

Article

Glycosylation Flux Analysis of Immunoglobulin G in Chinese Hamster Ovary Perfusion Cell Culture

Sandro Hutter ^{1,2,†}, Moritz Wolf ^{1,†} , Nan Papili Gao ^{1,2}, Dario Lepori ¹, Thea Schweigler ¹, Massimo Morbidelli ¹  and Rudiyanto Gunawan ^{1,2,3,*} 

¹ Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland; sandro.hutter@outlook.com (S.H.); moritz.wolf@chem.ethz.ch (M.W.); nan.papili-gao@chem.ethz.ch (N.P.G.); dario.lepori@bluewin.ch (D.L.); thea-schweigler@web.de (T.S.); massimo.morbidelli@chem.ethz.ch (M.M.)

² Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

³ Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA

* Correspondence: rudiyant@buffalo.edu

† These authors contributed equally to this work.

Received: 30 May 2018; Accepted: 18 September 2018; Published: 1 October 2018



Abstract: The terminal sugar molecules of the N-linked glycan attached to the fragment crystallizable (Fc) region is a critical quality attribute of therapeutic monoclonal antibodies (mAbs) such as immunoglobulin G (IgG). There exists naturally-occurring heterogeneity in the N-linked glycan structure of mAbs, and such heterogeneity has a significant influence on the clinical safety and efficacy of mAb drugs. We previously proposed a constraint-based modeling method called glycosylation flux analysis (GFA) to characterize the rates (fluxes) of intracellular glycosylation reactions. One contribution of this work is a significant improvement in the computational efficiency of the GFA, which is beneficial for analyzing large datasets. Another contribution of our study is the analysis of IgG glycosylation in continuous perfusion Chinese Hamster Ovary (CHO) cell cultures. The GFA of the perfusion cell culture data indicated that the dynamical changes of IgG glycan heterogeneity are mostly attributed to alterations in the galactosylation flux activity. By using a random forest regression analysis of the IgG galactosylation flux activity, we were further able to link the dynamics of galactosylation with two process parameters: cell-specific productivity of IgG and extracellular ammonia concentration. The characteristics of IgG galactosylation dynamics agree well with what we previously reported for fed-batch cultivations of the same CHO cell strain.

Keywords: N-linked glycosylation; perfusion cell culture; CHO cells; constraint-based modeling; monoclonal antibody

1. Introduction

Therapeutic recombinant monoclonal antibodies constitute the most important class of drugs in the biopharmaceutical industry, making up approximately half of the total revenue of biopharmaceutical products in 2013 [1]. The production of mAb drugs typically employs batch or fed-batch cultivations of mammalian cells. The state-of-the-art batch/fed-batch cell cultures are able to meet the large production volume requirement of mAbs, with reactor sizes of up to 25,000 L [2]. Nevertheless, the increasing number of mAb products entering different stages of clinical trials and the burgeoning market of biosimilars driven by impending patent expirations of blockbuster mAb drugs, give a strong motivation for the development of new production technology that is more robust and cost effective [3]. The US Food and Drug Administration (FDA) initiative on quality by design

and process analytical technology put further emphasis on implementing quantitative approaches for process improvements in biopharmaceutical manufacturing [4,5].

Continuous manufacturing technology offers an effective and flexible way for the large scale and robust production of drug compounds [6]. In the biopharmaceutical industry, the application of continuous cell culture technology has thus far been limited to the production of unstable products that require constant recovery [7]. Nevertheless, continuous perfusion cell cultures have previously been demonstrated to be capable of producing antibodies at a volumetric rate that matches or exceeds that of fed-batch cultures [8]. In addition to the high productivity, the stable steady state operation mode in conjunction with the short residence time of perfusion cultures translates to a tight maintenance of product quality. Whether or not product qualities and their key controlling parameters can be directly translated from the traditional batch/fed-batch cell cultures to the perfusion cell cultivations is still unresolved. Past work on converting mAb production from batch/fed-batch to perfusion cell cultures gave conflicting reports on product qualities, where a few studies demonstrated consistent product qualities [9,10], and many others showed differences between the two production modes [11–15].

Among the key critical quality attributes (CQAs) of therapeutic mAbs is the glycan structures of the Fc domain [16]. The N-linked glycosylation is a common post-translational modification of proteins, a process that occurs in the endoplasmic reticulum (ER) and Golgi apparatuses. The Fc glycan structure has been shown to impact protein folding [17], clearance [18], bioactivity [19] and efficacy [20], as well as immunogenicity [21]. Moreover, there exists naturally-occurring heterogeneity in the glycosylation of mAbs. Thus, the FDA approval of mAb drugs is given for a particular composition of mAb glycoforms [5]. Because of the importance of the N-linked glycosylation, the majority of recombinant mAb drugs are produced using mammalian host cells in order to achieve human-like glycan structures [22]. In particular, Chinese Hamster Ovary (CHO) cells have become the major expression host for the biopharmaceutical production of therapeutic mAbs [23].

The N-linked glycosylation of mAbs has been shown to depend on the host genetic background [24], expression vector [25], media composition [16], media supplements [26], and bioprocess parameters [27]. However, the mechanism of the above dependence remains to be established [28]. Toward closing this gap, mathematical models of the glycosylation network have previously been developed [29–36]. Many of these models have a large number of unknown and system-specific kinetic parameters that need to be fitted to experimental data [29–33]. A number of parameter-free models have also been proposed to study the N-linked glycosylation process based on the constraint-based modeling approach (i.e., stoichiometric models) [34–36]. Recently, we proposed a flux analysis method, called glycosylation flux analysis, which provides predictions of the rate or flux of intracellular glycosylation reactions using the stoichiometric model of the glycosylation network and the cell secretion rates of mAb glycoforms [36]. Similar to the well-known metabolic flux analysis (MFA) [37], the GFA is based on the molar balance equations of glycoforms involved in the glycosylation reaction network under the pseudo steady state assumption. The GFA further makes use of the relatively small number of enzymes involved in the glycosylation process to reduce the degrees of freedom in the flux estimation. More specifically, we assumed that glycosylation fluxes vary with time according to a (global) cell-specific factor and a (local) enzyme-specific factor. When applied to study the IgG production of CHO cells in fed-batch cultivations, the GFA was able to give insights on how changes in the media affect the intracellular IgG glycosylation reactions [36].

The formulation of the flux estimation in the GFA involves a nonlinear least-square regression, which requires computationally intensive global optimizations. One contribution of this work is a reformulation of the flux estimation in the GFA into an iterative procedure involving two linear regression problems, which provides higher computational efficiency. In this study, we applied the improved GFA to analyze the N-linked glycosylation of IgG in perfusion CHO cell cultures undergoing setpoint changes in the viable cell densities and perfusion rates. In the post-analysis of the GFA results, we employed a random forest regression analysis to elucidate the key process parameters that affect intracellular IgG glycosylation fluxes. Finally, we compared the characteristics of IgG glycosylation

between the traditional fed-batch and continuous perfusion cell cultures and identified similarities and differences.

2. Materials and Methods

Figure 1 illustrates the analysis of IgG glycosylation in our study. The perfusion cell culture experiments are summarized in Section 2.1. In the data preprocessing step, we computed the cell-specific secretion fluxes of IgG glycoforms, as well as other molecules such as glucose, lactate and ammonia, from cell culture measurements. The data preprocessing is described in Section 2.2. We applied the GFA to the secretion fluxes of IgG glycoforms calculated above using a constraint-based model of the glycosylation network. The GFA generates estimates of enzyme-specific glycosylation activities at different measurement time points. An improved iterative procedure for the GFA is detailed in Section 2.3. Finally, we performed a random forest regression analysis, as outlined in Section 2.4, to identify the process parameters that are able to explain the dynamic behavior of intracellular glycosylation activities.

