# A System Identification Approach for Developing Model Predictive Controllers of Antibody Quality Attributes in Cell Culture Processes

**Brandon Downey**
Capsugel, R&D, Bend, OR 97702

**John Schmitt**
Capsugel, R&D, Bend, OR 97702

**Justin Beller**
Capsugel, R&D, Bend, OR 97702

**Brian Russell**
Capsugel, R&D, Bend, OR 97702

**Anthony Quach**
Capsugel, R&D, Bend, OR 97702

**Elizabeth Hermann**
Capsugel, R&D, Bend, OR 97702

**David Lyon**
Capsugel, R&D, Bend, OR 97702

**Jeffrey Breit**
Capsugel, R&D, Bend, OR 97702

*DOI 10.1002/btpr.2537*
*Published online August 24, 2017 in Wiley Online Library (wileyonlinelibrary.com)*

*As the biopharmaceutical industry evolves to include more diverse protein formats and processes, more robust control of Critical Quality Attributes (CQAs) is needed to maintain processing flexibility without compromising quality. Active control of CQAs has been demonstrated using model predictive control techniques, which allow development of processes which are robust against disturbances associated with raw material variability and other potentially flexible operating conditions. Wide adoption of model predictive control in biopharmaceutical cell culture processes has been hampered, however, in part due to the large amount of data and expertise required to make a predictive model of controlled CQAs, a requirement for model predictive control. Here we developed a highly automated, perfusion apparatus to systematically and efficiently generate predictive models using application of system identification approaches. We successfully created a predictive model of %galactosylation using data obtained by manipulating galactose concentration in the perfusion apparatus in serialized step change experiments. We then demonstrated the use of the model in a model predictive controller in a simulated control scenario to successfully achieve a %galactosylation set point in a simulated fed-batch culture. The automated model identification approach demonstrated here can potentially be generalized to many CQAs, and could be a more efficient, faster, and highly automated alternative to batch experiments for developing predictive models in cell culture processes, and allow the wider adoption of model predictive control in biopharmaceutical processes.* © 2017 The Authors Biotechnology Progress published by Wiley Periodicals, Inc. on behalf of American Institute of Chemical Engineers *Biotechnol. Prog.*, 33:1647–1661, 2017
*Keywords: model predictive control, glycosylation, CHO, system identification, perfusion*

## Introduction

The production of biopharmaceuticals requires meeting high standards of purity, consistency, and quality to ensure the safety and effectiveness of products for patients. To maintain these high standards, regulatory agencies (in collaboration with industry) continue to adopt improved process

Additional Supporting information may be found in the online version of this article.

Correspondence concerning this article should be addressed to B. Downey at brandon.downey@capsugel.com.

filing approaches, including Quality by Design (QbD).[1–3] As a consequence of the development of better technologies and quality approaches, the expectations for demonstrating control of the quality of biopharmaceutical products has increased.

One goal of QbD is to lower product variability and increase product quality by increasing control of the production process and achieving a Quality Target Product Profile (QTPP). A common approach for controlling a QTPP is to develop a production process in which product Critical Quality Attributes (CQAs) are maintained within a predefined range by controlling Critical Process Parameters (CPPs). Because the production of biopharmaceutical products occurs via recombinant gene expression in cell hosts, many of the CQAs in these products are dictated by the state of the cell host in the cell culture production process.[4] CPPs are typically defined experimentally in the cell culture process by systematically manipulating recipe variables, such as media components or other culture conditions, and observing the ultimate outcome on the expressed protein CQAs. The experimental identification of CPPs requires a substantial use of resources prior to regulatory filing, and represents an area for possible acceleration of the biopharmaceutical commercialization process.

The biotherapeutic market is undergoing a transformation from the commercialization of predominantly large market "blockbuster" antibody products, to smaller volume high value "niche" products. In addition, increased expiration of patent protection has increased the development of low margin biosimiliars. This market trend means companies have more candidate products in development with more diverse protein formats. Increasing quality expectations, a greater diversity of products, and a demand for higher throughput in today's biopharmaceuticals market creates a unique challenge for the development and manufacturing of new products. Producing biotherapeutics that meet expectations for increases in quality across diverse protein formats pushes development organizations to quickly develop processes, manufacture clinical trial material, and provide tightly controlled product quality attributes for an increasing number of product candidates.

### *Product attribute control can allow more robust control of CQAs*

A routine approach in industry to determine how to control CPPs to achieve CQAs is to execute cell culture processes according to a predefined, tightly controlled recipe. This recipe (commonly including set points for CPPs, media/feed compositions, and feeding strategies) is determined from laboratory scale experiments, and is selected based on the recipe's ability (in conjunction with tight controls on raw material variation) to yield an acceptably low degree of variation in the product CQAs (among other process considerations, like product titer). This method of control is commonly referred to as open-loop control, where CPPs are set in the absence of any in-process measurement of product CQAs. Increasingly, emphasis is being placed on developing closed loop control strategies, where CPPs are actively manipulated during the process within a defined range to achieve the desired CQAs. This active control methodology has been called product attribute control.[5] In the context of QbD, product attribute control may allow development of processes with tighter CQAs, and thus greater control over the QTPP.

In addition to the identified CPPs, other complex interactions can impact a given CQA in a cell culture process (e.g., raw material variations, changes in cell metabolic state, dilution rates in perfusion). In addition, measurement of many CQAs is time-consuming, making frequent in-process measurement difficult. Because of this complexity, a model predictive controller is often the controller of choice for achieving product attribute control in cell culture and fermentation processes.[6,7] To develop a model predictive controller, a predictive model of how a given change in a culture variable affects the desired CQA over time must be developed. These models are typically constructed by developing some expert understanding of the fundamental underlying processes that lead to changes in the CQA, aided with directed experimentation. This sort of development usually yields a fundamentals based model that approximates the observed dynamics of the CQA in response to changes in the CPP. This model development methodology currently requires some understanding of the fundamental dynamics of the system by trained experts, and has been successfully demonstrated.[5]

### *Limitations of batch DoE approach to relate CPPs to CQAs*

Although the successful application of product attribute control has been demonstrated, the process of defining CPPs and constructing a control strategy for CQAs remains a time-consuming endeavor. Commonly, process optimization is carried out through many manually executed experiments (usually in a batch DoE format) followed by evaluation by expert-level scientists to understand the implications for process and media optimization. This approach takes a considerable amount of time and effort, and often produces data that is not ideal for developing a product attribute control scheme.

For instance, DoE style experimentation does not allow process dynamics that are related to product quality attributes to be deconvoluted from those which are unrelated in a time-varying system. Additionally, batch DoE experimental methods also rely on disparate sets of experiments to identify optimal process parameters, which requires many experiments to gather sufficient data to develop an effective control strategy. Furthermore, due to the amount of analytical effort typically required to take relevant product quality measurements, experimental design typically prioritizes minimized analytical measurements and leads to an insufficient amount of data generated. This insufficiency does not allow prediction of the effect of process parameters on those measurements, making control strategy design difficult. Finally, creation of a predictive model useful for control based on these datasets typically requires fundamental *a priori* knowledge of the underlying system, which does not allow for a reliable, systematic approach to be applied across a breadth of CQAs for development of model predictive controllers.

Taken together, this commonly used process development methodology, although usually ultimately effective, often makes it difficult and time-consuming to develop a robust control strategy. There remains a need for the application of a more systematic development approach in order to meet the challenge of developing a higher number of higher quality products on more diverse protein platforms.

### *Using system identification in automated perfusion apparatus to systematically relate CPPs to CQAs*

To address these challenges, we implemented a methodology for efficiently relating CPPs to product CQAs for development of model predictive controllers using an automated cell culture platform. This approach systematically subjects cells grown in a pseudo-steady-state perfusion apparatus to perturbations in CPPs, such as the concentration of feed components, and records the changes in the CQAs over time. This approach allows creation of predictive models predicting the effect of CPPs on CQAs and subsequent development of model predictive controllers in an automated, systematic fashion. This systematic approach is in principle generalizable to a diverse set of CPPs and CQAs, making it particularly well suited for an agile process development environment.

The strategy of using experiments wherein input variables are manipulated in a prescribed fashion to specifically determine the dynamic input-output relationship between CPPs and CQAs is known as system identification.[8] System identification approaches have the benefit of: requiring limited fundamental knowledge of the system *a priori,* reducing the total number of experiments required via simultaneous variation of multiple inputs, being systematic and amenable to automation, and are in principle generalizable to a diverse set of inputs and outputs. In addition, Systems ID approaches are designed to efficiently generate predictive models of processes that can be used for control. System identification approaches are applied successfully in many industrial applications including hybrid car battery management, unmanned flight, and supply chain management.[9–11]

Although system identification approaches have been applied successfully to build control strategies in many applications, application of these approaches to cell culture processes presents unique challenges including: the highly multivariate nature of cell culture processes, difficulty in measuring the CQAs of interest, and potential dynamic behavior of the processes which might limit the application of the technique to cell culture.

Systems identification approaches have previously been applied for batch processes.[12,13] However, cell culture processes are especially multivariate. In a typical batch experiment, a multitude of factors (metabolite concentrations, cell states, cell density, etc…) are all changing with time as the cells constantly consume the nutrients of the media and produce byproducts. This makes it difficult to observe the effects of an independently manipulated variable on the product in isolation of other changing variables. Additionally, due to the long time scales involved with mammalian cell culture processes (a typical run lasts about 2 weeks), gathering the amount of data required for application of a system identification approach would take a prohibitively long time.

The difficulty in measuring CQAs arises from the fact that typical CQA measurement methods for protein therapeutics often require a series of steps (separation of the cells from the culture broth, separation of the protein from the broth components, followed by additional digestions and dilutions) before ultimate measurement of the CQA. This measurement difficulty limits the amount of data that can be gathered practically in a system identification experiment. Additionally, due to the long timescales involved with mammalian cell culture, sampling must occur at very inconvenient times (such as the middle of the night).

Finally, the degree to which nonlinearities, time-variance, and hysteresis are dominant effects in the underlying processes is unknown. Each of these effects require additional model complexity to adequately represent the input-output relationship between CPPs and CQAs and could limit the successful application of the technique to cell culture

To address these challenges, we designed and constructed an automated, pseudo-steady-state apparatus based on perfusion technology to execute system identification experiments.

The pseudo-steady-state perfusion culture allows independent manipulation of input CPPs (like concentrations of media components) in a fashion that is more independent of the highly multivariate nature of the cell culture process, thereby greatly simplifying the detection of the effect of each CPP on the output CQA. The pseudo-steady-state apparatus also allows multiple manipulations of each CPP to be carried out in sequence, while not having to complete an entire batch process to detect the ultimate effect of the CPP on the endpoint CQA. Additionally, due to the fact that the culture is at a pseudo-steady-state, non-linearities, time-variance, and hysteresis effects can be detected by varying the order of manipulations of CPPs and looking for changes in outcome of CQAs over time.

