet was resuspended in 2 mL of ice cold 0.9% w/v sodium chloride and then immediately centrifuged at 800 rpm to remove supernatant residues. The cell pellet was resuspended in 400 µL of ice-cold 50% v/v aqueous acetonitrile (Sigma-Aldrich, Gillingham, UK). The solution was incubated on ice for 10 minutes before centrifugation (4 °C, 10,000× g, 5 min). The nucleotide and nucleotide sugar containing supernatant was dried using a SpeedVac (Savant Inc. Laboratory, MI) and resuspended in 150 µL water. The final supernatant sample was filtered using a 0.22 µm syringe filter and stored at -80 °C prior to analysis.

The HPAEC method described in del Val, Kyriakopoulos et al. (2013) was also used for intracellular nucleotide, nucleoside and nucleotide sugar quantification. Briefly, elutions were performed with gradients of 1.5 M sodium acetate solution (Sigma-Aldrich, Gillingham, UK) in 3 mM potassium hydroxide (Sigma-Aldrich, Gillingham, UK) with a maximum ion concentration of 1 M sodium acetate using a CarboPac PA1 column (Dionex, Bannockburn, IL, USA). A modified method was used for the quantification of GDP-Fuc and GDP-Man to prevent species co-elution with intracellular tryptophan. Elutions were performed with gradients of 1.5 M sodium acetate solution in 3 mM potassium hydroxide. Elutions of the sodium acetate solution were carried out using the following gradients: t_0 min = 20%, t_1 min = 66%, t_6 min = 66%, t_{23} min = 20%, t_{30} min = 20%. Detection was performed at 262.1 nm.

2.4 Antibody glycan analysis

Secreted antibody was purified from cell culture supernatant with a Protein A IgG Purification Kit (Thermo Scientific, Horsham, UK) and concentrated using Vivaspin® 2 ultracentrifugation spin columns 50kDa (Generon, Maidenhead, UK). Further glycan processing and glycan quantification was kindly performed by MedImmune (Cambridge, UK). The workflow involved the denaturation of purified and concentrated IgG samples, PNGase digestion and glycan labelling. Separation/quantification was performed using the LabChip® platform (PerkinElmer, Seer Green, UK) based on capillary electrophoresis.

2.5 Mathematical Modelling

Model construction, simulation, parameter estimation and optimization were conducted in gPROMS v.5.0.1 (Process System Enterprise Ltd, London, U.K., www.psenterprise.com/gproms). Unless otherwise stated, the model parameters were estimated using the maximum likelihood optimization formulation in gPROMS that maximises the probability that the mathematical model will predict the experimental measurements by assigning values to the physical and variance model parameters. The same solver was used to design the independent feeding strategy used for validation. The model framework comprises three sub-models: an unstructured cell growth, death and metabolism model, a model describing the NSD synthesis and an N-linked glycosylation model (del Val, Nagy et al. 2011, del Val, Fan et al. 2016). The cell culture model estimates the specific cell growth and specific mAb productivity that are used as inputs for the NSD synthesis model. The latter calculates the NSD concentration and the fluxes of NSDs from the cytosol to the Golgi

environment, which are fed to the N-linked glycosylation model to estimate the distribution of mAb glycoforms.

2.6 Statistical Analysis

The MedCalc Statistical Software v.18.2.1 (MedCalc Software bvba, Ostend, Belgium) was used for the estimation of the Concordance Correlation Coefficient (CCC) (Lawrence 1989) for a 95% confidence interval, to quantify the degree of agreement of the model framework with experimental results, alongside with the R^2. CCC and R^2 were used to quantify the agreement of two sets of variables, which in our case were the observed and the modelling results. The Pearson Correlation Coefficient (PCC) was additionally calculated and presented for completeness, even though it is limited to quantifying the correlation of two variable sets and lacks the ability to describe their agreement. OriginPro 2017 (OriginLab, Northampton, MA) was used for the calculation of PCC for a 95% confidence interval.

3. Model construction and parameter estimation

3.1 Cell culture model

The cell culture model framework was adapted from previous work on fed-batch hybridoma cell cultures (Kontoravdi, N. Pistikopoulos et al. 2010). The model is comprised of three modules: a CHO cell growth and death model, a CHO cell metabolism model and a mAb production model. The control, 10G, 10G5U, 10G20U and 50G5U experiments were used for model construction and parameter estimation.

3.1.1 CHO cell growth and death

CHO cell growth description was based on modified Monod equations that were defined by the limiting substrates and inhibiting metabolites. The equations of the model framework that describe CHO cell growth and death are listed below:

$$\frac{dV}{dt} = F_{in} - F_{out} \quad (1)$$

$$\frac{d(VX_v)}{dt} = (\mu - \mu_{death})VX_v - F_{out}X_v \quad (2)$$

Eq.(1) describes the material balance for the cell culture volume V (L) according to the inlet F_{in} (L·h^{-1}) and the outlet F_{out} (L·h^{-1}) flow rates. Eq. (2) describes the material balance for the cell density X_v (cell·L^{-1}), where μ (h^{-1}) and μ_{death} (h^{-1}) indicate the specific cell growth and death rate, respectively.

$$\mu = \mu_{max} f_{\text{lim}} f_{inh} \quad (3)$$

$$f_{\text{lim}} = \frac{[Glc]}{[Glc]+K_{Glc}} \frac{[Asn]}{[Asn]+K_{Asn}} \quad (4)$$

$$f_{\text{inh}} = \frac{KI_{Amm}}{[Amm]+KI_{Amm}} \frac{KI_{Lac}}{[Lac]+KI_{Lac}} \frac{KI_{Urd}}{[Urd]+KI_{Urd}} \quad (5)$$

Eq. (3), (4) and (5) describe the specific cell growth rate, where μ_{max} (h^{-1}) is the maximum specific cell growth rate, f_{lim} (-) is the substrate limiting factor and f_{inh} (-) is the metabolite inhibiting factor. Glucose and asparagine were regarded as the limiting substrates, while lactate, ammonia and uridine were considered to inhibit growth. The latter two were also regarded as toxic for cell growth. K_{Glc} (mM) and K_{Asn}(mM) are the Monod constants for glucose and asparagine, respectively. KI_{Amm} (mM), KI_{Lac} (mM), KI_{Urd} (mM) are the inhibiting constants for the ammonia, lactate

and uridine, respectively. $[Glc]$, $[Asn]$, $[Amm]$, $[Lac]$ and $[Urd]$ are the extracellular concentrations (mM) of glucose, asparagine, ammonia, lactate and uridine, respectively.

$$\mu_{death} = \mu_{death,max}\left(\frac{[Amm]}{[Amm]+K_{d,Amm}} + \frac{[Urd]}{[Urd]+K_{d,Urd}}\right) \quad (6)$$

Eq. (6) describes the specific cell death rate, were $\mu_{death,max}$ (h^{-1}) is the maximum specific cell death rate and $K_{d,Amm}$ (mM) and $K_{d,Urd}$ (mM) are the ammonia and uridine constants for cell death, respectively.