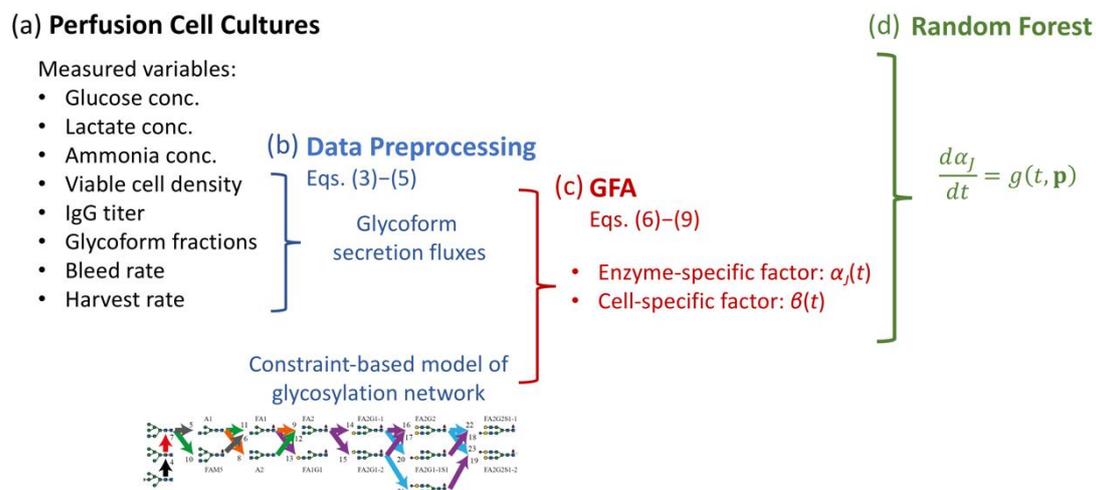


Figure 1. Analysis procedure of IgG glycosylation in CHO perfusion cell cultures. (a) Four perfusion cell cultures were performed (see Section 2.1). Measurements of glucose/lactate/ammonia concentrations, viable cell density, IgG titer, glycoform fractions, bleed rates and harvest rates were taken from each cell culture run. (b) In the data preprocessing step (see Section 2.2), the cell-specific secretion fluxes of IgG glycoforms, as well as the cell-specific secretion/uptake rates of IgG, glucose, lactate and ammonia, were computed. (c) The glycosylation flux analysis (GFA) was applied to the secretion fluxes of IgG and its glycoforms (see Section 2.3). The GFA is based on a constraint-based model of the glycosylation network where the glycosylation fluxes vary with time according to two multiplicative factors: the enzyme-specific factor $\alpha_j(t)$ and the cell-specific factor $\beta(t)$. The factor $\alpha_j(t)$ captures the dynamic changes in the enzyme-specific glycosylation activities, while the factor $\beta(t)$ describes the dynamic changes in the cell-specific IgG productivity. (d) A random forest regression analysis was carried out to rank the process parameters (\mathbf{p}) based on their ability to predict the dynamic changes of $\alpha_j(t)$.

2.1. Continuous Perfusion Cell Cultures

The detailed procedure of the perfusion cell culture is available elsewhere [38]. Here, we provide a brief summary of the experiments. A proprietary CHO cell line expressing IgG1 was cultured using a previously developed 1.5 L perfusion cell culture setup [39], as depicted in Figure 2. In total, four perfusion cell culture experiments were performed with different viable cell density (VCD) and perfusion rate (PR) set-point profiles, as shown in Figure 3. Each of the experiments was inoculated following the same procedure. CHO cells were held back in the reactor by a cell retention device so that only cell free reaction mixture left the reactor through the harvest stream. The feed flowrate (F) of fresh media into the reactor and the bleed (B) and harvest (H) flowrates out of the reactor were

balanced to keep the reactor volume constant. The perfusion rate (PR) is given by the flowrate through the reactor, as follows:

$$PR = H + B = F \quad (1)$$

The perfusion rate represents the rate of fresh media supplied to the cells. We define the cell-specific perfusion rate (CSPR) as the rate of fresh media per cell:

$$CSPR = \frac{PR}{VCD} \quad (2)$$

During the experiments, the following measurements were collected on a daily basis: cell counts and cell viability, concentrations of glucose, lactate and ammonia, IgG titer, and protein glycan distribution.

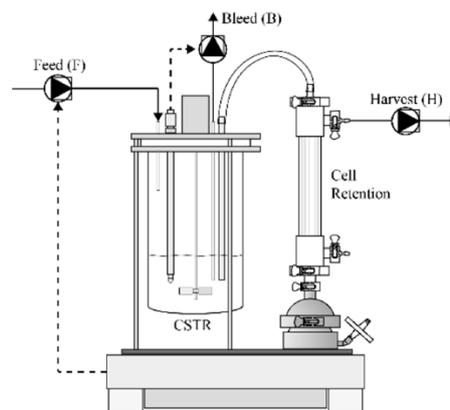


Figure 2. A schematic of perfusion cell culture reactor (Figure adapted from [39] with the permission from the authors). CHO cells were cultivated in suspension in a continuous stirred tank reactor with continuous feeding of fresh nutrients. Cell-free spent media was constantly collected in the harvest stream, while cells remained in the stirred tank reactor. A bleed stream removed a small fraction of the reactor mixture, including biomass, which was used to regulate viable cell density.

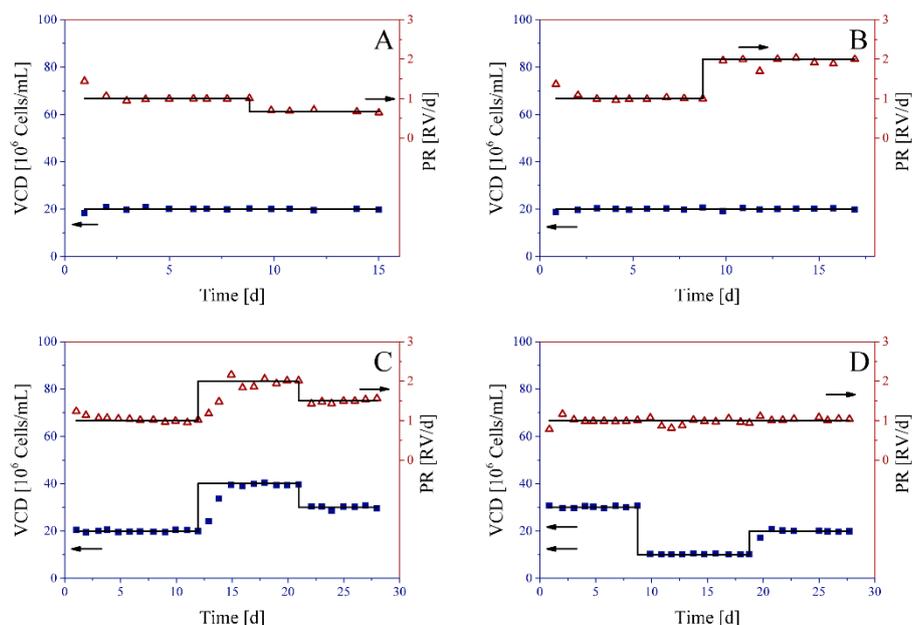


Figure 3. Perfusion cell culture experiments. Four perfusion cell culture experiments, labeled (A), (B), (C) and (D), were conducted with varying VCD and PR given set-points (solid lines). The experimental VCD and PR are shown as blue filled squares and red empty triangles, respectively.