Additionally, the perfusion system was automated to allow the increased data generation required by system identification techniques. Specifically, cell containing and cell free samples were sampled automatically using an aseptic autosampler to deliver samples to a liquid handling robot that automated the sample preparation. In addition, this system was designed to store samples at stable conditions during the night so human sampling was not required at inconvenient times.

### *Experimental design and rationale*

In this study, we sought to determine whether a system identification experiment performed on a pseudo-steady-state apparatus could be used to create a predictive model of a CQA given an arbitrary CPP manipulation in a fed-batch system.

Because of the importance of *N*-glycosylation as an important effector of the half-life, immunogenicity, and secondary effector function of therapeutic proteins,[14,15] we utilized *N*-glycan galactosylation as a model quality attribute. A systems identification approach for generating a predictive model of galactosylation was applied to a model Chinese Hamster Ovary (CHO) cell line grown in perfusion culture by manipulating the galactose concentration in the extracellular medium and recording the resulting galactosylation of the expressed antibody.

In the work presented here, we hypothesize that the emergent dynamics of a model CHO cell culture system are simple enough to yield quantifiable process behavior that can be used to construct a predictive model of process outputs such that the model can be used for model predictive control. To test the hypothesis, we used a pseudo-steady-state apparatus with a CHO cell line expressing a glycosylated monoclonal antibody (mAb) and subjected the culture to step-increases in galactose feed concentration. We then developed a predictive model of *N*-glycan galactosylation vs. galactose feed concentration and implemented the model as part of a model predictive controller in simulation. The performance of the controller was evaluated under simulated fed-batch conditions.

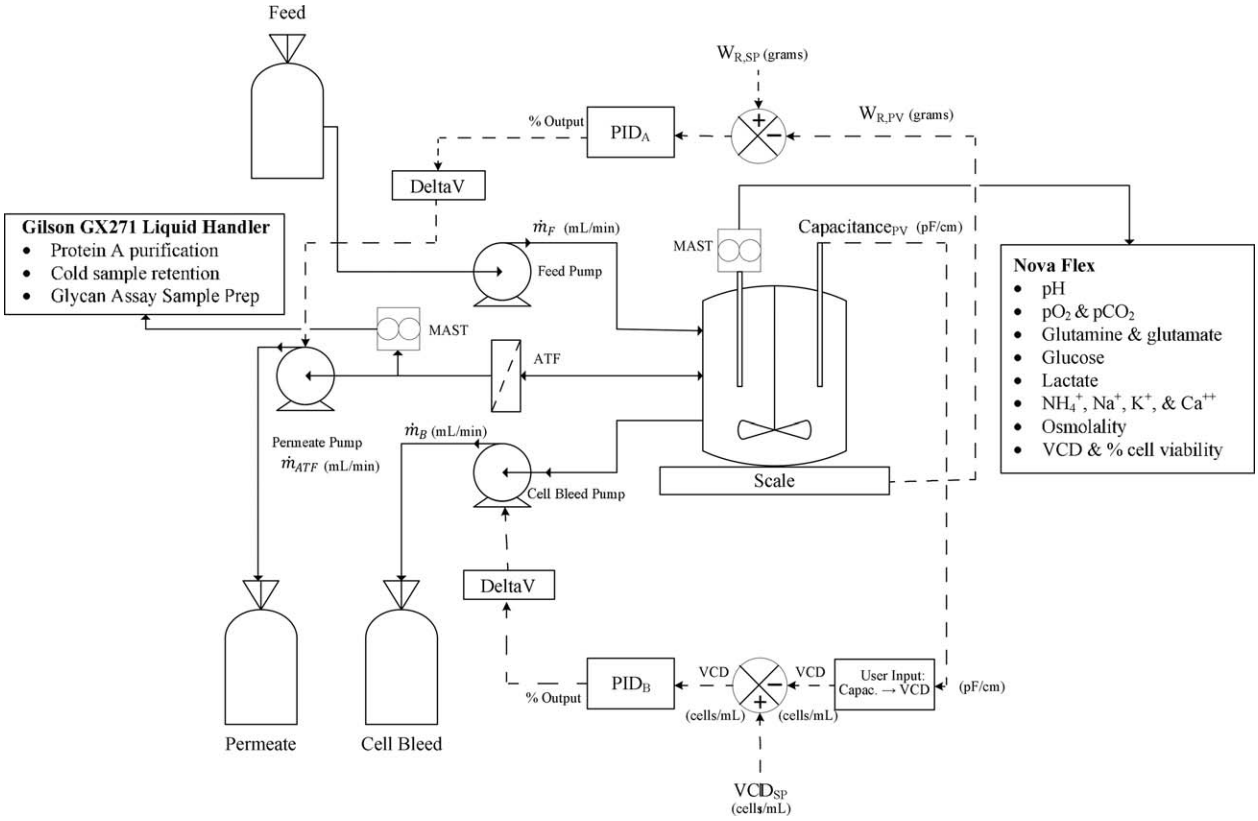


**Figure 1. Schematic representation of automated perfusion system used in this study.**

## Materials and Methods

### *Cell culture*

The CHOZN® GS−/− cell line producing a monoclonal antibody against the Rabies virus was used (MilliporeSigma; St Louis., MO). The CHO cell line was maintained in shake flask prior to inoculation of a bioreactor. The maintenance media for cell growth in the shake flasks was CD CHO Fusion (MilliporeSigma; St. Louis MO). Cells were grown to a density of $3 \times 10^6$ cells $mL^{-1}$ and then passaged into a new shake flask at $2 \times 10^5$ cells $mL^{-1}$. At the time of inoculation for bioreactor studies, bioreactors were inoculated at a density of $3 \times 10^5$ cells $mL^{-1}$. Media and feed of these reactors were comprised of the Ex-Cell® Advanced™ CHO Fed Batch System (MIlliporeSigma, St. Louis, MO).

### *Perfusion reactor*

Perfusion cultures were grown in 2-L glass bioreactors (Broadley-James, Irvine, CA) with a 1-L working volume. Perfusion bioreactors were equipped with a peripheral alternating tangential flow filter (ATF), containing a 0.2-$\mu$m pore cell-retention device (Refine Technology, Waltham, MA). Operational conditions were as follows: temperature 37°C ± 0.1°C, dissolved oxygen 30% ± 5% air saturation, agitation 138 ± 5 rpm, pitched blade impeller. Once the viable cell density of the culture reached $5.0 \times 10^6$ cells $mL^{-1}$ perfusion was initiated. During perfusion, a chemically defined nutrient feed, a mixture of 90% Ex-Cell® Advanced™ CHO Fed Batch System Media and 10% Ex-Cell® Advanced™ CHO Fed Batch System Feed was fed at a rate of 0.21 nL $cell^{-1}$ $day^{-1}$ (1065 mL $day^{-1}$). The media/feed blend ratio was determined based on the volumetric composition of the media equivalent to day 5 of a model fed-batch cell culture process (i.e., one 10 vol% bolus feed has been added by day 5). The perfusion rate was chosen to simulate the relative nutrient abundance present on day 5 of a model fed batch culture. This was accomplished using a previously described method.[16] Specifically the "equivalent CSPR" was calculated and used as the perfusion rate.

### *Cell density control of perfusion reactor*

Cell density was maintained at $5.0 \times 10^6$ cells $mL^{-1}$ for the duration of the experiment via control by dielectric spectroscopy, measured using a frequency scanning capacitance probe (Aber Futura, Aber Instruments, Aberystwyth, UK). The viable cell density was chosen to simulate a typical day 5 value observed in a model fed-batch cell culture process. In conjunction, the reactor volume was maintained at 1,000 mL via the reactor scale. A linear calibration curve between capacitance measured at 607 kHz and viable cell density was constructed for the cell line using data obtained during the growth phase of the culture, based on the method described here.[17] The probe was zeroed in media without cells prior to inoculation. Viable cell density was monitored and periodically checked against values predicted by capacitance. Very little drift in the calibration was observed during the course of the experiments.

Two PID loops were constructed to control the reactor volume and VCD. Figure 1 shows a schematic of the perfusion system. $PID_A$ represents the loop that controls the reactor volume. Because the nutrient feed is fed at a constant rate into the reactor, only the cell-free permeate and cell bleed pumps can be manipulated to hold reactor volume constant. $PID_A$ controls the overall reactor volume by measuring vessel weight and manipulating the permeate pump rate to achieve set point. Simultaneously, $PID_B$ controls the VCD

value predicted by the capacitance readout by manipulating the cell bleed pump rate to maintain set point.

### *Fed batch bioreactor operation*

Fed batch cultures were used to generate validation data sets to test the model generated from perfusion experiments. Cells were cultured in 3-L Broadley–James reactors with a 2-L working volume. Operating conditions were identical to the perfusion system, as described above. The culture was fed 10% v/v of nutrient feed containing enough concentrated glucose to raise the reactor concentration to 4 g $L^{-1}$. Feeding occurred on days 3, 6, and 9 from the start of the 14-day culture.

### *Sampling and data collection*

Perfusion and fed batch bioreactors were manually sampled every day of the culture between 8:00 AM and 12:00 PM. Prior to taking the sample, 3 mL of cell broth was removed to clear the dead volume in the sampling tube. A 5 mL sample was removed for gas, electrolyte, metabolite, nutrient, osmolality, and cell density analysis using the Nova Flex (Nova Biomedical Corporation, Waltham, MA) from the unfiltered bioreactor contents. In perfusion cultures, 30 mL of clarified permeate was taken from the shell side of the ATF and frozen at −80°C to be analyzed for glycan profile at a later time. In fed batch cultures, 30 mL whole cell broth was taken, centrifuged and supernatant removed and frozen for glycan analysis.

In addition, the MAST™ (Modular Aseptic Sampling Technology) system (Capsugel, Bend, OR) was used to take automated samples from the perfusion system when manual samples were inconvenient. This automated system uses pneumatics to draw samples from the reactor and send them to analytical instruments. Nightly, one sample was delivered directly to the Nova Flex and one sample to the Gilson GX271 liquid handler (Gilson, Middleton, WI). In addition, 30 mL of clarified cell broth was sampled and delivered to a chilled, 6°C rack on the liquid handler. The antibody was subsequently purified using an automated protein A purification method developed here for the Gilson liquid handler (HiTrap rProtein A FF, 1 mL column, GE Healthcare, Wauwatosa, WI). At the end of each automated sample, the full panel of Nova Flex (Nova Biomedical, Waltham, MA) data is collected and purified mAb sits on the chilled rack for storage and future glycan analysis.