3.1.2 CHO cell metabolism

Cell metabolism was described by the material balances of substrates, products and amino acids that were calculated according to Eq. (7):

$$\frac{d(V[Metabolite])}{dt} = F_{in}\left[Metabolite_{feed}\right] - F_{out}[Metabolite] + q_{metabolite}VX_v \quad (7)$$

where, $\left[Metabolite_{feed}\right]$ (mM) is the concentration of each metabolite in the feed, $[Metabolite]$ (mM) is the extracellular concentration of each metabolite in the culture and $q_{metabolite}$ ($mmol \cdot cell^{-1} \cdot h^{-1}$) is the specific production/consumption rate of each metabolite and is based on the intracellular metabolic links between the metabolites as discussed below. The sign convention used in the model for the specific rates is positive (+) for the secreted and negative (-) for the consumed metabolites, respectively. The specific production/consumption rates for each metabolite are listed below. In Eq. (8)-(16), $Y_{X_{Metabolite}}$ ($cell \cdot mmol^{-1}$) describes the yield of cell biomass on the metabolite and is assumed to remain constant for each metabolite during the cell culture process to avoid overparameterization.

$$q_{Glc} = (-\frac{\mu}{Y_{X_{Glc}}} - m_{Glc}) \cdot (\frac{Kc_{Gal}}{Kc_{Gal}+[Gal]})^{n_{Gal}} \quad (8)$$

$$n_{Gal} = 1 - f_{Gal}\frac{q_{Gal}}{q_{Glc}} \quad (9)$$

Eq. (8) and (9) describe the specific consumption rate of glucose. Moreover, the specific consumption rate of glucose was found to depend on the respective galactose rate, as described in Eq. (9). The link between the two metabolites is expected as galactose feeds into the glycolytic pathway through its enzymatic conversion into glucose 1-phosphate. The $\frac{Kc_{Gal}}{Kc_{Gal}+[Gal]}$ term in Eq. (8) is used as a regulator when galactose is fed in the culture. In the absence of galactose the above term is equal to 1 and does not affect the glucose specific rate. m_{Glc} (mmol·cell^{-1}·h^{-1}) is the maintenance coefficient of glucose for other metabolic pathways of the cell, Kc_{Gal} (mM) is the regulating concentration of galactose and n_{Gal} (-) and f_{Gal} (-) are factors that participate in regulating the specific uptake rate of glucose. n_{Gal} accounts for the fraction of the galactose specific consumption rate and the respective glucose consumption rate. In Eq. (9) when the galactose consumption rate increases, n_{Gal} decreases and consequently from Eq. (8) q_{Glc} increases. The majority of glucose transporters (GLUTs) in mammalian cells show affinity to both glucose and galactose and therefore a regulation of the glucose and galactose flux from the GLUTs is possible (Wieczorke, Dlugai et al. 2003, Zhao and Keating 2007).

$$q_{Gln} = \frac{\mu}{Y_{X_{Gln}}} + q_{Amm}Y_{Gln/Amm} \quad (10)$$

where, $Y_{Gln/Amm}$ ($mmol_{Gln}$·$mmol_{Amm}^{-1}$) is the yield of glutamine from ammonia.

$$q_{Lac} = \left(\frac{\mu}{Y_{X_{Lac}}} - Y_{Lac/Glc}q_{Glc}\right)\frac{(Lac_{max1}-[Lac])}{Lac_{max1}} + m_{lac}\frac{Lac_{max2}-[Lac]}{Lac_{max2}} \quad (11)$$

where, $Y_{Lac/Glc}$ ($mmol_{Lac} \cdot mmol_{Glu}^{-1}$) is the yield of lactate from glucose, Lac_{max1} (mM) and Lac_{max2} (mM) are kinetic constants for the lactate consumption that occurs during the stationary phase (Zagari, Jordan et al. 2013) and m_{lac} ($mmol \cdot cell^{-1} \cdot h^{-1}$) is the maintenance coefficient of lactate for other metabolic pathways of the cell.

$$q_{Amm} = \frac{\mu}{Y_{X_{Amm}}} - Y_{Amm/Urd}\, q_{Urd} \quad (12)$$

where, $Y_{Amm/Urd}$ ($mmol_{Amm} \cdot mmol_{Urd}^{-1}$) is the yield of ammonia from uridine.

$$q_{Glu} = -\frac{\mu}{Y_{X_{Glu}}} \quad (13)$$

$$q_{Gal} = -\frac{\mu}{Y_{X_{Gal}}} \frac{[Gal]}{[Gal]+K_{Gal}} \quad (14)$$

where, K_{Gal} (mM) is a Monod constant. The $\frac{[Gal]}{[Gal]+K_{Gal}}$ term functions as a regulator of the specific consumption rate of galactose.

$$q_{Urd} = \frac{\mu}{Y_{X_{Urd}}} \frac{[Urd]}{[Urd]+K_{Urd}} \quad (15)$$

where, K_{Urd} (mM) is a Monod constant that regulates the effect uridine concentration on the specific consumption rate of uridine, with the same way that was previously described for galactose.

$$q_{Asn} = -\frac{\mu}{Y_{X_{Asn}}} - Y_{Asn/Asp} q_{Asp} \quad (16)$$

where, $Y_{Asn/Asp}$ ($mmol_{Asn} \cdot mmol_{Asp}^{-1}$) is the yield of asparagine from aspartate.

$$q_{Asp} = -\frac{\mu}{Y_{X_{Asp}}} - Y_{Asp/Asn} q_{Asn} \quad (17)$$

where, $Y_{Asp/Asn}$ ($mmol_{Asp} \cdot mmol_{Asn}^{-1}$) is the yield of aspartate from asparagine.

3.1.3 mAb production

mAb synthesis was described with the material balance of the extracellular mAb concentration and the respective specific production rate:

$$\frac{d(V[mAb])}{dt} = -F_{out}[mAb] + q_{mAb}VX_v \quad (18)$$

$$q_{mAb} = Y_{mAb_X}\mu + m_{mAb} \quad (19)$$

where, $[mAb]$ (pg·L^{-1}) is the extracellular mAb concentration, q_{mAb} (pg·cell^{-1}·h^{-1}) is the specific productivity of mAbs, Y_{mAb_X} (pg·cell^{-1}) is the yield of mAbs from cell growth and m_{mAb} (pg·cell^{-1}·h^{-1}) is a non-growth associated term.