2.2. Estimation of Secretion and Uptake Fluxes

As inputs to the GFA, the cell-specific secretion flux of each IgG glycoform ($v_{E,i}$) was determined based on the molar balance in the reactor, as follows

$$V \frac{dc_{E,i}(t)}{dt} = Vv_{E,i}(t)VCD(t) - (H(t) + B(t))c_{E,i}(t) \quad (3)$$

where $c_{E,i}$ denotes the concentration of the i -th IgG glycoform and V denotes the reactor volume. Since the reactor volume is kept (approximately) constant, Equation (3) can be rearranged to give:

$$v_{E,i}(t) = \frac{\left(\frac{dc_{E,i}(t)}{dt} + (h(t) + b(t))c_{E,i}(t) \right)}{VCD(t)} \quad (4)$$

The variables $h(t) = \frac{H(t)}{V}$ and $b(t) = \frac{B(t)}{V}$ represent the specific harvest and bleed rate, respectively. The concentration of the IgG glycoform was calculated as the product of the measured glycoform fraction (f_i) and the IgG titer (T) according to:

$$c_{E,i}(t_k) = f_i(t_k)T(t_k) \quad (5)$$

where t_k denotes the k -th measurement time point.

For the estimation of the secretion fluxes in Equation (4), the computed concentration of IgG glycoforms $c_{E,i}(t_k)$ was first smoothed (as a function of time) using spline fitting. The time derivative $\frac{dc_{E,i}(t)}{dt}$ was then evaluated by taking the first order derivative of the spline curve. Equation (4) can be used to compute the cell-specific uptake/secretion fluxes of IgG, glucose, lactate and ammonia by replacing $c_{E,i}(t)$ with the concentration measurements of IgG, glucose, lactate and ammonia, respectively.

2.3. Glycosylation Flux Analysis

The glycosylation flux analysis is based on a constraint-based model of the protein glycosylation reaction network. Like other flux analysis methods, the GFA is developed for evaluating the intracellular (glycosylation) fluxes using the cellular (mAb glycoforms) secretion rates. In the GFA, a pseudo steady state assumption is taken to derive the stoichiometric model of the glycosylation network, as follows:

$$\frac{dc_I(t)}{dt} = Sv_I(t) - v_E(t) = 0 \quad (6)$$

where c_I denotes the vector of m intracellular IgG glycoform concentrations, v_I denotes the vector of n intracellular IgG glycosylation reaction fluxes (rates), v_E denotes the vector of secretion fluxes of IgG glycoforms estimated above, and S denotes the $m \times n$ stoichiometric matrix. The (i,j) -th element of S gives the number of the i -th glycoform molecule produced (if positive) or consumed (if negative) by the j -th glycosylation reaction. Since the number of reaction fluxes (i.e., the number of unknowns) typically exceeds that of glycoforms (i.e., the number of equations), the estimation of v_I from v_E in Equation (6) is underdetermined. In other words, there exist many v_I for the same experimentally determined v_E . In the method Flux Balance Analysis, the most plausible v_I is set to the vector that maximizes a cellular objective, such as the biomass production [37]. However, the appropriate cellular objective to use for glycosylation networks is not immediately obvious.

Instead of assuming a cellular objective, the GFA enforces constraints on how groups of glycosylation fluxes vary together. More specifically, in the GFA each glycosylation flux $v_{I,j}(t)$ is computed as the product of a reference flux value $v_{I,j}^{\text{ref}}$, an enzyme-specific factor $\alpha_j(t)$ and a cell-specific factor $\beta(t)$, as follows:

$$v_{I,j}(t) = \alpha_j(t)\beta(t)v_{I,j}^{\text{ref}} \quad (7)$$

The variables $\alpha_j(t)$ and $\beta(t)$ represent the fold-change amplification or attenuation and can therefore be normalized to 1 at a chosen reference time point t_{ref} .

- (1) generate a uniformly distributed random vector of $\mathbf{v}_I^{\text{ref}}$ within a biologically feasible range ($v_{I,j}^{\text{ref}} \in [0, 25]$; unit : pg/cell/day),
- (2) given $\mathbf{v}_I^{\text{ref}}$ from step (1), solve for or update α_j using Equation (10),
- (3) given α_j from step (2), solve for or update $\mathbf{v}_I^{\text{ref}}$ using Equation (11), and
- (4) repeat steps (2) and (3) until the change of Φ as described in Equation (9) becomes smaller than a threshold (default 10^{-10}).

Equations (10) and (11) were solved following the iterative procedure above using the MATLAB subroutine lsqin constraining α_j to values between zero and 20 and $\mathbf{v}_I^{\text{ref}}$ to be positive. In order to improve the chance of obtaining the global minimum of Φ , we adopted a multi-start strategy and ran the aforementioned iterative estimation for multiple random initial vectors of $\mathbf{v}_I^{\text{ref}}$. Among the results of the multi-start runs, we took the best α_j and $\mathbf{v}_I^{\text{ref}}$ values that correspond to the lowest Φ value. Note that the multi-start strategy is embarrassingly parallel and can be implemented on a multiprocessor computing platform.

2.4. Random Forest for Regression

For identifying explanatory variables of the dynamical changes in the galatossiltransferase activity $\alpha_{J,\text{GalT}}$, we formulated a regression problem using the change in the GalT enzyme-specific factors over time $d\alpha_{J,\text{GalT}}/dt$ as the response variable, and the experimental parameters of perfusion cell cultures as predictor variables. More specifically, we are interested the following regression problem:

$$\frac{d\alpha_{J,\text{GalT}}}{dt} = g(t, \alpha_{J,\text{GalT}}, \mathbf{p}) \quad (12)$$

which describes the dynamic change in $\alpha_{J,\text{GalT}}$ using the function $g(t, \alpha_{J,\text{GalT}}, \mathbf{p})$ that depends on time, $\alpha_{J,\text{GalT}}$, and process parameters \mathbf{p} . Note that the function $g(t, \alpha_{J,\text{GalT}}, \mathbf{p})$ is likely nonlinear in nature. Here, we employed Random Forest (RF) [40] to build the above regression model using data from all four perfusion cell culture experiments. RF regression involves building an ensemble of unpruned regression trees, in which each regression tree is created using a bootstrap sample of the original dataset. At each node of a tree, a subset of predictors is selected randomly to determine the best decision split of the samples. The final prediction of the regression trees in RF is obtained by averaging the predictions of the entire ensemble. Notably, RF regression is able to capture nonlinear dependencies of the response variable on the predictors.

In our work, we applied RF using normalized data of each variable, in which the data were centered and divided by the standard deviation. We created a RF regression model using an ensemble of 100 trees and employed one third of the total predictors for the decision split at each node. Predictor variables (features) were subsequently ranked based on their average impurity gain over all splits and all trees. Predictor variables with higher impurity gains contribute more to the variability in the prediction of the response variable, and hence are considered more important.