### *Effect of galactose concentration on GLUT1 and GLUT4 protein levels*

In a separate experiment, the CHOZN cells (Sigma–Aldrich; St. Louis, MO) were grown in the wells of deep 12-well plates. Like the perfusion reactor the cells were grown in a 90/10 blend of the Platform Media/Feed. The cells were treated with either 1, 5, or 10 mM galactose. Following 4 days in culture the cells were harvested, centrifuged, and the pellet flash frozen at −80°C until Western blot analysis.

### *Perfusion step-change experiments*

A system identification approach based on step changes in galactose feed concentration was employed to generate the necessary dynamic data for model construction. To accomplish this, the perfusion reactor was subjected (once steady-state was reached) to increasing or decreasing concentrations of galactose selected from (0, 0.1, 1, 2, 5, and 10 mM). The order of the steps was randomized, with the exception of the first perfusion experiment, where all steps were executed in order of increasing concentration. The data presented here was gathered from a total of three perfusion experiments, where each perfusion experiment consisted of a series of multiple step changes, and the order of the steps between each perfusion experiment was different. Before switching of galactose concentration, steady-state was confirmed by ensuring there was no greater deviation in the %galactosylation than ± 2%galactosylation for the previous six measurements. The sampling procedure included twice daily samples on days that there was not a step-change, and four daily samples for 2 days following a step change. This sampling frequency was determined through experimentation to define the data resolution necessary to capture the steady state and dynamic characteristics of the system. The step changes were performed by introducing a bolus of galactose to immediately raise the galactose concentration to target, and then changing the feed composition to the appropriate concentration of galactose.

### *Analytical methodology*

*Determination of Reactor Galactose Concentration.* To measure the galactose concentration in the perfusion reactor, the Galactose Assay Kit and protocol (Item # MAK012; Sigma–Aldrich, St. Louis MO) was followed. The assay is based on measuring the enzymatic oxidation of galactose to produce a product with an absorbance of 570 nm.

N-*Glycan Analysis.* A 500 μg of purified antibody was digested at 37°C with 2 μL of PNGaseF (NE Biolabs) overnight. Following digestion, liberated N-glycans were labeled with 2-aminobenzamide for 2–3 h at 65°C. Labeled glycans were then isolated using GlykoPrep G cartridges (Prozyme). Samples were then analyzed by LC/MS and HPLC-FLD analysis, according to the previously published method.[18]

Analysis of the glycoforms produced by the CHOZN® GS−/− cell line revealed that the predominant mAb bound glycans were the fucosylated non-galactosylated (G0F), fucosylated mono-galactosylated (G1F), and fucosylated di-galactosylated N-glycans (G2F) (Figure 2). Therefore, for determining the effect galactose concentration had on the galactosylation of mAbs, we focused on these three glycoforms. Specifically, we calculated the %galactosylation by taking the sum of G1F and G2F intensity and dividing by the sum of the G0F,G1F, and G2F intensity.

*Western Blot Analysis of GLUT1 and GLUT4 Protein Levels.* Cell lysate was prepared by homogenization with 20 passes through a glass dounce homogenizer in a Phosphate Buffer containing a protease inhibitor cocktail (Thermo Fisher; Waltham, MA). A 10 μg of protein were loaded per sample and separated using a 10% polyacrylamide gel (Bio-Rad). Following subsequent Western transfer to a PVDF membrane, blots were blocked and incubated with primary antibody against GLUT1 (1:1K, LSBio; Seattle, WA), GLUT4 (1:1K, LSBio; Seattle, WA), and GAPDH (1:10K, LSBio; Seattle, WA). Blots were subsequently incubated with corresponding secondary antibody conjugated to Cy5 dye (Jackson Immuno; West Grove, PA). Blots were imaged on a GE Typhoon scanner and analyzed using ImageJ.

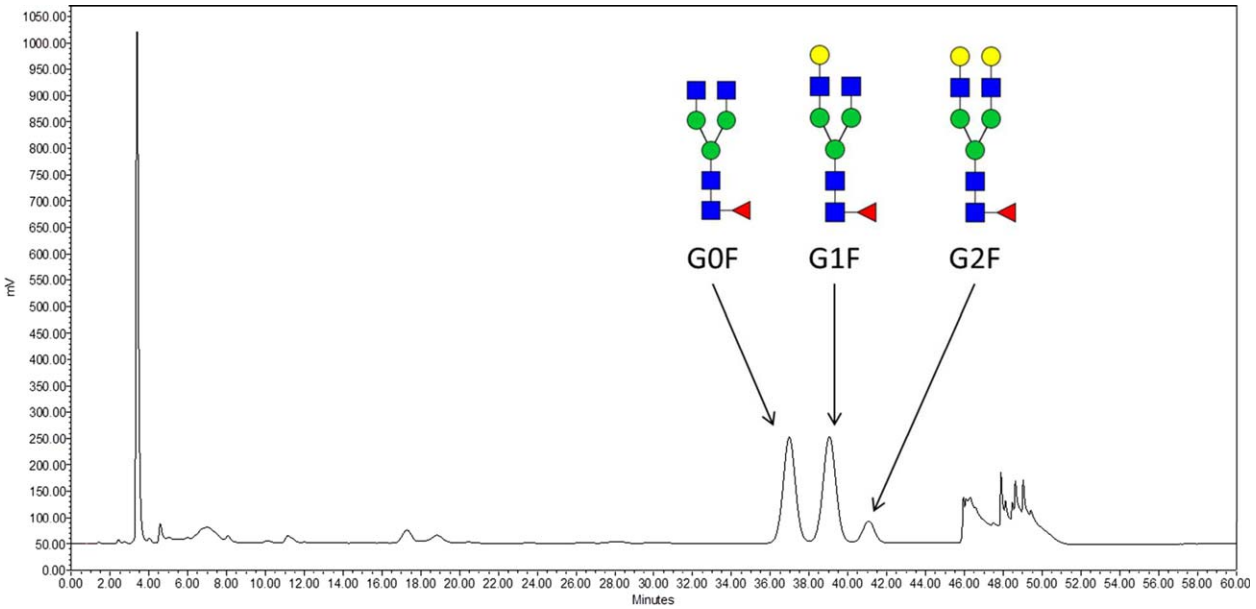


**Figure 2. Example chromatogram showing abundance of fucosylated glycans G0F, G1F, and G2F isolated from CHO cell culture system used in this study.**

Protein levels were normalized to the corresponding GAPDH intensity.

### *Data analysis*

*Calculation of %Galactosylation.* The cell line used in this study predominantly produced fucosylated *N*-glycans of the groups G0F (0 galactose), G1F (1 galactose), and G2F (2 galactose). The degree of galactosylation was calculated by taking the sum of the intensity of G1F and G2F, and dividing it by the sum of G0F, G1F, and G2F. Therefore, we are expressing the degree of galactosylation as the percentage of galactosylated *N*-glycans to all *N*-glycans. This attribute is referred to here as %galactosylation.

*Calculation of the Cell-specific Consumption Rate of Galactose.* During the steady-state phases of the experiment, the reactor was fed a constant concentration of galactose. Because of the cellular consumption of galactose, there is a difference between the concentration in the reactor and the concentration in the feed at steady-state in the presence of cells. This concentration difference was used to calculate the cell specific consumption of galactose as follows (Eq. 1):

$$\mathrm{CSCR_{gal,ss}} = \frac{F\mathrm{in} \times (C\mathrm{gal,feed,ss} - C\mathrm{gal,reactor,ss})}{\mathrm{VCD} \times \mathrm{V}}$$

$\mathrm{CSCR}_{gal,ss}$ = cell specific galactose consumption rate ($\mu$mol/cell/day)

$C_{\mathrm{gal,feed,ss}}$ = steady−state feed galactose concentration ($\mu$M)

$C_{\mathrm{gal,reactor,ss}}$ = Steady−state reactor galactose concentration ($\mu$M)

VCD = viable cell density $\left(\mathrm{cells\,mL^{-1}}\right)$

$V$ = reactor volume (mL)

$F_{\mathrm{in}}$ = feed flow rate $\left(\mathrm{mL\,day^{-1}}\right)$

(1)

*Determination of Model constants by Curve-Fitting.* As experimental responses to step changes in galactose feed concentration resembled that of a second order dynamic model, the dynamic step response data was fit to model outputs of a second order dynamic model (see Eq. 4). Specifically, $K_{\mathrm{cells}}$ and $\tau_{\mathrm{cells}}$ constants were optimized to achieve a minimum sum of least squares fit to experimental step response data. Optimization was performed using the fminsearch algorithm in MATLAB (version 2015a, Mathworks, Natick, MA) which is based on the Nelder-Mead simplex algorithm.[19]

*Calculation of Galactose Concentration vs. Time in Fed-batch System.* Because galactose is not a typically measured metabolite during mammalian cell culture processes, the galactose concentration profile was predicted in the fed-batch model as shown in Eq. 2. The equation uses knowledge of the uptake kinetics of galactose generated experimentally, combined with measurements of the viable cell density, reactor volume, and feed recipe to determine the galactose concentration in the reactor over time.

$$Gal_{\mathrm{t}} = \frac{Gal_{\mathrm{t-1}} \times V_{\mathrm{t}} - \frac{V_{\max}\mathrm{Gal_{t-1}}}{k_{\mathrm{m}} + \mathrm{Gal_{t-1}}} \times \mathrm{VCD_{t-1}} \times V_{\mathrm{t}} \times \Delta t + \frac{V_{\mathrm{bol,t}} C_{\mathrm{bol,t}}}{V_{\mathrm{t}}}}{V_{\mathrm{t}}} \quad (2)$$

*Simulation of Model Predictive Controller.* A model predictive controller (MPC) was constructed using a receding horizon approach with bias correction (further described in the Results). All simulations were performed in MATLAB (version 2015a, Mathworks, Natick, MA). In each MPC simulation, the model was either assumed to predict the actual output perfectly, or disturbances and measurement noise were added to the actual output to simulate model error. In the case where the model predicts the output perfectly, the measured values were assumed to be equal to the model predicted value at each measurement time.

Measurement error was added by adding Gaussian noise with a standard deviation equal to the average standard deviation of measured %galactosylation measurements at steady-state in the perfusion experiments ($\sigma/\mu = 0.009$) to the model forecasted values. Disturbances were added by adding a functional offset to the measured values.