3.2 NSD metabolic model

The NSD metabolic model uses the specific production rate of mAbs and the specific cell growth rate as inputs. Throughout the NSD model, the nucleotides are assumed to be in excess, as supported by our experimental measurements (data not shown) and therefore were not included in the rate calculations. The rates are based on simple Michaelis-Menten saturation kinetics and the substrates used are the intracellular NSD concentrations, the extracellular concentrations of glucose, galactose and uridine and the intracellular concentration of glutamine. Apart from glutamine, the intracellular concentrations of the metabolites were not estimated in order to avoid oversimplifications and overparameterization. The intracellular concentration of glutamine was linked to the respective extracellular through this simple equation:

$$[Gln_{intra}] = f_{Gln}[Gln] \quad (20)$$

where, $[Gln_{intra}]$ (mM) is the intracellular concentration of glutamine and f_{Gln} is the secretion factor of glutamine to the extracellular environment.

The rates were calculated according to the proposed NSD metabolic pathway (Fig. 1) and the kinetic mechanisms that have been reported for each enzyme (Kanehisa and Goto 2000, Schomburg, Chang et al. 2002).

The rates ($mmol_{NSD} \cdot L_{cell}^{-1} \cdot h^{-1}$) as shown in Fig. 1 are listed below:

$$r_1 = V_{max1} \frac{[Gln_{intra}]}{K_{M1_{Gln}} + [Gln_{intra}]} \quad (22)$$

$$r_{1_{sink}} = V_{max1_{sink}} \frac{[UDPGlcNAc]}{\left(K_{M1_{sink}} + [UDPGlcNAc]\right)\left(1 + \frac{[CMPNeu5Ac]}{KI_{1_{sink}}}\right)} \quad (23)$$

$$r_2 = V_{max2} \frac{[Glc]}{K_{M2_{Glc}} + [Glc]} \quad (24)$$

$$r_{2b} = V_{max2b} \frac{[UDPGal]}{K_{M2b_{UDPGal}}\left(1 + \frac{[UDPGlcNAc]}{KI_{2A}} + \frac{[UDPGalNAc]}{KI_{2B}} + \frac{[UDPGlc]}{KI_{2C}} + \frac{[UDPGal]}{KI_{2D}}\right) + [UDPGal]} \quad (25)$$

$$r_3 = V_{max3} \frac{[Glc]}{K_{M3_{Glc}} + [Glc]} \quad (26)$$

$$r_4 = V_{max4} \frac{[UDPGlcNAc]}{K_{M4_{UDPGlcNAc}} + [UDPGlcNAc]} \quad (27)$$

$$r_5 = V_{max5} \frac{[UDPGlcNAc]}{K_{M5_{UDPGlcNAc}}\left(1 + \frac{[CMPNeu5Ac]}{KI_5}\right) + [UDPGlcNAc]} \quad (28)$$

$$r_6 = V_{max6} \frac{[UDPGlc]}{K_{M6_{UDPGlc}}\left(1 + \frac{[UDPGlcNAc]}{KI_{6A}} + \frac{[UDPGalNAc]}{KI_{6B}} + \frac{[UDPGal]}{KI_{6C}}\right) + [UDPGlc]} \quad (29)$$

$$r_{6_{sink}} = V_{max6_{sink}} \frac{[UDPGal]}{K_{6_{sink}}\left(1 + \frac{[UDPGlc]}{KI_{6_{sink}}}\right) + [UDPGal]} \frac{[Gal]}{[Gal] + K_{regulator}} \quad (30)$$

$$r_7 = V_{max7} \frac{[GDPMan]}{(K_{M7_{GDPMan}} + [GDPMan])\left(1 + \frac{[GDPFuc]}{KI_7}\right)} \quad (31)$$

$$r_{7_{sink}} = V_{max7_{sink}} \frac{[GDPFuc]}{K_{M7_{sink}} + [GDPFuc]} \quad (32)$$

where, V_{max}, i ($mmol_{NSD} \cdot L_{cell}^{-1} \cdot h^{-1}$) is the maximum turnover rate of reaction i, $[NSD]$ (mM) are the cytosolic NSD concentrations, $K_{Mi_{NSD}}$ (mM) are the saturation constants of the NSD in the i reaction, KI_i (mM) is the inhibition constant in the i reaction and $K_{regulator}$ (mM) is a fixed parameter that deactivates the regulating $\frac{[Gal]}{[Gal] + K_{regulator}}$ term when no galactose is fed. The $r_{i_{sink}}$ terms (not shown in Fig. 1) describe the flux of NSD used in reactions that are not included in the chosen metabolic framework and were found to affect GDPFuc, UDPGal and UDPGlcNAc concentration.

The contribution of uridine and galactose to NSDs concentration was described with separate reaction rates ($mmol_{NSD} \cdot L_{cell}^{-1} \cdot h^{-1}$) that are described below:

$$r_{i_{Urd}} = V_{max,i_{Urd}} \frac{[Urd]}{K_{M,i_{Urd}} + [Urd]} \quad (33)$$

where, i stands only for reactions 1,2,4 and 6.

$$r_{6_{Gal}} = V_{max6_{Gal}} \frac{[Gal]}{K_{M6_{Gal}}\left(1 + \frac{[UDPGal]}{KI_{6D}} + \frac{[Gal]}{KI_{6E}} + \frac{[Urd]}{KI_{6F}}\right) + [Gal]} \quad (34)$$

The material balance of each NSD is derived from Fig. 1 and follows the general form:

$$\frac{d([NSD])}{dt} = \sum_i v_i r_i - F_{out_{NSD}} \quad (35)$$

where, v_i is the stoichiometric coefficient for reaction i, r_i is any of the reactions depicted in Fig. 1 and $F_{out_{NSD}}$ ($mmol_{NSD}·L^{-1}·h^{-1}$) describes the flux of each cytosolic NSD into the Golgi apparatus and represents the transport rate of each NSD consumed for the glycosylation of host cell protein (HCP), glycolipids and mAb product, as shown in Eq. (36):

$$F_{out_{NSD}} = \frac{[NSD]}{K_{TP_{NSD}}+[NSD]}\left(\frac{N_{HCP/Lipids_{NSD}}\mu}{V_{cell}} + \frac{N_{precursor_{NSD}}q_{mAb}}{V_{cell}} + r_{NSD_k}^{mAb,glyc}\right) \quad (36)$$

where, $K_{TP_{NSD}}$ (mM) is the transport protein saturation constant. The first sum term of Eq. (36) describes the flux of NSD used for HCPs and glycolipids synthesis, where V_{cell} (L) is the cellular volume and $N_{HCP/Lipids_{NSD}}$ ($mmol_{NSD}·cell^{-1}$) is the nucleotide sugar donor stoichiometric coefficient for HCP and glycolipids synthesis, as described by del Val, Polizzi et al. (2016). The second term describes the NSD demand for the formation of the precursor oligosaccharide ($Glc_3Man_9GlcNAc_2$), where $N_{precursor_{NSD}}$ ($mmol_{NSD}·cell^{-1}$) is the stoichiometric coefficient of the NSDs that participate in precursor oligosaccharide synthesis (del Val, Polizzi et al. 2016). The third term of Eq. (36) describes the fluxes of NSDs in the latter stages of N-linked glycosylation and are described by Eq.(37) (del Val, Nagy et al. 2011):

$$r_{NSD_k}^{mAb,glyc} = \left(\frac{V_{Golgi}}{V_{cell}}\right)\int_{z=0}^{z=1}\sum_{j=1}^{N.R.} v_{k,j}\cdot r_j(z)\,dz \quad (37)$$

where, $r_{NSD_k}^{Glyc}$ ($mmol_{NSD}·L^{-1}·h^{-1}$) is the rate of the NSD k that is used for mAb glycosylation, k can be any NSD, V_{Golgi} (L^{-1}) is the Golgi volume, $v_{k,j}$ ($mmol_{NSD}·mmol_{OS}^{-1}$) is the stoichiometric coefficient of NSD k required as a co-substrate in reaction j, z is the normalized Golgi length and $r_j(z)$ ($mmol_{OS}·L^{-1}·h^{-1}$) is

the rate of glycosylation reaction *j* along the length of the Golgi that is calculated in the N-linked glycosylation model.