3. Results

3.1. Perfusion Cell Culture Experiments

We performed four perfusion cell cultures (Experiment A, B, C and D) using different VCD and PR set-point profiles, as illustrated in Figure 3. In Experiments A and B, we kept the VCD set-point constant for the entire duration of the cell cultures but shifted the PR set-point between day nine and ten (down in Experiment A and up in Experiment B). In Experiment C, we changed the VCD and PR set-points in a manner that maintained the CSPR at the same value. Finally, in Experiment D, we varied the VCD set-point while leaving the PR set-point constant. As shown in Figure 3, the VCD and PR followed the set-points very well with only minor deviations (also see Table S1). As the control of VCD was done only by bleeding (i.e., removal of cells), the VCD unsurprisingly tracked a decrease in the set-points better than an increase. Supplementary Figures S1–S4 give a summary of the cell culture parameters,

including VCD, PR, CSPR, the concentrations of glucose (Glc), lactose (Lac), ammonia (Amm) and IgG, and the cell-specific growth rate (μ), for all experiments. As shown in Figure 4, the cell-specific productivities of IgG, i.e., the secretion rate of IgG divided by the VCD, in all experiments follows a decreasing trend over the course of the cell cultivation. Correspondingly, the experimental secretion rates of the IgG glycoforms computed in the data pre-processing decrease with time (see Figure 5 and Supplementary Figure S5).

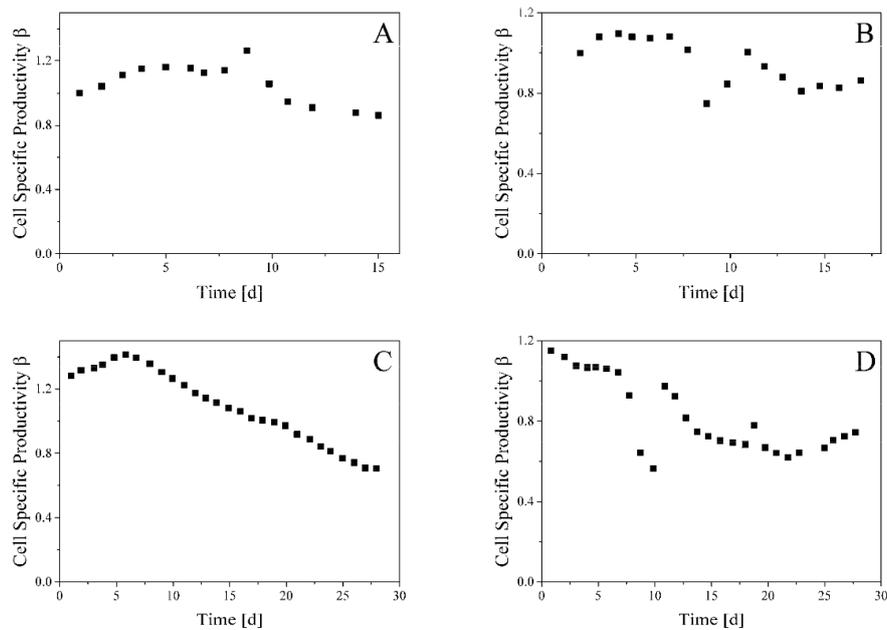


Figure 4. Cell-specific productivity. The cell-specific productivity generally decreases over the course of the four perfusion cell culture experiments (A), (B), (C) and (D).

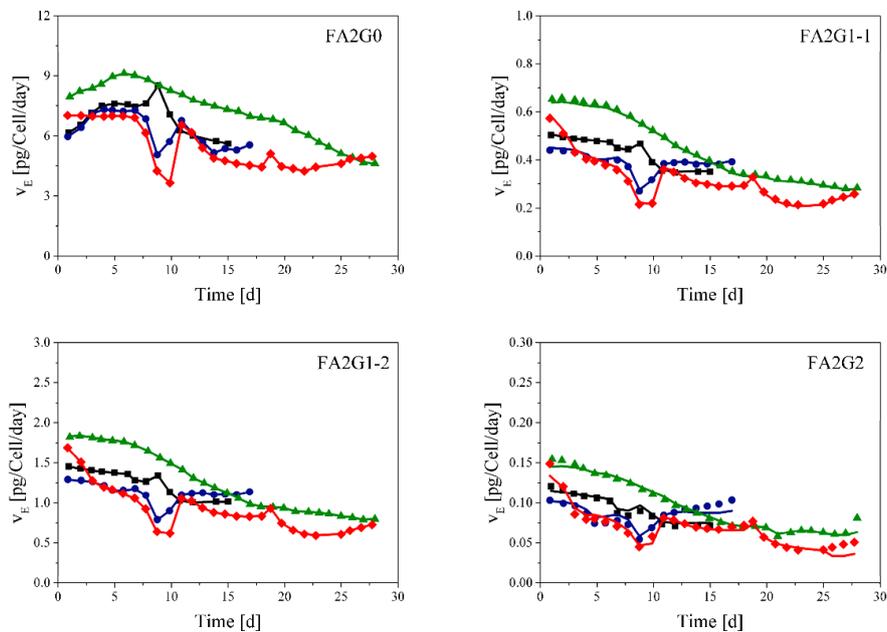


Figure 5. Secretion fluxes of the main IgG glycoforms. The solid symbols show the experimental secretion fluxes computed in the data preprocessing step, as outlined in Section 2.2 (Experiment A: black squares, Experiment B: blue circles, Experiment C: green triangle, Experiment D: red diamonds). The lines show the secretion fluxes from the fitting of α_j in the GFA, as outlined in Section 2.3.

3.2. Glycosylation Flux Analysis

For the GFA, we employed the IgG glycosylation reaction network depicted in Figure 6, which consists of 19 IgG glycoforms and 25 IgG glycosylation reactions. The glycosylation reaction network was based on a previously published network [41], where we omitted glycosylation reactions and molecules corresponding to IgG glycoforms that are not detected in our experiments and do not participate as intermediate species. Since each perfusion cell culture was started in the same manner, we used the same reference IgG glycosylation flux vector $\mathbf{v}_1^{\text{ref}}$ in the GFA of all experiments, more specifically using day one of experiment A as the reference time sample point (i.e., $\mathbf{v}_1^{\text{ref}}$ of day one in Experimental A is set to the vector of 1 s). Figure 5 shows the fitting of IgG glycoform secretion fluxes in the GFA for the major glycoforms (see Supplementary Figure S5 for the rest of IgG glycoforms).

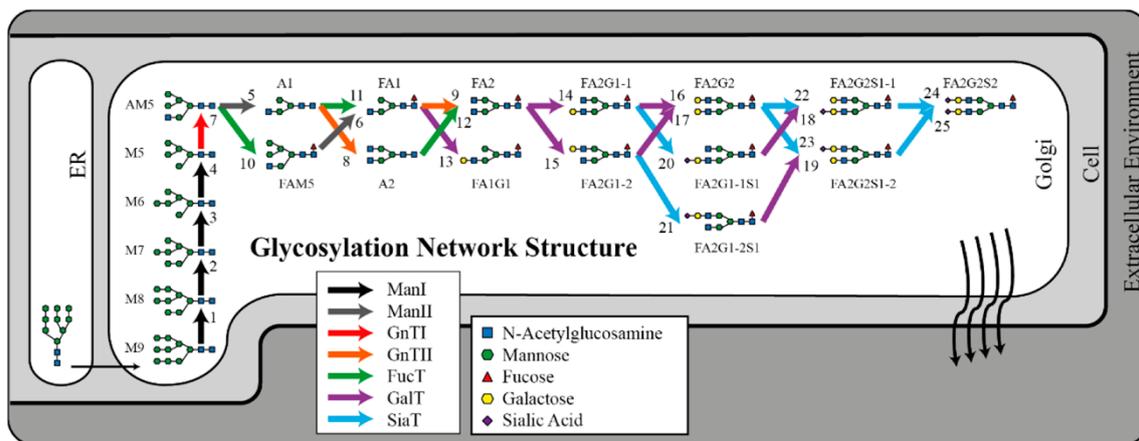


Figure 6. Glycosylation network for the GFA of immunoglobulin G in CHO-S. The enzyme names are abbreviated as follows: α -Mannosidase I and II (Man I/II), *N*-Acetylglucosaminyltransferase I and II (GnT I/II) and Fucosyltransferase (FucT), Galactosyltransferase (GalT) and Sialyltransferase (SiaT). The glycan labels are provided in Supplementary Table S2.