Disturbance functions tested in this study were ±1% constant offset and $\pm 0.1 \times t$ offset (a linearly increasing offset). Both the disturbances and measurement errors were

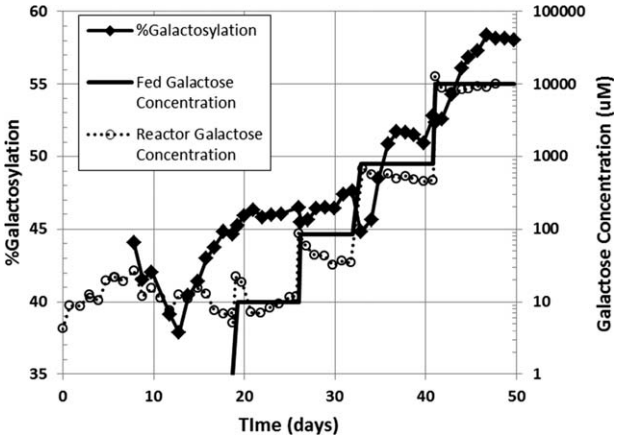


**Figure 3. Example step-response experiment.**

Reactor galactose concentration is manipulated in a step-wise fashion and the resulting change in %galactosylation and reactor galactose concentration is recorded. Increases in %galactosylation are observed at galactose feed concentrations of 100 $\mu$M and above. No galactose was fed prior to day 19.

calculated by iteratively adding each to a zero bias model predicted %galactosylation for each sampling instant, and then implementing a feed pump rate based on those values with disturbances and error added.

Conditions used in the simulations were:

- Initial reactor volume = 1 L
- Bolus nutrient feeds (not containing galactose) on days 3.2, 6.2, and 9.2 of 100 mL each
- Galactose feed concentration = 100 g $L^{-1}$
- No galactose in the basal media
- %Galactosylation set point = 50%
- Sampling interval = 24 h
- Simulation time step = 0.1 days
- Measurement lag = 24 or 48 h

The optimizer used in the MPC was fmincon, a constrained nonlinear optimization function in MATLAB.[20]

## Results

### *Pseudo-steady-state experiments*

To create a predictive model of %galactosylation, training data was generated to determine both the steady-state and dynamic response of %galactosylation to specified inputs. In this study, a pseudo-steady-state perfusion system was used to subject the cell culture to step changes in galactose concentration. The resulting data were then used to obtain the effect of each step change on galactose consumption and %galactosylation, ultimately supporting the development of a predictive model.

Specifically, a pseudo-steady-state cell culture was subjected to step increases in reactor galactose concentration. The resulting response of %galactosylation of mAb *N*-glycans was measured. Each step-response was allowed to reach steady-state with respect to %galactosylation, defined as six consecutive measurements within 2% galactosylation of each other. Switching of the step (galactose concentration) was not initiated until the above criteria were met (Figure 3).

Each step change caused the reactor galactose concentration to immediately increase (due to a bolus addition of galactose, as described in Methods) and then stabilize at a steady-state. The difference between the galactose concentration of the feed media and the reactor was used to calculate the cellular consumption of galactose ($CSCR_{gal,ss}$), as outlined in the methods. In addition, each step change of galactose concentration of >0.1 mM produced a corresponding increase in %galactosylation of secreted antibody. The %galactosylation increased dynamically in response each bolus addition of galactose, with all responses reaching a new steady-state value within 12–48 h.

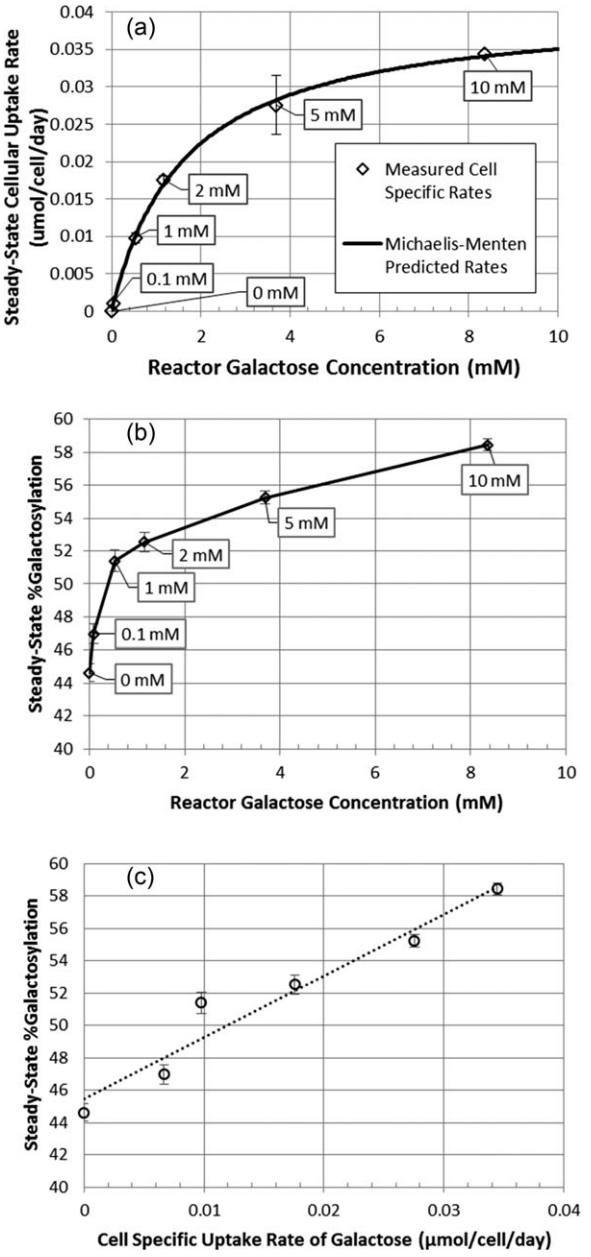


**Figure 4. (a) Steady-state cell specific galactose consumption rate as a function of reactor galactose concentration. (b) Steady-state %galactosylation as a function of reactor galactose concentration. (c) Linear correlation of steady-state %galactosylation as a function of galactose uptake rate. Steady-state cell specific galactose consumption rates were fit to a Michaelis–Menten kinetic model. The steady-state %galactosylation was linearly correlated with the cell specific galactose consumption rates for corresponding galactose concentrations. The data labels in Figures 4a,b are the galactose concentrations in the reactor feed, whereas the x axis is the measured galactose concentration in the reactor.**

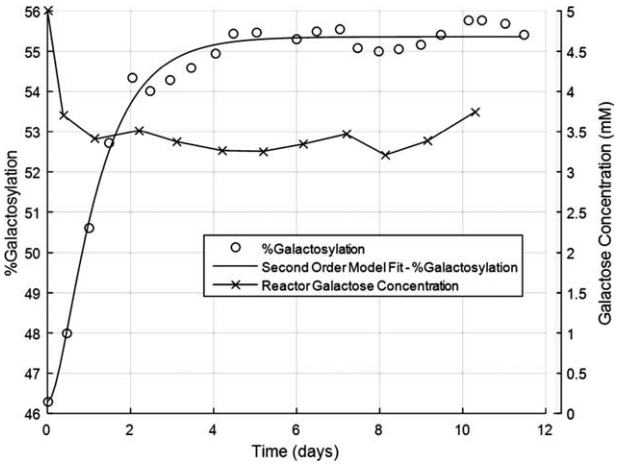


**Figure 5. Example dynamic step response of %galactosylation in response to step increase in galactose concentration.**

A second order step-response model (solid line) was fit to the observed %galactosylation data (o). The galactose concentration (x) was raised to 5 mM initially using a bolus addition at $t = 0$ days, and then the reactor was continuously fed media containing galactose at 5 mM. The actual reactor concentration of galactose is lower than the feed concentration due to cellular consumption.

### *Steady-state galactosylation characterization*

The values obtained from calculating the steady-state galactose consumption ($CSCR_{gal,ss}$) and %galactosylation were analyzed against the respective reactor concentration of galactose, to determine how the steady-state consumption rates and steady-state %galactosylation were impacted by the concentration of galactose (Figures 4a,b).

As evidenced by the asymptote in the $CSCR_{gal,ss}$ vs. galactose concentration curve, higher concentrations of galactose feed (5 and 10 mM Galactose) exhibited saturating behavior (Figure 4a). This kinetic behavior suggests an enzymatic regulation of galactose consumption and was therefore analyzed using Michaelis–Menten kinetics. The cellular consumption rate data was fit using least squares regression to a Michaelis–Menten model, yielding a Vmax of $4.07 \times 10^{-2}$ $\mu$mol cell$^{-1}$ day$^{-1}$ and a $K_m$ of 1.63 mM.

The steady-state %galactosylation as a function of galactose concentration exhibited similar behavior to the galactose consumption, suggesting that %galactosylation may be directly correlated to the galactose consumption rate. When analyzed, the correlation of steady-state %galactosylation vs galactose uptake results in a direct, linear correlation (slope = 380. %*cell*day/$\mu$mol, intercept = 45.5%, $R^2 = 0.944$) (Figure 4c).

### *Model development—dynamic response of %galactosylation to galactose concentration*

To use the data obtained from the galactose step change experiments to train a predictive model of %galactosylation for an arbitrary galactose concentration profile, the dynamics of the %galactosylation as a function of changes in galactose concentration must be determined independently of any convoluting effects which are unique to the apparatus.

By analyzing the steady-state $CSCR_{gal,ss}$ and %galactosylation the steady-state differences in response between the various step-changes are defined. However, the period prior to reaching each steady state reveals information about the kinetic response of %galactosylation to the changing concentration of galactose. As an example the kinetic response of %galactosylation to step changes in galactose concentration from 0 to 5 mM is graphed in (Figure 5). The increase in %galactosylation does not occur immediately, in contrast to the behavior of the galactose reactor concentration. In fact it takes days (2) to reach the new steady-state %galactosylation values after initiation of a step-change. These kinetics appear to follow an over-damped second order response (Figure 5).

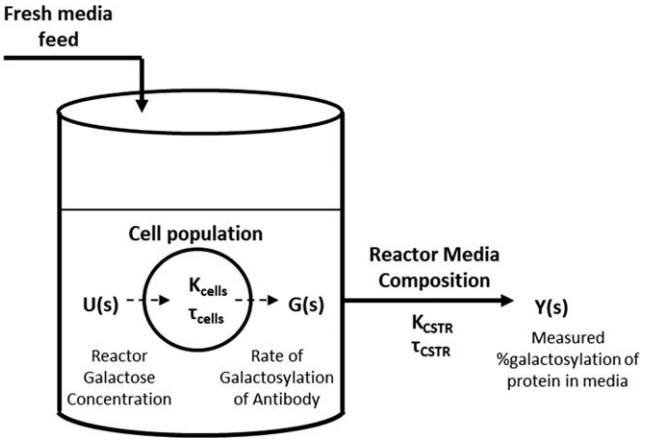


**Figure 6. Schematic representation of physical situation present in pseudo-steady-state apparatus.**

The measured %galactosylation ($Y$(s)) is a function of the cell specific production rate of galactosylated antibody ($G$(s)) and the fresh media dilution rate. $K_{CSTR}$ and $\tau_{CSTR}$ are functions of the perfusion apparatus dilution rate, and are known, whereas $K_{cells}$ and $\tau_{cells}$ are unknowns.