However, the use of Eq. (37) in the current modelling framework substantially increases simulation time because the numerical solution requires convergence between the fluxes calculated by Eq. 36 and the integral over *z*. Eq. (38) is proposed to reduce the simulation time while maintaining a mechanistic relationship between the NSD and glycosylation sub-models. NSD flux for mAb glycosylation is assumed to be a function of oligosaccharide's concentration in the end of the normalized Golgi length (z=1) as the residence time of an oligosaccharide molecule in the Golgi (22min, (del Val, Nagy et al. 2011)) is significantly lower than the time scale within which experimentally observed changes in NSD concentrations occur.

$$r_{NSD_k}^{mAb,glyc} = Vel_{Golgi}\left(\frac{V_{Golgi}}{V_{cell}}\right)\sum_{i=1}^{N.OS} \nu_{k,i} \cdot OS_i(1) \quad (38)$$

where, Vel_{Golgi} (Golgi length·min^{-1}) is the linear velocity with which the molecules travel through the Golgi apparatus, $\nu_{k,i}$ (mmol$_{NSD}$·mmol$_{OS}$ $^{-1}$) is the stoichiometric coefficient of NSD k required to produce the oligosaccharide i and $OS_i(1)$ (mmol$_{OS}$·L^{-1}) is the concentration of the oligosaccharide i in the end of the normalized Golgi length as calculated by the N-linked glycosylation model.

3.3 N-linked glycosylation model

The estimated NSD concentrations are fed to the N-linked mAb glycosylation model that was adapted from del Val, Nagy et al. (2011), in order to estimate the mAb glycoform distribution. The adapted model relates the cytosolic concentration of NSDs with the glycan distribution, considering that the concentration of NSDs inside the Golgi apparatus is 20 times greater than the cytosolic concentration (Waldman and

Rudnick 1990). The measured NSD concentration is the total intracellular NSD concentration that accounts both the concentration of NSDs in the cytosolic and Golgi environment respectively. As proposed by del Val et al. (2016), the current model assumes a constant Golgi volume and a variable linear velocity through Golgi (Vel_{Golgi}), to describe the different rates at which secretory cargo (the mAb) traverses the Golgi apparatus (Presley, Cole et al. 1997).

3.4 Parameter estimation

3.4.1 Culture and NSD model

The CHO cell culture and NSD metabolic model include 102 unknown parameters, from which 85 parameters were estimated by fitting the model equations to the available experimental data (close to 900 data points in total), using the Parameter Estimation entity in gPROMS v.5.0.1. The remaining 17 parameters were assigned fixed values according to literature. In order to obtain satisfactory parameter confidence intervals, the model was separated in modular sections and parameter estimation was carried out sequentially as described below: The parameters of the CHO cell culture model and the NSD metabolic model that were not related to galactose and uridine were estimated based only on the control experiment. Using the values defined in the first step, the parameters that were introduced due to galactose and uridine feeding were estimated based on the four feeding experiments. Therefore, parameter estimation was strongly based on the control experiment and no parameter differs between the control and the feeding experiments. Subsequently and as a final step, 18 parameters of the N-linked glycosylation model were estimated as further described in section 3.4.2.

The standard deviation of each experimentally measured variable was used to calculate a variance model for that measurement, which was used in the estimation of related model parameters. In all kinetic rate expressions, both maximum turnover rates and Monod saturation constants were unknown, making the problem of parameter estimation structurally unidentifiable because the linear dependence between the rate and the substrate concentration below saturation only gives us an estimate of their ratio. Therefore, the majority of Monod constants were hypothesised not to differ more than one order of magnitude from the metabolite or NSD concentration used in each rate, in order to estimate more realistic values for both K_m and V_{max}. However, the estimated values are not intended to accurately represent these kinetic parameters owing to the lack of structural and therefore also numerical identifiability.

3.4.2 N-linked Glycosylation model

The method used for estimating the unknown parameters of the glycosylation model was based on a previously presented method by del Val, Fan et al. (2016). However, a slightly modified method was used for the estimation of the GalT distribution (no sialylation was observed). The enzyme parameters were estimated by obtaining the minimum amount of enzyme required to achieve the experimental Galactosylation Index (*GI*) that is described by Eq. (37):

$$GI = \frac{0.5*\Delta[monogalactosylated\ OS]+1*\Delta[digalactosylated\ OS]}{\Delta[mAb]} \quad (39)$$

where, the numerator describes the secreted concentration (mM) of galactosylated mAb glycoforms that were secreted during the chosen interval, and the denominator

describes the total secreted mAb during that interval. The calculation of the secreted mAb glycoforms in each interval was based on Fan, Jimenez Del Val et al. (2015).

In addition to the estimation of two dissociation constants ($K_{d,i}$) for GalT, the dissociation constants for FucT and GnTI were also estimated, owing to the high concentrations of fucosylated glycans and the low concentrations of Man5 glycans that were observed experimentally. The data from the interval that presented the highest specific mAb productivity was used for $K_{d,i}FucT$ and $K_{d,i}GnTI$ estimation. Considering also the high accumulation of UDPGal that was observed experimentally during the feeding strategies, the estimation of two $K_{d,k}GalT$ was found to be of great importance. The interval that presented the highest galactosylation was chosen for $K_{d,k}GalT$ estimation. The remaining parameters were set to their nominal values (del Val, Nagy et al. 2011).

4. Results

4.1 Model Fitting

The model was simulated for all feeding strategies that were used for model construction and parameter estimation to evaluate the fitting of the model to the training data.

CHO cell culture model simulation

The simulation results present very good agreement with the experimental measurements for all feeding strategies as can be seen in the obtained CCC and R^2 values in Table 1. Selected estimated kinetic parameters that were used for model simulation are depicted in Table 2 (all parameter values can be found in Table S1 in

the Supplementary Material). Cell density, ammonia, lactate and glutamine extracellular concentrations match closely with the experimental measurements and indicate that the model framework construction and parameter estimation were successfully conducted as shown in Fig. 2 and Fig. S1 in the Supplementary Material. Importantly, the model captures the metabolic shift from lactate production to consumption in the end of the exponential growth phase and the ammonia increase due to uridine addition. The model is also able to describe the inhibitory impact of galactose feeding on glucose consumption. Finally, product titre is in quantitative agreement for all the feeding experiments indicating that mAb specific production rate follows a linear relationship with the specific growth rate as assumed in Eq. (19).