We compared the computational speed of the iterative GFA in analyzing the perfusion cell culture datasets with that of the original implementation. On a test computer (3.33 GHz Intel Xeon W3680, 18 GB RAM), the iterative GFA converged within 5 min using 250 random starting points. Of note, among the 250 random starts, 40% converged to the same minimum Φ value. Meanwhile, the original implementation of GFA using the global optimization toolbox MEIGO [42] required at least 42 min to converge. However, among ten repeated independent runs of MEIGO, only one converged to the same minimum Φ value as the iterative GFA, while the other runs converged to higher Φ values.

Figure 7 gives the time profiles of the enzyme-specific factors $\alpha_j(t)$ for each of the enzymes in the IgG glycosylation network. Most of the enzyme-specific factors, particularly those of α -Mannosidase I and II (Man I/II), *N*-Acetylglucosaminyltransferase I and II (GnT I/II) and Fucosyltransferase (FucT), maintain a constant activity level throughout the cell cultivation (i.e., $\alpha_j(t) \approx 1$). Meanwhile, the galactosyltransferase (GalT)-specific factor decreases during the beginning of the four experiments and varies with changes in the VCD and PR set-points. Since the fluxes catalyzed by sialyltransferase (SiaT) are close to zero, the estimate of the corresponding $\alpha_j(t)$ becomes unreliable due to high sensitivity to experimental noise, and thus is omitted from further analysis.

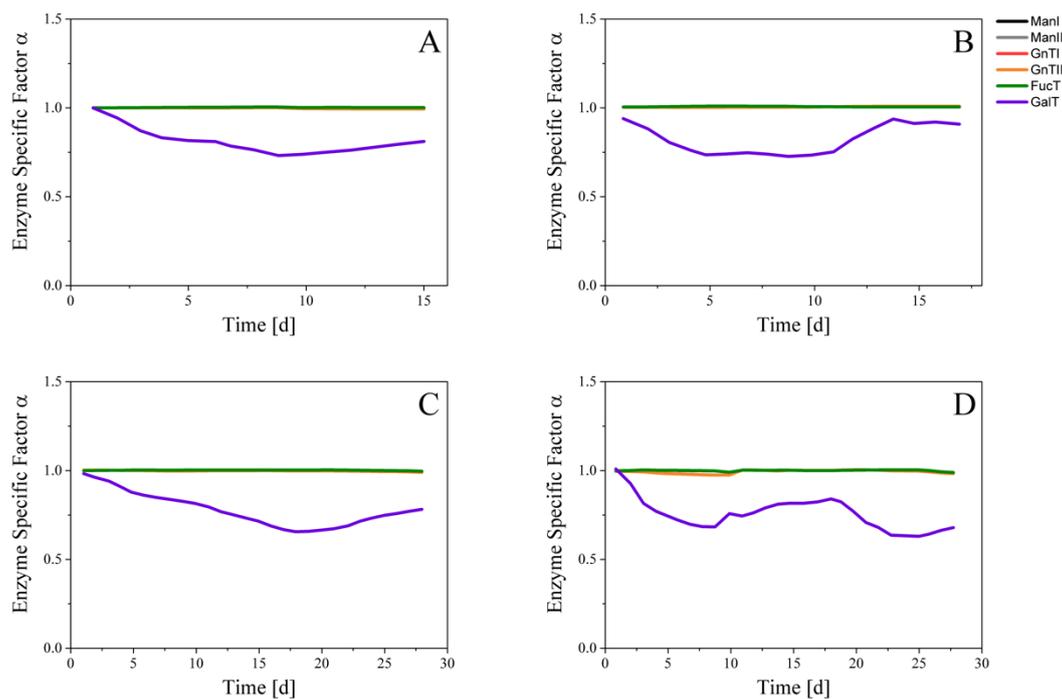


Figure 7. Predicted enzyme-specific factors. The activity of fluxes catalyzed by Man I (black), Man II (grey), GnT I (red), GnT II (orange) and FucT (green) remain relatively constant in all experiments. However, the fluxes catalyzed by GalT (purple) shows significant variation over the course of the four perfusion cell culture experiments (A), (B), (C) and (D).

3.3. Effects of Process Parameters on Glycosylation

As mentioned above, the enzyme-specific factor of IgG galactosylation fluxes displays the most dynamical change during the perfusion cell culture. However, the relationship between the changes in the enzyme-specific factor $\alpha_{GalT}(t)$ and the other process parameters is difficult to discern by a simple observation of the experimental data. For this reason, we employed a random forest regression analysis using the change of galactosyltransferase-specific factor over time $\frac{d\alpha_{GalT}(t)}{dt}$ as the response variable and using fourteen process parameters as the predictor variables (see Section 2.4). A RF regression model should be able to capture any nonlinear dependencies between the response and predictor variables. Here, we considered the following predictors: $\alpha_{GalT}(t)$, VCD, PR, B, CSPR, time and the concentrations of IgG, Glc, Lac and Amm, and the specific productivities of IgG, Glc, Lac and Amm (i.e., q_{IgG} , q_{Glc} , q_{Lac} and q_{Amm} , respectively). Furthermore, we excluded data from the startup period of the cell culture (i.e., days one to three of each experiment), as we were more interested in the regulation of IgG glycosylation during the steady state operations and setpoint changes of the perfusion cell culture. Finally, we ranked the predictor variables in decreasing magnitudes of the impurity gains. A higher impurity gain points to a predictor variable with higher importance in explaining the response variable.

Figure 8 gives the ranking of the predictor variables in decreasing impurity gains. The specific productivity of IgG (q_{IgG}) and the concentration of ammonia (Amm) are the two most important predictors of the dynamical changes in the galactosyltransferase specific activity. Indeed, when we repeated the RF regression using only q_{IgG} and Amm as the predictor variables, we observed a similar quality of data fitting to the response variables (see Supplementary Figure S6). Kolmogorov-Smirnov (KS) test and Wilcoxon rank sum test further confirmed that the residuals of the RF regression models using all 14 predictors and those using only q_{IgG} and Amm, are not statistically different (KS test p -value = 0.857; Wilcoxon rank sum test p -value = 0.824). Following q_{IgG} and Amm in the ranking are the glucose uptake rate (q_{GLC}) and bleed rate (B), indicating that the specific growth rate has a moderate contribution to the changes in the IgG galactosylation activity. The influence of growth

rate on IgG galactosylation is somewhat expected, as the nucleotide sugar UDP-Galactose is used for different forms of cellular glycosylation, not solely for IgG [43].