In the pseudo-steady-state apparatus used in this study, the observed dynamics are a convolution of both (1) the change in %galactosylation of the protein produced by cells (which is what we are trying to determine), and (2) the dilution/residence time effects of the expressed protein in the continuously stirred tank reactor (CSTR) apparatus, as shown schematically in Figure 6.

Where $U$(s) is the input galactose concentration, $K_{cells}$ and $\tau_{cells}$ are the gain and time constant of the change in %galactosylation of expressed antibody from the cell population, $K_{CSTR}$ and $\tau_{CSTR}$ are the gain and time constant of the change in %galactosylation of measured antibody in the system due to CSTR dynamics, $G$(s) is the production rate of galactosylation antibody over time, the $Y$(s) is the measured %galactosylation in the reactor over time.

Specifically, a cell population in this system that is subjected to a change in galactose concentration ($U$(s)) will, in response, change the %galactosylation of the protein it is expressing ($G$(s)). However, in order to measure this change in the CSTR apparatus, we must measure the %galactosylation present in the bulk media ($Y$(s)). At the time of a change in %galactosylation of the protein produced by the cell population ($G$(s)), the reactor still contains protein with a %galactosylation resulting from the previous galactose concentration. Therefore, a measurement of the %galactosylation of protein in the bulk media ($Y$(s)) will not reach a new steady-state until the protein from the previous steady-state has been diluted out of the system by the continuous addition of fresh media and replaced by the production of the protein with the new galactosylation (G(s)) level. This results in a series convolution of the galactosylation dynamics of the cells and the dilution dynamics of the CSTR

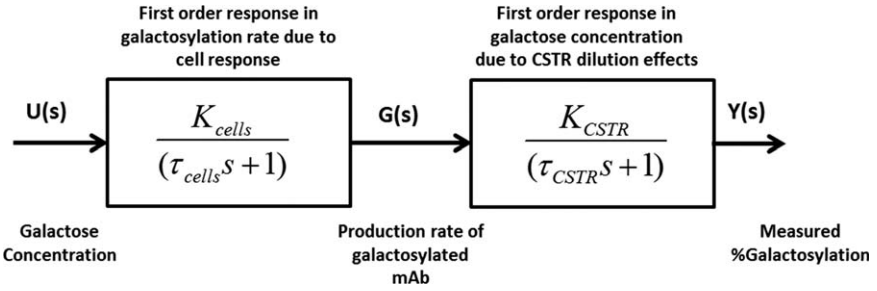


**Figure 7. Theoretical block diagram depicting series convolution of constant CSTR kinetics and %galactosylation rate cellular kinetics.**

The cell specific production rate of galactosylation mAb ($G$(s)) as a function of galactose concentration ($U$(s)) is the relationship to be determined in the analysis.

apparatus. Therefore, in order to calculate the dynamic response rates of %galactosylation to concentration changes, the dilution kinetics of the apparatus (CSTR) and the kinetics of the change in %galactosylation need to be considered together.

This series relationship is formalized in a block diagram (Figure 7).

Because any observed dynamic output in a CSTR will be due to a series combination of the CSTR kinetics and the actual change in %galactosylation driven by the cell response, the response of each step change was assumed to be second order. Note that although the CSTR dynamics are known to be first order, the cell response dynamics (resulting in output $G$(s)) were assumed to be first order. The dynamics of either block can be of any order, but this assumption simplifies the data analysis and fits the observed responses for this particular experiment. This analysis gives the following equation for a dynamic step-response of %galactosylation for a given step increase in galactose concentration in the Laplace domain Eq. 3:

$$Y(s)=\frac{\mathrm{KM}}{(\tau_{\mathrm{cells}}s+1)(\tau_{\mathrm{CSTR}}s+1)} \tag{3}$$

Representing Eq. 3 in the time domain for the over-damped case results in Eq. 4

$$y(t)=KM\left(1-\frac{\tau_{\mathrm{cells}}e^{-t/\tau_{\mathrm{cells}}}-\tau_{\mathrm{CSTR}}e^{-t/\tau_{\mathrm{CSTR}}}}{\tau_{\mathrm{cells}}-\tau_{\mathrm{CSTR}}}\right)+y_0 \tag{4}$$

Where y(t) is the observed %galactosylation measured in the reactor at any time, $K$ is the system gain (%galactosylation/mM galactose), $M$ is the step magnitude (mM), $y_0$ is the starting %galactosylation prior to initiation of the step change. Note $K$ is the product of $K_{\mathrm{cells}}$ and $K_{\mathrm{CSTR}}$.

Because the feed rate and reactor volume were held constant for all experiments (steady-state assumption), we can assume that the time constant associated with CSTR dynamics ($\tau_{\mathrm{CSTR}}$) was also constant for the data collected. In this case, the CSTR time constant $\tau_{\mathrm{CSTR}}$ is simply the mean residence time of the reactor Eq. 5.

$$\tau_{\mathrm{CSTR}}=\mathrm{Residence\ Time}=\frac{V}{F_{\mathrm{in}}} \tag{5}$$

The resulting fit constants for the cell dynamic portion of the response can then be obtained by fitting $K$ and $\tau_{\mathrm{cells}}$ to the %galactosylation response data using least squares regression. The cell response time constant ($\tau_{\mathrm{cells}}$) as a function of galactose concentration can then be determined (Figure 8).

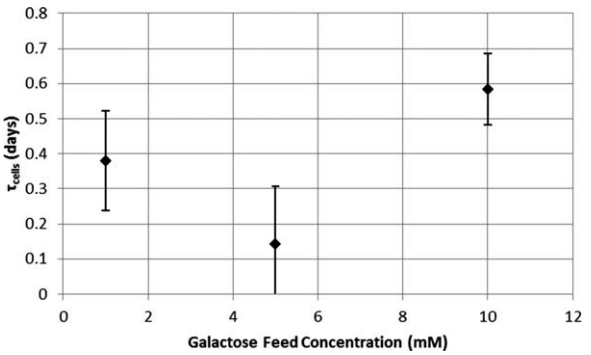


**Figure 8. Cellular %galactosylation response time ($\tau_{\mathrm{cells}}$) vs. galactose feed concentration.**

The %galactosylation response time was determined by fitting a second order model (Eq. 4) to %galactosylation data obtained from step response experiments.

The response time for cellular %galactosylation shows a decrease in response time between galactose concentration of 1 and 5 mM, followed by an increase in response time between 5 and 10 mM. The fact that the response time is not constant as a function of galactose feed concentration suggests that the underlying dynamics of the system are changing as a function of the galactose feed concentration in a manner which is not entirely captured by the simple linear, first order approximation used here. This nonconstant response time was captured in the model by allowing $\tau_{\mathrm{cells}}$ to vary as a function of galactose concentration.

We hypothesized that the biological basis of this change in the process time constant was due to a change in GLUT1 and GLUT4 protein levels. GLUT1 and GLUT4 are the major transporters responsible for the uptake of galactose into cells. To test this hypothesis, cultures were grown for 4 days in either 1, 5, or 10 mM concentrations of galactose and the protein levels of GLUT1 and GLUT4 measured by Western blot. When the normalized expression of GLUT1 and 4 are represented as a function of the corresponding response times ($\tau_{\mathrm{cells}}$), an inverse linear relationship is observed (Figure 9). As the protein expression of GLUT1 and 4 increases the process time constant decreases. This implies that the shorter process time constant is partially driven by the increased protein expression of GLUT1 and 4, and thus a potential for more rapid galactose uptake. Further, since longer response times are observed at conditions with lower galactose transporter expression, we conclude that changes in expression of hexose transporters are a major determiner of changes in characteristic response time ($\tau_{\mathrm{cells}}$).

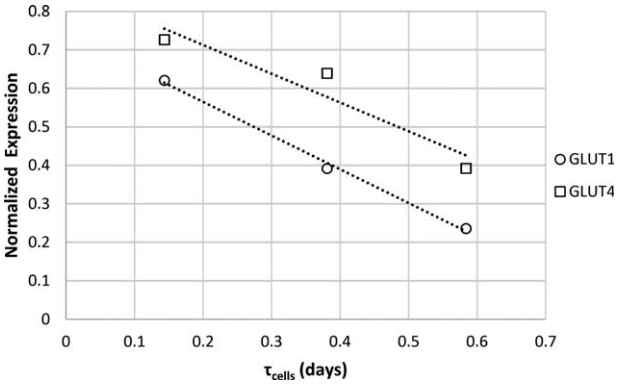


**Figure 9. Cellular galactosylation rate time constant ($\tau_{cells}$) vs. GLUT1 and GLUT4 transporter expression at varying galactose feed concentrations.**

Normalized expression of GLUT1 and GLUT4 values were calculated by measuring intensity of the respective GLUT signal divided by the intensity of GAPDH signal. These normalized expression values were graphed as a function the response times calculated for the same galactose concentrations derived from pseudo-steady-state experimentation. Linear regression revealed an inverse linear correlation between GLUT expression levels and response time.

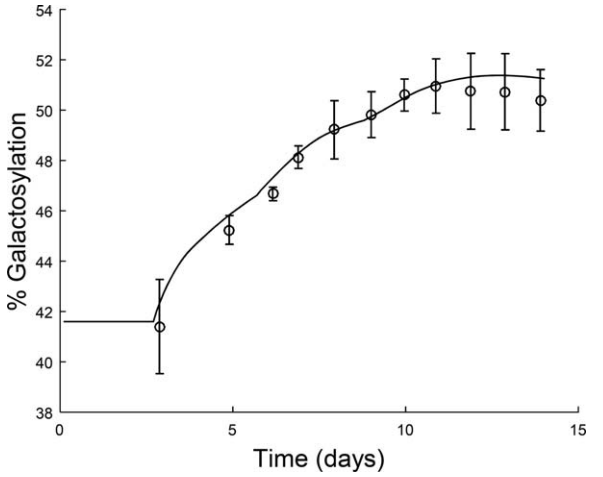


**Figure 10. Predicted (-) and actual (o) %Galactosylation vs. time for a fed-batch culture with galactose feeds.**

Galactose was fed in combination with nutrient feeds in boluses on days 3, 6, and 9. Actual %galactosylation is presented as a mean of three separate fed-batch runs. Error bars represent 1 standard deviation.