The CCCs, R^2 and the PCCs were calculated for cell density and mAb concentration, in order to quantify the agreement of the model results with the experimental measurements. CCC is used to correlate the values of the model results and the experimental measurements and PCC to correlate the trends of the simulated and experimental curves. The calculated CCC and PCC coefficients and R^2 are presented in Table 1.

As shown in Table 1, all agreement coefficients for cell density and mAb concentration are above 0.9 with the exception of cell density in the 10G experiment, where CCC is just below 0.9 and the 10G20U experiment, for which the model is unable to capture the steep death phase. The calculated PCCs are higher than the respective CCCs, as expected, and show that the model correctly captures the trend of the examined variables versus time.

NSD synthesis model simulation

Simulation results for UDPGal, UDPGlc, UDPGlcNAc and UDPGalNAc concentrations are presented in Fig. 3. The model captured the increase of UDPGal concentration due to galactose feeding when no uridine was added. Moreover, the increase of UDPGal concentration in the experiments that included uridine addition was well described in the simulations. Higher concentrations of UDPGlcNAc owing to uridine addition were successfully calculated while GDPMan and GDPFuc estimated concentrations were in the same order of magnitude as the experimental measurements (Fig. S2, Supplementary Material). The significant increase of UDPGalNAc that is not used as a sugar donor in N-linked glycosylation and UDPGlc was also described. Regarding the concentration of UDPGlc in the 10G (B) experiment, the model indicated a slight increase of the concentration over time while the experimental results followed a fluctuating trend leading to a 50-75% difference between the simulation and experimental results in the stationary and death phase of the cell culture. However, the concentration of UDPGlc is fluctuating in a narrow range when compared to the other feeding experiments and therefore this discrepancy is not regarded as significant. The estimated maximum turnover rates for the reactions between NSDs as depicted in Fig. 1 are displayed in Table 2.

The calculated CCCs, R^2 and PCCs of UDPGal and UDPGlcNAc concentrations shown in Table 1 are above 0.9 for all experiments with the exception of the agreement coefficient of UDPGal concentration in the 10G experiment that is 0.897. The high values of the coefficients and R^2 indicate that the NSD model can successfully describe the NSD synthesis pathway and capture the significant alterations caused by precursor feeding.

Glycoform distribution

Fig. 4 compares the simulated and experimental distributions of G0, G0F, G1F, G2, G2F and Man5 mAb glycoforms on days 7, 9, 11 and 12 of cell culture. The majority of glycoform distributions that were calculated by the N-linked glycosylation model were within ±5% of the corresponding experimental measurements. Only the distribution of the G1F glycoform on day 7 in the 50G5U experiment presented a 6.8% underestimation compared to experimental results. Experimentally, specific production rate of galactosylated mAb (Supplementary Table S3) was reduced alongside with the decreasing cell viability (Supplementary Figure S3) post galactose and uridine feeding.

The simulation results presented a general underestimation of the experimental measurements probably due to insufficient estimated enzyme availability. Enzyme distribution parameters were estimated only according to the control experiment to reduce overfitting. Therefore, the increased availability of GlcNAc and galactose was not depicted in enzyme concentrations and resulted in slight underestimations for some distributions in the feeding experiments. The estimated enzyme distribution parameters and dissociation constants are presented in Supplementary Table S2. The remaining parameters of the glycosylation model have been assigned to their nominal values (del Val, Nagy et al. 2011). Glycoform distributions could not be subjected to statistical analysis, as the number of measurements (4) was relatively small.

4.2 Model predictive capability

A model-based dynamic feeding strategy aimed to maximise galactosylation was implemented experimentally and the data was used to verify the predictive capabilities of the developed mathematical model. The dynamic optimisation was

carried out in the relevant function within the gPROMS modelling environment. The objective function we sought to maximise was the concentration of galactosylated mAb species ($mg \cdot L^{-1}$), as described in Eq. (40).

$$galactosylation = 1 \times [monogalactosylated\ mAb] + 2 \times [digalactosylated\ mAb] \quad (40)$$

There were three control variables: the feed flow rate and the concentrations of galactose and uridine in the feed. The problem formulation allowed for the feed to be added or not (binary decision) on alternate days starting on day 2. The concentration of galactose and uridine in the feed was allowed to range between 0 and 500 mM and 0 and 100 mM, respectively. The resulting feeding strategy includes four pulse additions of varying feed concentrations (details shown in Figure 5). Fig. 5 compares model predictions with measured data for that experiment. Cell density and titre concentration predictions are in good agreement with the experimental measurements.

The model captures the timing of the different growth phases and the impact of galactose and uridine feeding on cell density with sufficient accuracy. Despite the addition of these two precursors, which can negatively affect growth, the cell density remains unaffected compared to the control thanks to the inclusion of these dependencies in the mathematical formulation used for optimisation. Glycan distributions measured on days 7, 9, 11 and 12 are then compared to those predicted by the model (Fig. 5). Model accuracy is within ±3.5% with the exception of day 12 measurement of G0F that is overestimated by 5.3%. The experimentally measured galactosylated-mAb concentration at day 12 of cell culture exceeds the respective model prediction by 33% due to the unexpected higher total mAb concentration and G1F distribution observed. We hypothesise that the underestimation of the

galactosylation percentage is due to intracellular regulation of glycosyltransferase enzymes and/or NSD transporter proteins, which are not included in the model. Indeed, if the concentration of GalT used for model simulation is increased by only 15%, the simulation results are in significantly better agreement with the experimental findings, particularly for G0F and G1F species (Fig. 5). The CCCs, R^2 and PCCs of cell density and mAb concentration presented in Table 1, are above 0.95 indicating that the model successfully predicts cell density and mAb concentration over the culture period. The proposed experimental/modelling platform can therefore be used to design optimal NSD precursor feeding strategies that control glycosylation.

To this end, Fig. 6 illustrates the impact of the optimized feeding strategy on the percentage and the concentration of the galactosylated mAb. In both cases, the optimized feeding strategy presents higher galactosylation compared to the control experiment. Unlike the control experiment where the galactosylation percentage decreases through time, the optimized feeding strategy leads to an increasing galactosylation percentage and galactosylated mAb concentration during the cell culture period.