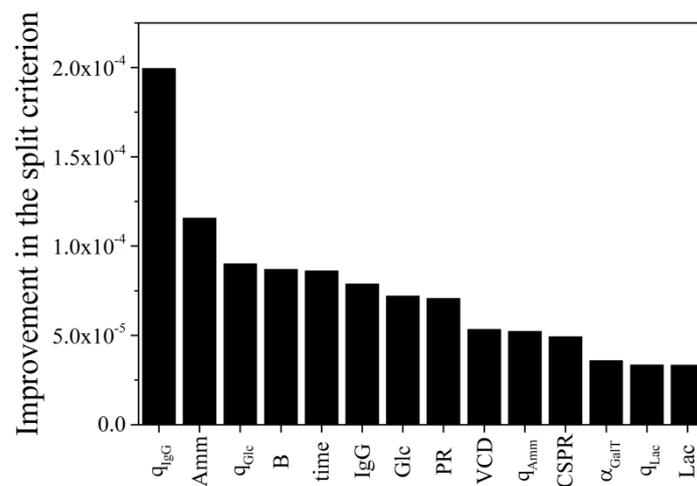


Figure 8. Ranking of predictors based on importance. The predictors are sorted in decreasing impurity gains as measured by the improvement in the split criterion.

4. Discussion

In this work, we improved the computational efficiency of the GFA and applied the GFA to analyze the IgG glycosylation of four perfusion CHO cell culture experiments. The GFA relies on a stoichiometric model of the IgG glycosylation network to estimate IgG intracellular glycosylation reaction fluxes from the secretion fluxes of IgG glycoforms [36]. The GFA assumes that the intracellular IgG glycosylation fluxes vary with time according to two multiplicative factors: α_j and β_j . The enzyme-specific factor α_j represents the relative change of the fraction of IgG that can be processed by a specific enzyme, while the cell-specific factor β_j describes the relative change of the specific productivity of IgG. In particular, α_j captures dynamical alterations in the intracellular IgG glycosylation fluxes that cannot be explained by increases or decreases in the amount of IgG entering the glycosylation network.

We caution against making an inference on the intracellular IgG glycosylation fluxes directly from the cell culture data—i.e., without using the appropriate mass and flux balances. For example, a larger fraction of an IgG glycoform could result from a higher secretion rate of the specific IgG glycoform or from lower secretion rates of the other glycoforms. Furthermore, as noted in Section 2.3, the flux balance around the cells in Equation (6) is underdetermined, implying that multiple solutions there exist for the intracellular IgG glycosylation fluxes for the same observed secretion fluxes. The GFA gets around such an issue by assuming that reactions catalyzed by the same enzyme vary with time commensurately. As demonstrated in the study, by quantifying dynamic intracellular glycosylation activities in an enzyme-specific manner, the GFA enables the identification of key process parameters that are indicative of and potentially contribute to the changes in the intracellular IgG glycosylation process.

In contrast to other approaches that rely on kinetic modeling of the glycosylation network, the GFA does not require specifying a large number of kinetic parameters, whose values are uncertain and likely system specific. However, the GFA requires time-series data for computing the secretion fluxes of different glycoforms, which include viable cell density, IgG titer and glycoform fractions. Such data may not always be available during the production of IgG. The GFA is also unable to provide mechanistic insights for the dynamic changes in the enzyme-specific factors α_j , at least not directly. In this study, we performed a post-analysis using a random forest regression to identify the key process parameters that are able to explain the variability in $\alpha_j(t)$.

The four perfusion cell cultures in this study differed among each other in the VCD and PR set-point profiles, which were designed with the goal of understanding how IgG productivity and

glycosylation vary with the set-points and process variables during steady-state operations. The cell culture conditions (i.e., media composition, seeding density, pH, temperature) of the four experiments were kept being as similar as possible, so that the effect of changes in operating set-points VCD and PR on the glycosylation process could be examined.

In our previous study using the same CHO cells, we reported that the cell-specific productivity of IgG in fed-batch cultivations increases with time, and that such an increase is associated with lower α_{jS} for the upstream enzymes ManI/II, GnTI/II and FucT [36]. We attributed the decreasing α_{jS} to the inability of these enzymes to process the increasing amount of IgG entering the intracellular glycosylation network, beyond the processing capacity of each enzyme. In contrast to fed-batch cultivations, the cell-specific productivity of IgG generally decreases over the course of the perfusion cell culture, as shown in Figure 4. The GFA shows that the enzyme-specific factors of the aforementioned upstream enzymes stay approximately constant at roughly 1 throughout the experiments. With α_j staying near 1, the IgG intracellular glycosylation fluxes associated with these enzymes thus vary proportionally with the cell-specific productivity $\beta(t)$. This trend is not surprising, considering that the amount of IgG that needs to be processed decreases over time.

The enzyme-specific factor with the most dynamical changes in the perfusion cell cultures corresponds to galactosyltransferase (GalT), which parallels that in the fed-batch cultivations. However, unlike the reduction in the GalT-specific factor $\alpha_{GalT}(t)$ over time during the fed-batch cultivations [36], $\alpha_{GalT}(t)$ dynamics in perfusion cell cultures showed both up and downward trends during the different steady-state operation periods, as shown in Figure 7. Because of the rich dynamics of $\alpha_{GalT}(t)$ in the perfusion cell culture, the relationship among the GalT-specific factor $\alpha_{GalT}(t)$ and process variables was not immediately obvious. By applying a random forest regression analysis, we identified the specific productivity of IgG and the concentration of ammonia as the two variables with the highest importance in explaining the dynamic changes of IgG galactosylation (see Figure 8). The accumulation of ammonia and its inhibitory activity on GalT have previously been reported as one of the main reasons for decreasing IgG galactosylation in fed-batch cultivations [36,44]. Unlike ammonia, the connection between the IgG specific productivity and galactosylation is not immediately clear. The partial correlation between $\frac{d\alpha_{GalT}(t)}{dt}$ and q_{IgG} is small and negative (partial correlation = -0.037), suggesting that the relationship between the two variables as revealed by RF analysis is likely to be nonlinear. The negative partial correlation further implies that keeping all other process parameters the same, a higher q_{IgG} is concomitant with decreasing α_{GalT} with time. Such a trend is in general agreement with how α_j of the upstream enzymes vary with the cell-specific productivity of IgG as explained earlier.

Taken together, the key process variables that are associated with IgG galactosylation in our perfusion cell cultures resemble those in the fed-batch cultivations of the same CHO cells. Note that the aforementioned similarity between fed-batch and perfusion cell cultures cannot be concluded by a direct observation of the process data and is only apparent through the combination of the GFA and random forest regression analysis. Such similarity suggests the possibility of using extensive data and knowledge from traditional batch/fed-batch production to inform or predict the IgG glycosylation in continuous perfusion cell cultures, which is a topic of interest in our research groups. In a separate publication, we noted comparable IgG glycoform patterns between our perfusion cell culture process at steady state and the fed-batch cultivation beyond the stationary growth phase [15]. In the literature, past studies produced contradictory conclusions on the IgG glycosylation when comparing the two cell culture modes, where some studies demonstrated similarities [9,10] and others showed discrepancies in the IgG glycan structures [13,14]. A recent report of a head-to-head comparison between perfusion and fed-batch cell cultures noted a strong similarity in the IgG glycosylation dynamics [45], an observation that aligns well with the result of our analysis.