### *Creation of predictive model of %galactosylation for translation to industry relevant fed-batch platform*

Using the step-response data to determine both the steady-state gain ($K_{\text{cells}}$) and dynamic response time ($\tau_{\text{cells}}$), we then produced a model that could predict the effect of a fed-batch strategy on the resulting degree of mAb galactosylation. The resulting model was used to predict %galactosylation vs. time for a fed-batch culture system.

Because galactose is not a traditionally measured metabolite in mammalian cell culture processes, any controller which uses a model relying on this measurement would be difficult to implement. Therefore, the galactose concentration profile was calculated using knowledge of the galactose uptake kinetics and galactose feed profiles (Eq. 2 in the Methods).

To predict the %galactosylation at a given timepoint for an arbitrary cell culture system, the response of the cells must be determined for an arbitrary galactose input. Because the kinetics of %galactosylation due to the cells were assumed to be first order in this analysis (the first block in Figure 7), the following model applies Eq. 6:

$$\tau_{\text{cells}}\frac{dG(t)}{dt}+G(t)=K_{\text{cells}}u(t) \tag{6}$$

where $G$(t) is the galactosylation of antibody expressed at time (t), the quantity $K_{\text{cells}}$u(t) is the steady-state %galactosylation at a given galactose concentration ($u$(t)), determined by the steady-state analysis above, $\tau_{\text{cells}}$ is the time-constant as determined in the dynamic analysis above, and dG(t)/dt is the rate of change of the %galactosylation of expressed antibody at time (t). Note for small values of $\tau_{\text{cells}}$, the system approximates steady-state.

Solving for $G$(t) with the initial condition $G(\text{t}=0)=G_0$ yields the following equation for the %galactosylation over time:

$$G(t)=[G_0-K_{\text{cells}}u(t)]e^{-t/\tau_{\text{cells}}}+K_{\text{cells}}u(t) \tag{7}$$

We determined previously that the steady-state %galactosylation varies linearly as a function of the cell specific galactose uptake rate (Figure 4c). This relationship was used to calculate the gain-dependent portion of (Eq. 7).

$$K_c(u(t))=K_{cells}u(t)=a\left(\frac{V_{\max}u(t)}{k_m+u(t)}\right)+b \tag{8}$$

where $K_{\text{c}}$(u(t)) is the steady-state %galactosylation as a function of galactose concentration, $V_{\max}$ and $k_{\text{m}}$ are the maximum galactose uptake rate and Michaelis constant, and $a$ and $b$ are the slope and intercept of the %galactosylation vs. cell specific galactose uptake rate relationship determined in Figure 4c.

We also previously determined from the step-change data that the response time ($\tau_{\text{cells}}$) varies as a function of galactose concentration. During model calculation, $\tau_{\text{cells}}$ was interpolated from the $\tau_{\text{cells}}$ vs. galactose concentration data determined previously.

Substituting Eqs. 8 into Eqs. 7 and allowing $\tau_{\text{cells}}$ to vary as a function of galactose concentration yields the final form Eq. 9.

$$G(t)=[G_0-K_{\text{c}}(u(t))]e^{-t/\tau_{\text{cells}}(u(t))}+K_{\text{c}}(u(t)) \tag{9}$$

To use the pseudo-steady-state apparatus in process development and model predictive control, one must be able to generate predictions for industry relevant cell culture platforms that may be different than those used to generate the model (i.e., Perfusion vs. Fed-Batch Reactor). To test the prediction quality of this approach, we applied predictions to a fed-batch culture that was fed boluses of supplementary galactose, as described in Methods.

In a fed-batch culture, unlike the perfusion system, the viable cell density (VCD) of the reactor changes over time, and thereby changing the rate of galactose consumption. In addition, all secreted protein remains in the bioreactor. The %galactosylation in a fed-batch reactor over time is given by Eq. 10.

$$\%Gal_{\text{cum}}(t)=\frac{[mAb_{\text{Gal}}]}{[mAb_{\text{tot}}]}=\frac{\int_0^t V(t)\times VCD(t)\times Q_{\text{p}}(t)\times G(t)dt}{\int_0^t V(t)\times VCD(t)\times Q_{\text{p}}(t)dt} \tag{10}$$

where $V$ is the vessel volume, VCD is the viable cell density, $Q_{\text{p}}$ is the cell specific production rate of antibody, and

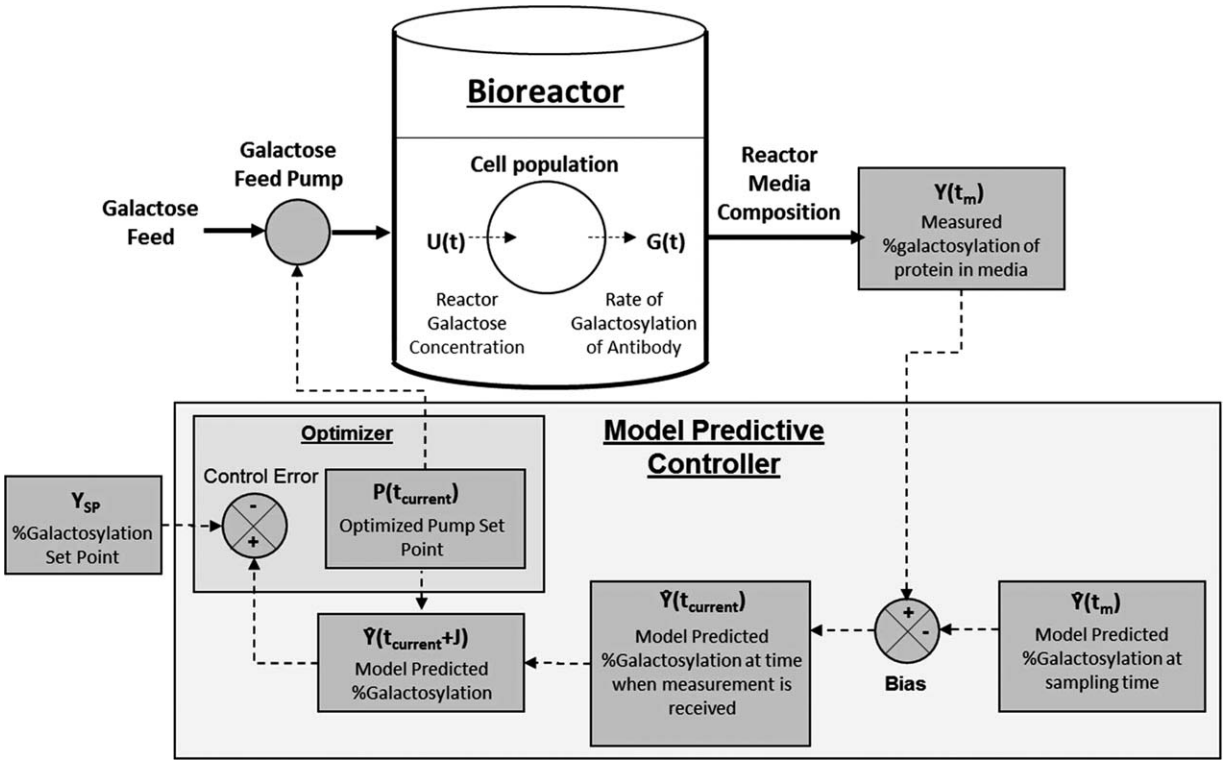


**Figure 11. Schematic representation of the control loop simulated in the model predictive controller.**

The controller used sparse measurements of %galactosylation combined with model forecasts to control a galactose feed pump. $Y$ is the model predicted %galactosylation, $t_m$ is the time at which the bioreactor was sampled, $t_{current}$ is the time at which the %galactosylation measurement for the sample is received, $J$ is a tunable parameter that dictates how aggressive the controller is, $Y_{sp}$ is the %galactosylation set point, and $P$ is the galactose pump set point.

$G$ is the %galactosylation of expressed antibody, given by Eq. 7.

To evaluate the prediction quality of the fed-batch model (Eq. 10), the model was evaluated for the conditions present in the fed-batch experiments. Inputs to the model were vessel volume, viable cell density, cell specific productivity of mAb, galactose feed concentrations and feed volumes. We observed a good prediction of %galactosylation over the course of the fed-batch reactor (Figure 10).

### *Creation of model predictive controller for %galactosylation and simulation in fed-batch platform*

To demonstrate how the fed-batch predictive model could be implemented for control of %galactosylation, we created a model predictive controller in simulation using the fed-batch predictive model outlined previously to control %galactosylation to defined set point. The physical control loop we simulated is outlined in Figure 11.

The model predictive controller was designed based on a technique outlined previously,[21] based on a receding horizon approach incorporating bias correction to account for model inaccuracies and measurement lag. A summary of the controller execution steps is outlined in Figure 12.

For each sampling instant, a measurement lag of 24 h was assumed. This is equivalent to the fastest turnaround between sampling and measurement which can be achieved for the glycan measurement used in this study. Here the sampling time is labeled $t_m$ and the time at which the %galactosylation measurement of the sample is received is labelled $t_{current}$.

Once a %galactosylation measurement is received, the difference between the measured %galactosylation ($y(t_m)$) and the model predicted %galactosylation ($\hat{y}(t_m)$) is calculated. This difference is added to all remaining model forecasts for the current model iteration, here called the bias ($B = y(t_m) - \hat{y}(t_m)$).

The controller then creates a model forecast including the bias for the timeframe $t = t_m$ to $t = t_{current}$. The controller estimates the %galactosylation at $t_{current}$ using this model forecast. The initial conditions for this forecast are equal to the conditions of the previously generated model forecast at $t = t_m$.

Once an estimated %galactosylation at the current time ($t_{current}$) is generated, the controller then optimizes the galactose feed pump set point to minimize the error between the %galactosylation set point and a model predicted value at some time ($t_{current}+J$) in the future. $J$ is a tunable parameter that for this simulation was set to 96 h. The galactose feed pump set point was constrained to 0–10 mL day$^{-1}$ to prevent excessive feeding. The optimized pump set point is then implemented in the simulation and not changed until a new measurement is received.

We implemented this model predictive control scheme in simulation to gauge its' ability to effectively control %galactosylation to a set point by manipulating a galactose feed pump rate.

In simulation, the controller was able to achieve a %galactosylation within 0.5% of set point by the end of each batch for a measurement lag of 24 h for all scenarios tested (Figure 13). A measurement lag of 48 h was also tested, with similar results (data not shown). The controller successfully kept the galactose concentration in the range explored by the steady-state-apparatus (0–10 mM).