5. Discussion

A mechanistic mathematical model to describe cell metabolism, mAb production, NSD synthesis and N-linked glycosylation has been proposed. The model was trained to data from a wide range of culture conditions in order to capture the responses of all described pathways to different feeding strategies of galactose and uridine addition. Then, the model's predictive capability was successfully validated with an independent experiment. The predictive experiment included different additions of galactose and uridine in a range of concentrations and on different culture days

compared to the experiments that were used for model training. The model results for the predictive experiment showed very good agreement with the experimental measurements. The agreement was quantified by calculating the corresponding CCCs and R^2, which were generally over 0.95 (Table 1). Moreover, the majority of the predicted glycoform distributions deviated up to ±3.5% from the measured data (Fig. 4). The high values of all CCCs, R^2 and PCCs indicate that model successfully reproduces the experimental measurements and can therefore be potentially used as a predictive and optimization tool. To the authors' knowledge, this is one of the first studies to dynamically capture variations in mAb glycoform distribution and the first to accurately reproduce and predict the impact of different NSD precursor feeding strategies.

The model suggests that a relationship between the glucose flux through the glycolytic pathway and galactose supplementation exists and is regulated from the fraction of galactose and glucose consumption rates. Experimental results indicate that galactose addition results in a decreased glucose consumption rate, an effect that has been previously observed in CHO cells (Wong, Wati et al. 2010, Gramer, Eckblad et al. 2011). In addition, a simple model for mAb production was used to avoid overparameterization. The results of mAb concentration tracked closely the experimental measurements indicating that mAb production rate is linearly correlated to specific cell growth rate.

The proposed NSD synthesis model was used to calculate the concentrations of the sugar donors that participate in N-linked glycosylation, in particular UDPGal and UDPGlcNAc, that are extensively affected by the addition of galactose and uridine. Jedrzejewski et al. (2014) attempted to describe a detailed NSD metabolic pathway

for hybridoma cells but this led to an extensive and overparameterized reaction network with 34 material balances and 60 reactions. The reduced metabolic network that was used in this modelling approach includes only 7 material balances and 16 reactions, including the effects of galactose and uridine in NSD pools. Monod type equations were used to describe the NSD synthesis reaction rates, including the competitive or non-competitive inhibitions from other NSDs.

However, the NSD model presents issues of structural unidentifiability and therefore the calculated values of the parameters are not a representation of their actual values, specifically with respect to $K_{Mi_{NSD}}$. Moreover, the model is limited to a relatively small number of metabolites and amino acids, which, however, are routinely monitored industrially, and could be further expanded to account the majority of amino acids that participate in cell metabolism and the metabolic intermediates of the TCA cycle. As previously mentioned, the occasional deviations in glycan distributions are attributed to the limitations set by the chosen parameter estimation method that does not include the effect of galactose and uridine feeding on the expression of glycosyltranferases and transport proteins of nucleotide sugars (Clark, Griffiths et al. 2005).

Unstructured mathematical models are usually limited to the cell line and experimental data that are used for model construction and parameter estimation. Although the proposed modelling framework can adapt to different feeding concentrations of galactose and uridine, of its applicability to other cell lines and products remains to be tested. As a first step, a parameter sensitivity analysis can be performed to select the most significant parameters to be re-estimated based on a relatively smaller set of experimental data for a different system.

Despite these drawbacks, the model presented herein successfully captures the impact of galactose and uridine feeding on the secreted mAb glycoform distribution through the use of mechanistic relationships in cell metabolism and NSD synthesis. It captures a wide range of dynamic feeding regimes and was successfully used to design an operating strategy that enhances the percentage of galactosylation up to 64% and the concentration of galactosylated mAb up to 93% with no impact on other key performance indicators, in contrast to the experiments used for model development in which we saw a significant decrease in viable cell density and, consequently, antibody concentration as a result of galactose and uridine feeding. Therefore, the proposed modelling platform can be a vehicle for applying the QbD paradigm to optimize the design space for the production of the desired glycan distribution.

Acknowledgments

PK gratefully acknowledges his funding from the PhD scholarship scheme of the Department of Chemical Engineering, Imperial College London. PMJ and SNS thank the Engineering and Physical Sciences Research Council and the Biotechnology and Biological Sciences Research Council, respectively, for their studentships. The authors thank MedImmune for their funding, providing the cell line used in this study as well as access to their analytical facilities.

References

Clark, K. J. R., et al. (2005). "Gene-expression profiles for five key glycosylation genes for galactose-fed CHO cells expressing recombinant IL-4/13 cytokine trap." Biotechnology and Bioengineering **90**(5): 568-577.

del Val, I. J., et al. (2016). "Dynamics of immature mAb glycoform secretion during CHO cell culture: An integrated modelling framework." Biotechnology Journal **11**(5): 610-623.

del Val, I. J., et al. (2013). "An optimized method for extraction and quantification of nucleotides and nucleotide sugars from mammalian cells." Analytical Biochemistry **443**(2): 172-180.

del Val, I. J., et al. (2011). "A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus." Biotechnol Prog **27**(6): 1730-1743.

del Val, I. J., et al. (2016). "A theoretical estimate for nucleotide sugar demand towards Chinese Hamster Ovary cellular glycosylation." Scientific Reports **6**: 28547.

Fan, Y., et al. (2015). "A multi-pronged investigation into the effect of glucose starvation and culture duration on fed-batch CHO cell culture." Biotechnology and Bioengineering **112**(10): 2172-2184.

Goetze, A. M., et al. (2011). "High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans." Glycobiology **21**(7): 949-959.

Grainger, R. K. and D. C. James (2013). "CHO cell line specific prediction and control of recombinant monoclonal antibody N-glycosylation." Biotechnology and Bioengineering **110**(11): 2970-2983.

Gramer, M. J., et al. (2011). "Modulation of antibody galactosylation through feeding of uridine, manganese chloride, and galactose." Biotechnology and Bioengineering **108**(7): 1591-1602.

Hodoniczky, J., et al. (2005). "Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro." Biotechnol Prog **21**(6): 1644-1652.

Hutter, S., et al. (2017). "Glycosylation flux analysis reveals dynamic changes of intracellular glycosylation flux distribution in Chinese hamster ovary fed-batch cultures." Metabolic Engineering **43**: 9-20.

Jedrzejewski, P., et al. (2014). "Towards Controlling the Glycoform: A Model Framework Linking Extracellular Metabolites to Antibody Glycosylation." International Journal of Molecular Sciences **15**(3): 4492.

Kanehisa, M. and S. Goto (2000). "KEGG: Kyoto Encyclopedia of Genes and Genomes." Nucleic Acids Research **28**(1): 27-30.

Kaveh, O., et al. (2013). "Novel Dynamic Model to Predict the Glycosylation Pattern of Monoclonal Antibodies from Extracellular Cell Culture Conditions." IFAC Proceedings Volumes **46**(31): 30-35.

Kontoravdi, C. and I. J. del Val (2018). "Computational tools for predicting and controlling the glycosylation of biopharmaceuticals." Current Opinion in Chemical Engineering **22**: 89-97.

Kontoravdi, C., et al. (2010). "Systematic development of predictive mathematical models for animal cell cultures." Computers & Chemical Engineering **34**: 1192-1198.