5. Conclusions

The application of glycosylation flux analysis to IgG production data from four perfusion CHO cell culture experiments shed light on the possible controlling factors of IgG glycosylation. The activity of galactosyltransferase and thus intracellular IgG galactosylation fluxes displayed the most time-varying changes during the perfusion cell cultivation. The GFA, coupled with a random forest regression analysis, pointed to the cell-specific productivity of IgG and the concentration of ammonia in the media as the most important process parameters for explaining the dynamical alterations in GalT activity. These observations parallel our previous report on IgG glycosylation in fed-batch production. The similarity between fed-batch and perfusion cell culture, especially in how IgG galactosylation flux activity varies with the cell culture parameters, suggests the possibility of using extensive data and knowledge from batch/fed-batch production to predict the dynamics of IgG glycosylation in the continuous perfusion cell cultures.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2227-9717/6/10/176/s1>, Table S1: Summary of the reactor set points and the corresponding measured values, Table S2: Glycan structures in glycosylation network, Figure S1: Process measurements of Experiment A, Figure S2: Process measurements of Experiment B, Figure S3: Process measurements of Experiment C, Figure S4: Process measurements of Experiment D, Figure S5: Secretion flux fitting of other IgG glycoforms, Figure S6: Residuals of random forest regression.

Author Contributions: S.H., M.W., M.M. and R.G. conceived and designed the study; S.H. performed the computational analysis; M.W. performed the wet lab experiments; N.P.G. performed the random forest regression analysis; S.H., D.L., T.S. and R.G. developed the improved GFA; M.M. supervised the wet lab experiments; S.H., M.W. and R.G. wrote the paper.

Funding: This research was funded by Swiss Economic Ministry KTI (CTI) program grant number 19190.2 PFIW—IW and by ETH Zurich.

Acknowledgments: The authors would like to thank Ernesto Scibona and Thomas Villiger for their help and explanations regarding the glycosylation process.

Conflicts of Interest: No conflicts of interest are declared.

References

1. Ecker, D.M.; Jones, S.D.; Levine, H.L. The therapeutic monoclonal antibody market. *MAbs* **2015**, *7*, 9–14. [[CrossRef](#)] [[PubMed](#)]
2. Kelley, B. Industrialization of MAb production technology. *MAbs* **2009**, *1*, 443–452. [[CrossRef](#)] [[PubMed](#)]
3. Li, F.; Vijayasankaran, N.; Shen, A.; Kiss, R.; Amanullah, A. Cell culture processes for monoclonal antibody production. *MAbs* **2010**, *2*, 466–479. [[CrossRef](#)] [[PubMed](#)]
4. Rathore, A.S. Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends Biotechnol.* **2009**, *27*, 546–553. [[CrossRef](#)] [[PubMed](#)]
5. Food and Drug Administration. *FDA Guidance for Industry. PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*; Food and Drug Administration: Rockville, MD, USA, 2004.
6. Reay, D.; Ramshaw, C.; Harvey, A. *Process Intensification: Engineering for Efficiency, Sustainability and Flexibility*, 2nd ed.; Elsevier: New York, NY, USA, 2013.
7. Boedeker, B.G.D. Recombinant factor VIII (Kogenate[®]) for the treatment of Hemophilia A: The first and only world-wide licensed recombinant protein produced in high-throughput perfusion culture. In *Modern Biopharmaceuticals*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2013; pp. 429–443.
8. Clincke, M.F.; Molleryd, C.; Zhang, Y.; Lindskog, E.; Walsh, K.; Chotteau, V. Very high density of CHO cells in perfusion by ATF or TFF in WAVE bioreactorTM, Part I: Effect of the cell density on the process. *Biotechnol. Prog.* **2013**, *29*, 754–767. [[CrossRef](#)] [[PubMed](#)]
9. Meuwly, F.; Weber, U.; Ziegler, T.; Gervais, A.; Mastrangeli, R.; Crisci, C.; Rossi, M.; Bernard, A.; von Stockar, U.; Kadouri, A. Conversion of a CHO cell culture process from perfusion to fed-batch technology without altering product quality. *J. Biotechnol.* **2006**, *123*, 106–116. [[CrossRef](#)] [[PubMed](#)]

10. Lee, S.-Y.; Kwon, Y.-B.; Cho, J.-M.; Park, K.-H.; Chang, S.-J.; Kim, D.-I. Effect of process change from perfusion to fed-batch on product comparability for biosimilar monoclonal antibody. *Process Biochem.* **2012**, *47*, 1411–1418. [[CrossRef](#)]
11. Ryll, T.; Dutina, G.; Reyes, A.; Gunson, J.; Krummen, L.; Etcheverry, T. Performance of small-scale CHO perfusion cultures using an acoustic cell filtration device for cell retention: Characterization of separation efficiency and impact of perfusion on product quality. *Biotechnol. Bioeng.* **2000**, *69*, 440–449. [[CrossRef](#)]
12. Lüllau, E.; Kanttinen, A.; Hassel, J.; Berg, M.; Haag-Alvarsson, A.; Cederbrant, K.; Greenberg, B.; Fenge, C.; Schweikart, F. Comparison of Batch and Perfusion Culture in Combination with Pilot-Scale Expanded Bed Purification for the Production of Soluble Recombinant β -Secretase. *Biotechnol. Prog.* **2003**, *19*, 37–44. [[CrossRef](#)] [[PubMed](#)]
13. Lipscomb, M.L.; Palomares, L.A.; Hernández, V.; Ramírez, O.T.; Kompala, D.S. Effect of production method and gene amplification on the glycosylation pattern of a secreted reporter protein in CHO cells. *Biotechnol. Prog.* **2005**, *21*, 40–49. [[CrossRef](#)] [[PubMed](#)]
14. Zhuang, C.; Zheng, C.; Chen, Y.; Huang, Z.; Wang, Y.; Fu, Q.; Zeng, C.; Wu, T.; Yang, L.; Qi, N. Different fermentation processes produced variants of an anti-CD52 monoclonal antibody that have divergent in vitro and in vivo characteristics. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 5997–6006. [[CrossRef](#)] [[PubMed](#)]
15. Karst, D.J.; Steinebach, F.; Morbidelli, M. Continuous integrated manufacturing of therapeutic proteins. *Curr. Opin. Biotechnol.* **2018**, *53*, 76–84. [[CrossRef](#)] [[PubMed](#)]
16. Berger, M.; Kaup, M.; Blanchard, V. Protein glycosylation and its impact on biotechnology. In *Genomics and Systems Biology of Mammalian Cell Culture*; Hu, W.S., Zeng, A.-P., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 165–185.
17. Aebi, M. N-linked protein glycosylation in the ER. *Biochim. Biophys. Acta-Mol. Cell Res.* **2013**, *1833*, 2430–2437. [[CrossRef](#)] [[PubMed](#)]
18. Solá, R.J.; Griebenow, K. Glycosylation of therapeutic proteins: An effective strategy to optimize efficacy. *BioDrugs* **2010**, *24*, 9–21. [[CrossRef](#)] [[PubMed](#)]
19. Jefferis, R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev.* **2009**, *8*, 226–234. [[CrossRef](#)] [[PubMed](#)]
20. Goh, J.S.Y.; Liu, Y.; Liu, H.; Chan, K.F.; Wan, C.; Teo, G.; Zhou, X.; Xie, F.; Zhang, P.; Zhang, Y.; et al. Highly sialylated recombinant human erythropoietin production in large-scale perfusion bioreactor utilizing CHO-gmt4 (JW152) with restored GnT I function. *Biotechnol. J.* **2014**, *9*, 100–109. [[CrossRef](#)] [[PubMed](#)]
21. Harding, F.A.; Stickler, M.M.; Razo, J.; DuBridg, R.B. The immunogenicity of humanized and fully human antibodies: Residual immunogenicity resides in the CDR regions. *MAbs* **2010**, *2*, 256–265. [[CrossRef](#)] [[PubMed](#)]
22. Matasci, M.; Hacker, D.L.; Baldi, L.; Wurm, F.M. Protein therapeutics Recombinant therapeutic protein production in cultivated mammalian cells: Current status and future prospects. *Drug Discov. Today Technol.* **2008**, *5*, 37–42. [[CrossRef](#)] [[PubMed](#)]
23. Jayapal, K.P.; Wlaschin, K.F.; Hu, W.-S.; Yap, M.G.S. Recombinant protein therapeutics from CHO cells—20 years and counting. *Chem. Eng. Prog.* **2007**, *103*, 40–47.
24. Wright, A.; Morrison, S.L. Effect of glycosylation on antibody function: Implications for genetic engineering. *Trends Biotechnol.* **1997**, *15*, 26–32. [[CrossRef](#)]
25. Moremen, K.W.; Ramiah, A.; Stuart, M.; Steel, J.; Meng, L.; Forouhar, F.; Moniz, H.A.; Gahlay, G.; Gao, Z.; Chapla, D.; et al. Expression system for structural and functional studies of human glycosylation enzymes. *Nat. Chem. Biol.* **2017**, *14*, 156–162. [[CrossRef](#)] [[PubMed](#)]
26. Radhakrishnan, D.; Robinson, A.S.; Ogunnaike, B.A. Controlling the glycosylation profile in mAbs using time-dependent media supplementation. *Antibodies* **2018**, *7*, 1. [[CrossRef](#)]
27. Ivarsson, M.; Villiger, T.K.; Morbidelli, M.; Soos, M. Evaluating the impact of cell culture process parameters on monoclonal antibody N-glycosylation. *J. Biotechnol.* **2014**, *188*, 88–96. [[CrossRef](#)] [[PubMed](#)]
28. Varki, A. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology* **1993**, *3*, 97–130. [[CrossRef](#)] [[PubMed](#)]
29. Umaña, P.; Bailey, J.E. A Mathematical model of N-linked glycoform biosynthesis. *Biotechnol. Bioeng.* **1997**, *55*, 890–908. [[CrossRef](#)]