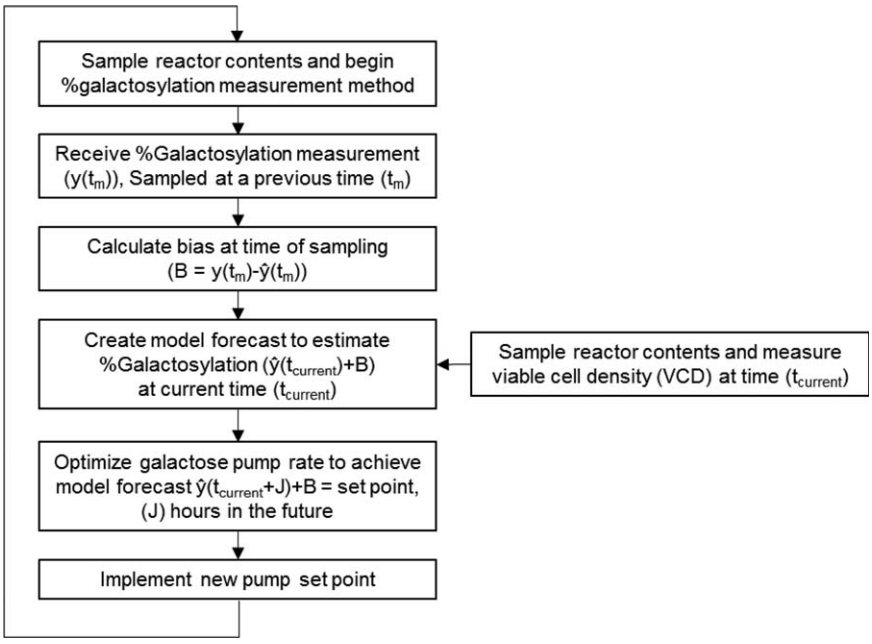


**Figure 12. Execution steps for model predictive controller developed using fed-batch %galactosylation forecast model.**

The steps are executed every time a new sample measurement of %galactosylation is received, and a new galactose pump set point is implemented.

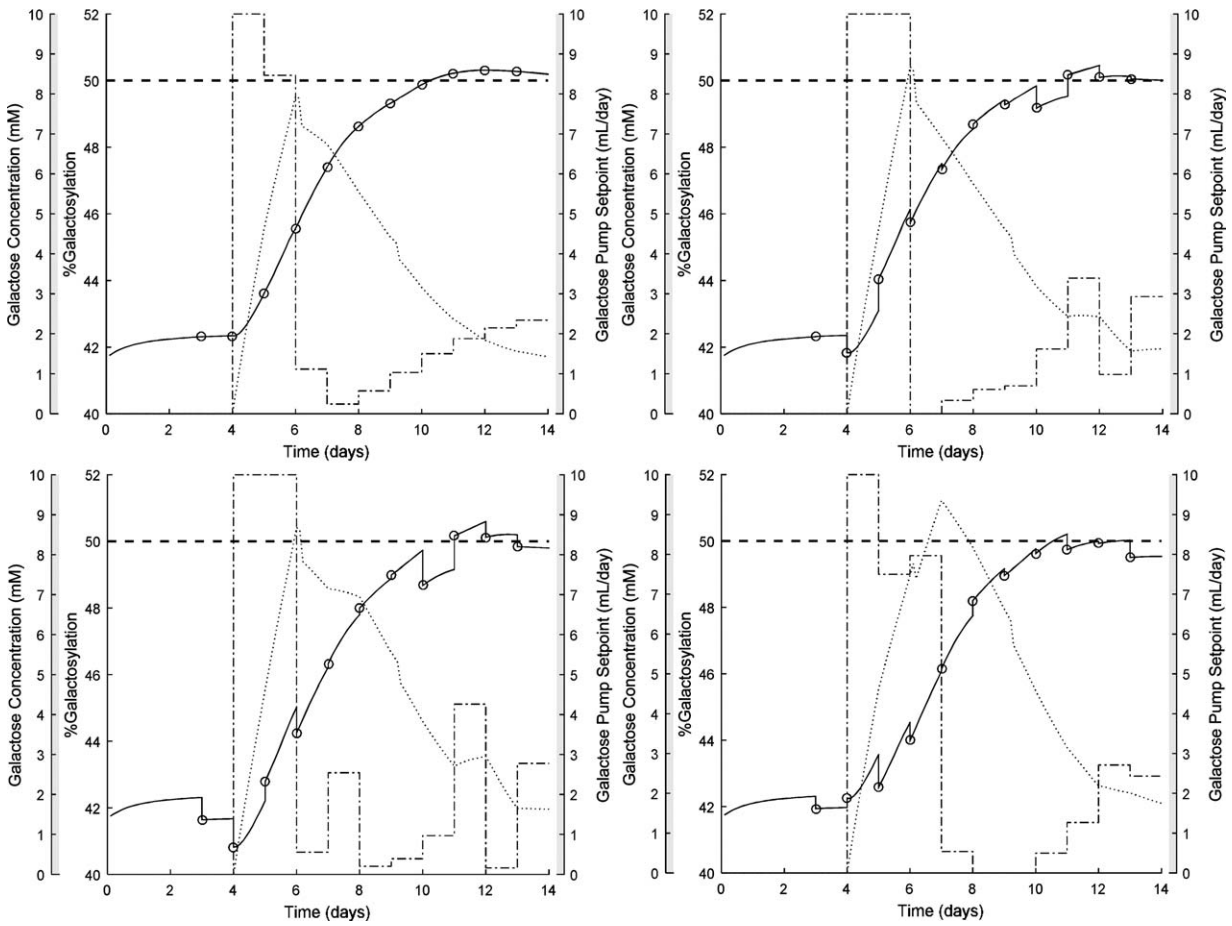


**Figure 13. Simulated MPC runs for controlling %galactosylation in a fed-batch run with: (a) no measurement noise or disturbances, (b) measurement noise added, (c) measurement noise and constant −1% offset disturbance, (d) measurement noise and linearly increasing $-0.1 \times t$ offset disturbance. Measured %galactosylation (o), model predicted %galactosylation (solid line), galactose pump rate (dash dot line), model predicted reactor galactose concentration (dotted line).**

## Discussion

### *Feeding of galactose increases the %galactosylation of a mAb*

By increasing the extracellular concentration of galactose we observed an increase in the %galactosylation of mAb. This result agrees with prior work, as treatment with galactose (as well as uridine and manganese chloride) increased the degree of galactosylation of a mAb in two different cell-lines.[22] The %galactosylation was directly correlated with galactose consumption rate which followed classic Michaelis–Menten kinetics. This observation, combined with the fact that the protein levels of GLUT1 and GLUT4 were inversely correlated with the characteristic response times, suggests that the bottleneck in galactose utilization for galactosylation is galactose transport, specifically through hexose transporters. We observed an increase in the protein levels of GLUT1 and GLUT4 when cultured in 1 and 5 mM galactose relative to no galactose. At 10 mM galactose we did not observe a change in GLUT1 and GLUT4 transporters. This may suggest that the regulatory signal for hexose transporters has a switch-point between 5 and 10 mM galactose. Indeed, this biphasic response to hexose concentration by hexose transporters has been genetically defined in yeast.[23] Furthermore, the effect of high levels of hexose on the downregulation of hexose transporters has been reported in cultures derived from rat skeletal muscle.[24] The effect of hexose transporter expression on galactose uptake was demonstrated through the stimulation of GLUT1 expression by glutamate treatment and the concomitant increase in galactose uptake, suggesting that GLUT expression is a major modulator of galactose uptake.[25] These observations reinforced choise of galactose as a control handle for %galactosylation, and supported the use of a varying response time as a function of galactose concentration in the model.

### *System identification with pseudo steady-state apparatus is a viable technique for predictive model development*

In this study, we demonstrated the ability to use a model constructed from pseudo-steady-state apparatus data to predict the %galactosylation of a mAb in an industry relevant fed-batch process. By analyzing both the steady-state galactose consumption and %galactosylation, in combination with the dynamics of the response, we developed a predictive model for levels of galactosylation for a given galactose feeding regimen. We then demonstrated the successful control in simulation of a realistic fed-batch scenario using the predictive model in a model predictive control scheme.

One implication of our study, which has limited precedent in the literature, is how well a model trained using data in a pseudo-steady-state system would be able to predict in a fed-batch context. Here we've shown that for galactosylation this translation between systems is possible. Prior work demonstrated the ability to develop a model predictive controller for high mannose species in a semi-perfused system using only data obtained through small microplate experimentation.[5] This suggests that development of a MPC may be possible through experimentation in different culture systems and scales, than that which the MPC is ultimately targeted for. More studies need to be conducted to demonstrate the wide applicability of this approach on other product quality attributes (e.g., fragments, reduced species, O-glycosylation and metabolic byproducts).

### *Pseudo steady-state apparatus has advantages over fed-batch for controller development*

One of the major goals for this study was to adopt the recent advances in bioprocess automation technologies to reduce laborious process development activities (like developing controllers for product CQAs). In this respect, the ability to generate a predictive model of cell culture process outputs in a pseudo-steady-system has some distinct advantages over other methods.

Most importantly, the step-response system identification method used in this study is systematic, continuous, and in principle generalizable to many product attributes. This means that regardless of the product quality attribute being studied, an experimental design can be generated algorithmically and executed continuously, allowing the process to be highly automated.

Even though a fed-batch or other laboratory DOE study design can also be automated,[12] the actual execution of the study design remains difficult to automate. Although technologies are being developed that may lower this technical barrier for automation of fed-batch experiments (for example, the ambr® system), these solutions remain very capital intensive, and usually require additional automation solutions to test the number of variables required to generate a suitable predictive model. Application of expert knowledge has allowed some groups[5] to significantly decrease the number of experiments required to develop a predictive model for better understood quality attributes, but this only applies to CQAs where fundamental understanding already exists. Finally, in a batch DOE the setup and takedown of experiments would need to be automated in addition to the actual experiment itself in order to realize the full benefits of automation. In contrast, for a continuous apparatus, execution of many experiments can be accomplished by currently existing technologies, such as manipulation of feeds by incorporation of pumps with defined media, maintaining cell density with an automated cell bleed (as accomplished here with dielectric spectroscopy), and automated sampling technology combined with automated analytics to record the output.