Krambeck, F. J. and M. J. Betenbaugh (2005). "A mathematical model of N-linked glycosylation." Biotechnology and Bioengineering **92**(6): 711-728.

Kremkow, B. G. and K. H. Lee (2018). "Glyco-Mapper: A Chinese hamster ovary (CHO) genome-specific glycosylation prediction tool." Metab Eng **47**: 134-142.

Lai, T., et al. (2013). "Advances in Mammalian Cell Line Development Technologies for Recombinant Protein Production." Pharmaceuticals **6**(5): 579-603.

Lawrence, I. K. L. (1989). "A Concordance Correlation Coefficient to Evaluate Reproducibility." Biometrics **45**(1): 255-268.

Liu, B., et al. (2014). "The availability of glucose to CHO cells affects the intracellular lipid-linked oligosaccharide distribution, site occupancy and the N-glycosylation profile of a monoclonal antibody." Journal of Biotechnology **170**: 17-27.

Nolan, R. P. and K. Lee (2011). "Dynamic model of CHO cell metabolism." Metabolic Engineering **13**(1): 108-124.

Presley, J. F., et al. (1997). "ER-to-Golgi transport visualized in living cells." Nature **389**: 81.

Schomburg, I., et al. (2002). "BRENDA, enzyme data and metabolic information." Nucleic Acids Research **30**(1): 47-49.

Shinkawa, T., et al. (2003). "The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity." Journal of Biological Chemistry **278**(5): 3466-3473.

Sou, S. N., et al. (2015). "How does mild hypothermia affect monoclonal antibody glycosylation?" Biotechnology and Bioengineering **112**(6): 1165-1176.

Spahn, P. N., et al. (2016). "A Markov chain model for N-linked protein glycosylation – towards a low-parameter tool for model-driven glycoengineering." Metabolic Engineering **33**: 52-66.

Thomann, M., et al. (2016). "Fc-galactosylation modulates antibody-dependent cellular cytotoxicity of therapeutic antibodies." Mol Immunol **73**: 69-75.

Urquhart, L. (2018). "Market watch: Top drugs and companies by sales in 2017." Nat Rev Drug Discov **17**(4): 232.

Waldman, B. C. and G. Rudnick (1990). "UDP-GlcNAc transport across the Golgi membrane: electroneutral exchange for dianionic UMP." Biochemistry **29**(1): 44-52.

Walsh, G. (2014). "Biopharmaceutical benchmarks 2014." Nature Biotechnology **32**: 992.

Wieczorke, R., et al. (2003). "Characterisation of Mammalian GLUT Glucose Transporters in a Heterologous Yeast Expression System." Cellular Physiology and Biochemistry **13**(3): 123-134.

Wong, N. S. C., et al. (2010). "An investigation of intracellular glycosylation activities in CHO cells: Effects of nucleotide sugar precursor feeding." Biotechnology and Bioengineering **107**(2): 321-336.

Yang, M. and M. Butler (2000). "Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation." Biotechnology and Bioengineering **68**(4): 370-380.

Yu, M., et al. (2012). "Production, characterization and pharmacokinetic properties of antibodies with N-linked Mannose-5 glycans." mAbs **4**(4): 475-487.

Yusufi, F. N. K., et al. (2017). "Mammalian Systems Biotechnology Reveals Global Cellular Adaptations in a Recombinant CHO Cell Line." Cell Systems **4**(5): 530-542.e536.

Zagari, F., et al. (2013). "Lactate metabolism shift in CHO cell culture: the role of mitochondrial oxidative activity." New Biotechnology **30**(2): 238-245.

Zhao, F.-Q. and A. F. Keating (2007). "Functional Properties and Genomics of Glucose Transporters." Current Genomics **8**(2): 113-128.

Figures

Figure 1. Simplified NSD metabolic pathway. The reactions are indicated by solid lines and the transport rates by dotted lines. The transport of extracellular metabolites is not included in model formulation.

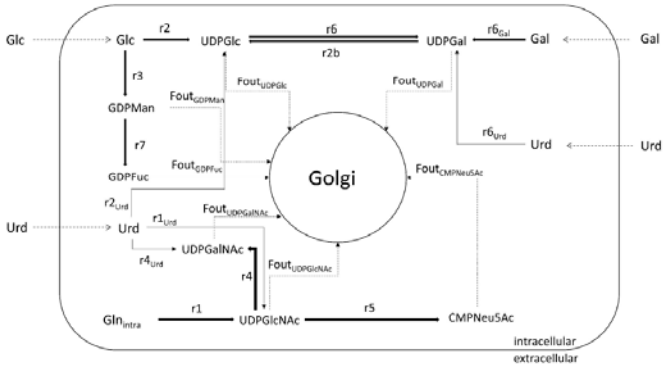


Figure 2. Model simulations (red line) and experimental measurements (black squares) for the viable cell density, extracellular mAb, glucose and lactate concentrations. Row A presents the fitting of the model to the control experiment, row B the fitting of the model to the 10G experiment, row C the fitting of the model to the 10G5U experiment, row D the fitting of the model to the 10G20U experiment and row E the fitting of the model to the 50G5U experiment.

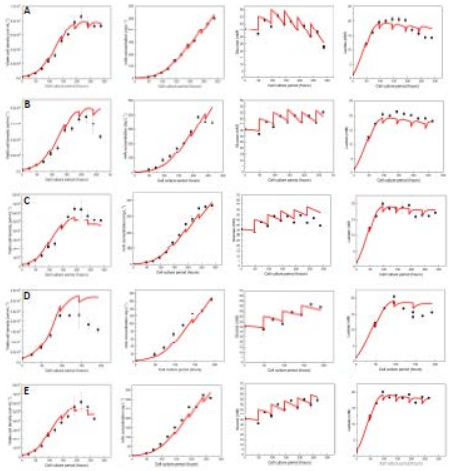


Figure 3. Model simulations (red line) and experimental measurements (black squares) for UDPGal, UDPGlc, UDPGlcNAc and UDPGalNAc concentration. Row A presents the fitting of the model to the control experiment, row B the fitting of the model to the 10G experiment, row C the fitting of the model to the 10G5U experiment, row D the fitting of the model to the 10G20U experiment and row E the fitting of the model to the 50G5U experiment.

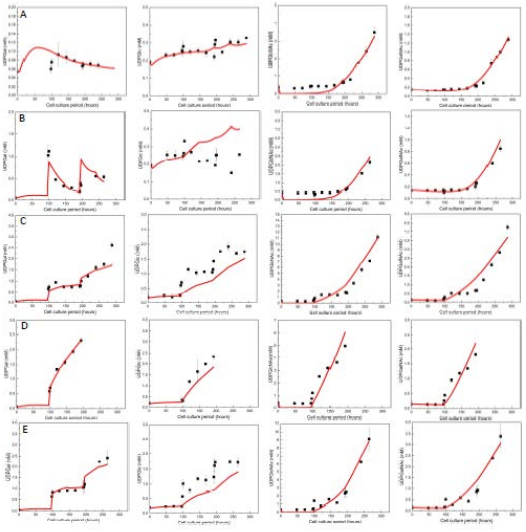


Figure 4. Model simulation of glycan distribution for all feeding experiments, including the control. Experimental data for the 10G (288h), 10G20U (216, 264 and

288h) and 50G5U (288h) experiments is missing as the respective cell cultures were terminated earlier due to reduced (below 70%) cell viability.