30. Krambeck, F.J.; Bennun, S.V.; Narang, S.; Choi, S.; Yarema, K.J.; Betenbaugh, M.J. A mathematical model to derive N-glycan structures and cellular enzyme activities from mass spectrometric data. *Glycobiology* **2009**, *19*, 1163–1175. [[CrossRef](#)] [[PubMed](#)]
31. Jimenez del Val, I.; 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. Prog.* **2011**, *27*, 1730–1743. [[CrossRef](#)] [[PubMed](#)]
32. Jiménez del Val, I.; Constantinou, A.; Dell, A.; Haslam, S.; Polizzi, K.M.; Kontoravdi, C. A quantitative and mechanistic model for monoclonal antibody glycosylation as a function of nutrient availability during cell culture. *BMC Proc.* **2013**, *7*, O10. [[CrossRef](#)]
33. Jedrzejewski, P.M.; Jiménez del Val, I.; Constantinou, A.; Dell, A.; Haslam, S.M.; Polizzi, K.M.; Kontoravdi, C. Towards controlling the glycoform: A model framework linking extracellular metabolites to antibody glycosylation. *Int. J. Mol. Sci.* **2014**, *15*, 4492–4522. [[CrossRef](#)] [[PubMed](#)]
34. Spahn, P.N.; Hansen, A.H.; Henning, G.; Arnsdorf, J.; Kildegaard, H.F.; Lewis, N.E.; Arnsdorf, J.; Kildegaard, H.F.; Lewis, N.E.; Markov, A. A markov chain model for N-linked protein glycosylation—Towards a low-parameter tool for model-driven. *Metab. Eng.* **2015**, *33*, 52–66. [[CrossRef](#)] [[PubMed](#)]
35. Spahn, P.N.; Hansen, A.H.; Kol, S.; Voldborg, B.; Lewis, N.E. Predictive glycoengineering of biosimilars using a Markov chain glycosylation model. *Biotechnol. J.* **2017**, *12*, 1–8. [[CrossRef](#)] [[PubMed](#)]
36. Hutter, S.; Villiger, T.K.; Brühlmann, D.; Stettler, M.; Broly, H.; Soos, M.; Gunawan, R. Glycosylation flux analysis reveals dynamic changes of intracellular glycosylation flux distribution in Chinese hamster ovary fed-batch cultures. *Metab. Eng.* **2017**, *43*, 9–20. [[CrossRef](#)] [[PubMed](#)]
37. Antoniewicz, M.R. Methods and advances in metabolic flux analysis: A mini-review. *J. Ind. Microbiol. Biotechnol.* **2015**, *42*, 317–325. [[CrossRef](#)] [[PubMed](#)]
38. Wolf, M. Development and Optimization of Mammalian Cell Perfusion Cultures for Continuous Biomanufacturing. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2018.
39. Karst, D.J.; Serra, E.; Villiger, T.K.; Soos, M.; Morbidelli, M. Characterization and comparison of ATF and TFF in stirred bioreactors for continuous mammalian cell culture processes. *Biochem. Eng. J.* **2016**, *110*, 17–26. [[CrossRef](#)]
40. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
41. Villiger, T.K.; Scibona, E.; Stettler, M.; Broly, H.; Morbidelli, M.; Soos, M. Controlling the time evolution of mAb N-linked glycosylation—Part II: Model-based predictions. *Biotechnol. Prog.* **2016**, *32*, 1135–1148. [[CrossRef](#)] [[PubMed](#)]
42. Egea, J.A.; Henriques, D.; Cokelaer, T.; Villaverde, A.F.; MacNamara, A.; Danciu, D.-P.; Banga, J.R.; Saez-Rodriguez, J. MEIGO: An open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. *BMC Bioinform.* **2014**, *15*, 136. [[CrossRef](#)] [[PubMed](#)]
43. Del Val, I.J.; Polizzi, K.M.; Kontoravdi, C. A theoretical estimate for nucleotide sugar demand towards Chinese Hamster Ovary cellular glycosylation. *Sci. Rep.* **2016**, *6*, 28547. [[CrossRef](#)] [[PubMed](#)]
44. Gawlitzek, M.; Ryll, T.; Lofgren, J.; Sliwkowski, M.B. Ammonium alters N-glycan structures of recombinant TNFR-IgG: Degradative versus biosynthetic mechanisms. *Biotechnol. Bioeng.* **2000**, *68*, 637–646. [[CrossRef](#)]
45. Walther, J.; Lu, J.; Hollenbach, M.; Yu, M.; Hwang, C.; McLarty, J.; Brower, K. Perfusion cell culture decreases process and product heterogeneity in a head-to-head comparison with fed-batch. *Biotechnol. J.* **2018**, e1700733. [[CrossRef](#)] [[PubMed](#)]