In addition to the ability of the pseudo-steady-state apparatus to be highly automatable, there is also a potential for time-savings when compared to fed-batch experimentation. A typical fed-batch study relies on run-to-run optimization, meaning each fed-batch run is essentially one data point in a DOE-style study design. Because each fed-batch run takes 12–14 days for cell culture processes (plus another typically 2 to 3 days for reactor setup and takedown), the potential exists for shorter experiments to be performed in a pseudo-state-apparatus. In the present study, the average step-change experiment (equivalent to a full fed-batch condition for model development purposes), was 4–8 days, which is significantly shorter than a typical fed-batch study. In addition, the current study design allowed each step change to proceed well into its' new steady-state, a very conservative approach. Many system identification studies using the step-change methodology have relied on datasets that are not required to reach steady-state and were successful at quantifying the system dynamics.[26] In these studies, typically the output was allowed to reach greater than 63% of the projected final steady-state value (also known as the characteristic response time). Application of these more aggressive system identification methodologies to this experimental design could yield even more time-savings; for the current study, this would

have shortened the step change experiments to 2 days each. Future work will demonstrate these more aggressive approaches in practice for our system. A final time-savings with the pseudo-state-apparatus comes from the ability to serialize step-change experiments. After the initial startup phase, there is no turn-around time between step change experiments in the pseudo-steady-state apparatus. In the current study, we demonstrated stability of the perfusion system to >120 days, meaning many experiments could be serialized in two virtually continuously operating vessels (data not shown).

### *Pseudo-steady-state system allows observability of phenomena which are difficult to detect in fed-batch experiments*

We showed that a pseudo-steady-state system allows individual variables to be separated and manipulated independently of each other (in this study just changing the concentration of galactose at a constant cell-density). Here, we investigated the effect of galactose on protein galactosylation without the confounding effects of consumption or production of other metabolites in a non-steady-state culture that may have impacted galactosylation. Although single media components can be manipulated independently in this system, a resulting change in a CQA is not necessarily solely due to a change in that component. Rather, the change in the CQA could be a combination of other confounding variables stemming from a change in cellular metabolism which occurs as a result of the manipulated media component such that the steady-state concentrations of other nutrients and byproducts are also altered. For the purposes of developing a model predictive controller, however, we showed that knowledge of the simple single input single output (SISO) relationship between galactose and %galactosylation was sufficient to predict and control %galactosylation. The pseudo-steady-state system allows easy identification of this SISO relationship, which would be otherwise difficult to obtain in a batch system. The ability to manipulate variables (especially nutrient concentrations) independently represents an attribute of this system that is crucial for adoption in a more complex nutrient-to-output mechanism.

In addition, the pseudo-steady-state system allows dynamic manipulation of input variables, and allows the kinetic response of the output of interest to be recorded and characterized in terms of the manipulated input. The ability to selectively manipulate specific input variables (i.e., [galactose]) using a well-defined input function allows the observed dynamic outputs of the system in response to the input function to be easily characterized using established techniques. This ability allows one to extract not only steady-state information from the system, but also characterize the time-dependence of how the output changes in relation to the input.

Another benefit of the pseudo-steady-state system is that due to the continuous nature of the experiments, effects of hysteresis can be easily detected by simply modulating the order of dynamic experiments and recording the effect on the output. Any effect of the past experiments or culture conditions can be detected in this manner. Cell culture processes typically exhibit non-linear behavior which is known to result in hysteresis-type behavior. Knowledge of hysteresis is important to maintain a culture in a favorable state in order to actively control the output (i.e., Product Quality Attributes).

In fact, in the current study, hysteresis effects were detected for cultures exposed to >10 mM galactose. Specifically, steady-state galactosylation values at galactose concentrations of less than 10mM were increased after exposure to >10 mM galactose feed. To avoid this complication, step-responses occurring after the system was exposed to galactose concentrations >10 mM were omitted. The galactosylation controller was also designed to avoid concentrations above 10 mM.

### *Prediction obtained using this method is suitable for use in model predictive controller*

In this study we were able to create a predictive model based on data from pseudo-steady-state experiments and successfully predict the galactosylation behavior of a fed-batch model system. We were also able to successfully control galactosylation in simulation when the model was implemented as part of a model predictive control scheme. The model we constructed was designed intentionally to use only inputs that could be easily automated using existing autosampling technologies. To generate a prediction of the %galactosylation profile vs. time for our system, one must only supply the galactose feeding regimen, reactor volume, viable cell density, and cell specific mAb productivity. This allowed the use of the model in a model predictive controller to control galactosylation using only inputs which are readily measured in a typical cell culture process and %galactosylation measurements.

A model predictive controller was built incorporating this forecast model in a receding horizon method. In this approach, long lead time measurements (such as many product quality measurements, including %galactosylation) are only measured sporadically during a cell culture process. When a new output measurement is received, the controller creates an input profile in such a way that a model forecast of the desired output trajectory is achieved. The controller implements control based on this model forecast until a new output measurement is received, and an updated model forecast is generated.

Here, we simulated control of %galactosylation in a fed-batch cell culture by controlling a galactose feed pump. We simulated long lead-measurements (24–48 h lead time) measured daily and found that control could be achieved despite reasonable simulated measurement error and disturbances. More work remains to demonstrate this type of controller in practice for this system, but has been successfully demonstrated on a functionally similar system (mannose).[5]

## Conclusions

In this study we demonstrated the application of a system identification approach for creating a predictive model of %galactosylation for an arbitrary galactose feeding scheme, and the use of that model in a model predictive controller to control %galactosylation in a simulated fed-batch process. The model was generated on cells growth in a pseudo-steady-state perfusion apparatus subjected to step increases in galactose concentration. The process used to generate the data for training the predictive model and the generation of the predictive model itself is highly automatable. Combined with the ability to serialize many step change conditions

during a single pseudo-steady-state experiment, this method represents an efficient way to generate predictive models for use in model predictive control schemes for galactose, and may be generalized for controlling many CQAs.

As knowledge of criticality, and the ability to measure more quality attributes for biotherapeutics increases in the future, so will the need for tighter control of more CQAs. Model predictive control is an attractive method for achieving this control. Using current model generation methodologies, the amount of work required by process development groups to develop model predictive controllers will become prohibitive. This controller development method is one technology that can help lower the barrier for development of MPC schemes, and may catalyze wider adoption of active control of CQAs as a method for achieving QTPPs of the biopharmaceutical products of the future.

## Acknowledgments

The authors acknowledge Clint Pepper, Dave Hansen, and the rest of the MAST™ team at Capsugel for their support in implementing the MAST™ autosampling technology in the perfusion system. They acknowledge MilliporeSigma for the use of their CHOZN® GS−/− cell line for this work.

## Literature Cited

1. FDA. Q8(R2) pharmaceutical development. *ICH Harmon Tripart Guidel*. 2009.
2. FDA. Q9 quality risk management. *ICH Harmon Tripart Guidel*. 2006.
3. FDA. Q10 pharmaceutical quality system. *ICH Harmon Tripart Guidel*. 2009.
4. Jimenez del Val I, Nagy JM, 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.
5. Zupke C, Brady LJ, Slade PG, Clark P, Caspary RG, Livingston B, Taylor L, Bigham K, Morris AV, Bailey RW. Real-time product attribute control to manufacture antibodies with defined *N*-linked glycan levels. *Biotechnol Prog*. 2015;31:1433–1441.
6. Aehle M, Bork K, Schaepe S, Kuprijanov A, Horstkorte R, Simutis R, Lübbert A. Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach. *Cytotechnology*. 2012;64:623–634.
7. Craven S, Whelan J, Glennon B. Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller. *J Process Control*. 2014;24:344–357.
8. Katayama T. *Subspace methods for system identification*. Springer Science & Business Media; Berlin/Heidelberg, Germany. 2006.
9. Plett GL. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 2. Modeling and identification. *J Power Sources*. 2004;134:262–276. doi: 10.1016/j.jpowsour.2004.02.032.
10. Mettler B, Tischler MB, Kanade T. System identification modeling of a small-scale unmanned rotorcraft for flight control design. *J Am Helicopter Soc*. 2002;47:50–63.
11. Angerhofer BJ, Angelides MC. System dynamics modelling in supply chain management: research review. In: *Simulation Conference, 2000. Proceedings. Winter*. Vol 1. IEEE; 2000:342–351.
12. St. Amand MM, Radhakrishnan D, Robinson AS, Ogunnaike BA. Identification of manipulated variables for a glycosylation control strategy. *Biotechnol Bioeng*. 2014;111:1957–1970.
13. Asprey SP, Macchietto S. Designing robust optimal dynamic experiments. *J Process Control* 2002;12:545–556.
14. Zheng K, Bantog C, Bayer R. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs*. 2011;3:568–576.
15. Higel F, Seidl A, Sörgel F, Friess W. *N*-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins. *Eur J Pharm Biopharm*. 2016;100:94–100.
16. Konstantinov K, Goudar C, Ng M, Meneses, R Thrift, J, Chuppa S, Matanguihan C, Michaels J, Naveh, D. The "push-to-low" approach for optimization of high-density perfusion cultures of animal cells. *Adv Biochem Eng Biotechnol*. 2006;101:75–98.
17. Downey BJ, Graham LJ, Breit JF, Glutting NK. A novel approach for using dielectric spectroscopy to predict viable cell volume (VCV) in early process development. 2013; Biotech. Prog.:479–487.
18. Prater BD, Connelly HM, Qin Q, Cockrill SL. High-throughput immunoglobulin G N-glycan characterization using rapid resolution reverse-phase chromatography tandem mass spectrometry. *Anal Biochem*. 2009;385:69–79.
19. Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J Optim*. 1998;9:112–147.
20. Byrd RH, Gilbert JC, Nocedal J. A trust region method based on interior point techniques for nonlinear programming. *Math Prog*. 2000;89:149–185.
21. Qina SJ, Badgwellb TA. A survey of industrial model predictive control technology. *Univ Texas Austin*. 2002;11:733–764.
22. Gramer MJ, Eckblad JJ, Donahue R, Brown J, Shultz C, Vickerman K, Priem P, van den Bremer ET, Gerritsen J, van Berkel PH. Modulation of antibody galactosylation through feeding of uridine, manganese chloride, and galactose. *Biotechnol Bioeng*. 2011;108:1591–1602.
23. Ozcan S, Johnston M. Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. *Mol Cell Biol*. 1995;15:1564–1572.
24. Sasson S, Cerasi E. Substrate regulation of the glucose transport system in rat skeletal muscle. Characterization and kinetic analysis in isolated soleus muscle and skeletal muscle cells in culture. *J Biol Chem*. 1986;261:16827–16833.
25. Loaiza A, Porras OH, Barros LF. Glutamate Triggers Rapid Glucose Transport Stimulation in Astrocytes as Evidenced by Real-Time Confocal Microscopy. *J Neurosci*. 2003;23:7337–7342.
26. Keesman KJ. System response methods. In: Grimble MJ, Johnson MA, editors. *System Identification: An Introduction*. London: Springer-Verlag London Limited; 2011:17–29.

Manuscript received Apr. 20, 2017, and revision received July 14, 2017.