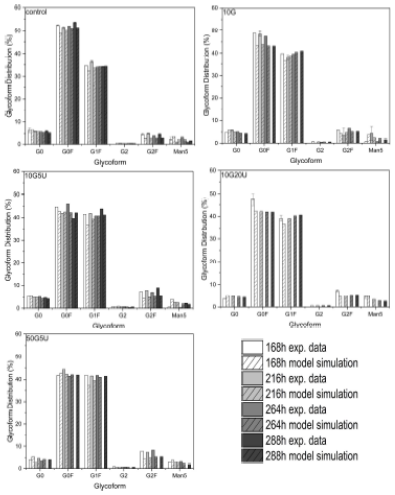


Figure 5. Model prediction and experimental measurements for the optimization experiment used to validate the model. The graph presents the model simulation (red line) and experimental data (black squares) for (A) cell density and (B) mAb concentration while the arrows indicate the days that galactose and uridine were added. Graph (C) depicts the glycan distribution in the optimization experiment, comparing the experimental data with the model predictions for unchanged GalT concentration and for a 15% increase in the concentration of GalT. Experimental conditions: V_{feed} = 10% $V_{culture}$; $[Gal]_{feed,day2}$ = 0 mM, $[Gal]_{feed,day4}$ = 6.53 mM, $[Gal]_{feed,day6}$ = 9.30 mM, $[Gal]_{feed,day8}$ = 8.98 mM and $[Gal]_{feed,day10}$ = 8.69 mM; $[Urd]_{feed,day2}$ = 0 mM, $[Urd]_{feed,day4}$ = 0.76 mM, $[Urd]_{feed,day6}$ = 1.34 mM, $[Urd]_{feed,day8}$ = 2.81 mM and $[Urd]_{feed,day10}$ = 10 mM.

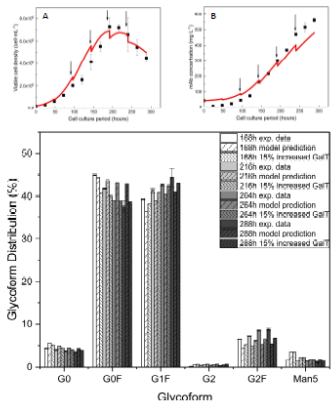
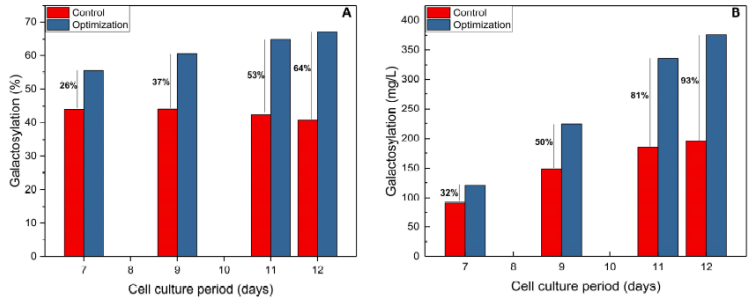


Figure 6. (A) Experimental degree of galactosylation (%) in the control (blue bars) and the optimization (red bars) experiment. (B) Experimentally measured concentration of galactosylated species (mg·L^{-1}) in the control (blue bars) and the optimization (red bars) experiment.



Tables

Table 1. Concordance correlation coefficients to evaluate the agreement of experimental measurements and model results

	Concordance correlation coefficients			
Experiment	CCC_{Xv}	CCC_{mAb}	CCC_{UDPGal}	$CCC_{UDPGlcNAc}$
Control	0.986	0.998	0.965	0.973
10G	0.892	0.958	0.897	0.919
10G5U	0.960	0.990	0.942	0.968
10G20U	0.741	0.974	0.982	0.941
50G5U	0.971	0.986	0.995	0.990
Predictive experiment	0.959	0.968	-	-

Pearson correlation coefficients

Experiment	**PCC_{Xv}**	**PCC_{mAb}**	**PCC_{UDPGal}**	**$PCC_{UDPGlcNAc}$**
Control	0.990	0.999	0.971	0.985
10G	0.971	0.970	0.930	0.983
10G5U	0.978	0.994	0.949	0.982
10G20U	0.946	0.988	0.989	0.965
50G5U	0.975	0.991	0.997	0.992
Predictive experiment	0.972	0.996	-	-

R squared (R^2)

Experiment	R^2_{Xv}	R^2_{mAb}	R^2_{UDPGal}	$R^2_{UDPGlcNAc}$
Control	0.992	0.999	0.998	0.968
10G	0.979	0.968	0.963	0.922
10G5U	0.984	0.995	0.961	0.976
10G20U	0.970	0.988	0.998	0.952
50G5U	0.980	0.992	0.990	0.985
Predictive experiment	0.984	0.984	-	-

Table 2. Selected estimated parameters of the CHO cell culture and NSD metabolic models

Estimated parameter	Value	Units	95% Confidence Interval
μ_{max}	6.50×10^{-2}	h^{-1}	7.44×10^{-4}
$\mu_{death,max}$	1.50×10^{-2}	h^{-1}	2.56×10^{-3}
K_{Glc}	14.04	mM	5.05×10^{-1}
K_{Asn}	2.62	mM	7.18×10^{-2}
KI_{Amm}	3.17	mM	2.49×10^{-1}
KI_{Lac}	1×10^{3}	mM	N/A
KI_{Urd}	41.09	mM	2.42
$K_{d,Amm}$	14.28	mM	3.00
$K_{d,Urd}$	27.86	mM	3.71
$Y_{mAb,X}$	3.39	$(pg\cdot cell^{-1})$	3.21×10^{-1}
m_{mAb}	4.10×10^{-1}	$(pg\cdot cell^{-1}\cdot h^{-1})$	6.57×10^{-3}
V_{max1}	9.22×10^{-1}	$(mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1})$	1.24×10^{-2}
V_{max2}	1.70×10^{-2}	$(mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1})$	9.78×10^{-4}
V_{max2b}	59.48	$(mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1})$	6.10
V_{max3}	5.50×10^{-2}	$(mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1})$	2.60×10^{-3}

V_{max4}	2.65×10^{-2}	($mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1}$)	2.12×10^{-3}
V_{max5}	1×10^{-4}	($mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1}$)	7.80×10^{-6}
V_{max6}	5.13	($mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1}$)	1.95×10^{-1}
V_{max7}	4.60	($mmol_{NSD}\cdot L_{cell}^{-1}\cdot h^{-1}$)	4.78×10^{-1}