TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Robust and Predictable Mammalian Cell Culture Bioprocesses

Ausgeführt am Institut für

Verfahrenstechnik (166)
der Technischen Universität Wien

unter der Anleitung von Christoph Herwig

durch

Viktor Konakovsky

Porzellangasse 47/15 1090 Wien

II

Acknowledgments

I want to thank Christoph Herwig for shaping me while carrying out my doctoral thesis at his working group and beyond. I am well aware of the unusual academic freedom I could experience under your tutelage and want to thank you for all the interesting discussions, guidance and our special relationship which made me a better scientist, colleague, and person.

I am indebted to my industrial colleagues Christoph Clemens, Martina Berger, Markus Michael Müller, Jan Bechmann, Stefan Schlatter and many more who I had the pleasure to meet in Biberach and Vienna. Your unconditional help and support during all times made this project not only my own, but our shared success, and for this I want to thank you.

I wish to express my gratitude for friendship and comradery experienced at Vienna University of Technology. I very much appreciate the international spirit and social activities that welded us closer together. It would have been half as fun without you and your company.

Hard preliminary work enabled me to join such a well-equipped lab. My thanks go to my knowledgeable and skilled colleagues Christian Dietzsch, Patrick Wechselberger, who set things in motion before my time, and to Andrea Egger, Paul Kroll, Tobias Klein, Matthias Brunner, and all the students who came later to build, develop and maintain the cell culture group at our institute.

My special thanks go to Wei-Shou Hu, who invited me to join his group for a scientific exchange that was generously sponsored by the Austrian Marshall Plan Foundation and supported by the CATMAT doctoral program. The friendly atmosphere at your institute was amazing and I hope I can return the kindness one day. Thank you Tung Le, Andrew Yongky, Kyoungho Lee, Dong Seong Cho, David Chau, and Arpan Bandyopadhyay and all the old and new members of CEMS who I did not mention here for giving me the opportunity to learn from you and for the good time also outside office hours.

Lastly, I want to thank all the people who I did not mention here but had a great influence on me, my research and my professional development before and after this thesis. I appreciate your kindness, experience and early mentorship. I promise to pass on my own knowledge like you did.

Finally, I would like to encourage my fellow colleagues to reach out whenever you can to benefit from a similar experience. It will allow you to see the big picture when the vision is not clear. Thank you all for all lessons learned, I will not forget them.

Dedication

This body of work is dedicated to my two families:

The one I always had and the other one I found.

Abstract (EN)

Within this thesis we implemented online process monitoring and control in our labs using PAT (Process Analytical Technology) tools and scale-free control variables in mammalian CHO (Chinese Hamster Ovary) cell culture processes to improve productivity, robustness and predictability at different scales. Predictive models for cell physiology and sensor signals were established based on historical data analysis of industrial process development data at lab (2L) and pilot plant scale (80L). The developed models proved to be scale-independent and transferable to other CHO clones, which allowed their application on different processes with limited prior information. Optimization of feeding in fed-batch mode was achieved by controlling the specific glucose consumption rate within a narrow range in real time using PAT tools, such as an online metabolic analyser and a capacitance probe for monitoring and control purposes. This led to very stable glucose, lactic acid and pH profiles, improving productivity and robustness of the platform process with scale-free parameters.

Mechanistic, statistical and *in-silico* models under dynamic fed-batch conditions were used to gain novel insights into cell metabolism, and allowed a predictive run forecast at 2L and 12000L scale. The established methodologies facilitate and improve process transfer and scale-up of industrial mAb (monoclonal Antibody) platform processes through advanced process monitoring and control. This is in line with recommendations from the FDA (Food and Drugs Administration) and EMA (European Medicines Agency) to implement PAT & QbD (Quality by Design) approaches in the biopharmaceutical industry to ensure consistent quality of medicines for the safety of patients.

Keywords: Mammalian CHO Cell Culture Metabolism, PAT, Scale-up, Process Modelling, Monitoring and Control

Kurzfassung (DE)

In dieser Arbeit wurden PAT (Process Analytical Technology) Systeme zur Prozessüberwachung und Prozesskontrolle für tierische CHO (Chinese Hamster Ovary) Zellkulturprozesse in unseren Laboratorien eingesetzt. Die Datenanalyse von historische Prozessdaten ermöglichte die Entwicklung von prädiktiven Modellen für Zellphysiologie und Sensorsignalen, welche für die Optimierung von Feedkontrolle eingesetzt wurden und neue Erkenntnisse über den Prozess lieferte. Die Kombination von fortgeschrittener statistischer, mechanistischer und *in-silico* modelierung vertiefte einerseits unser Prozessverständnis und ermöglichte andererseits Prognosen über den Zellmetabolismus sowie optimale Feedingstrategien unter dynamischen fed-batch Bedingungen. Die etablierten Methodolgien vereinfachen und verbessern Transfer und Scale-up von industriellen mAb (monoclonal Antibody) Platformprozessen durch erweitertes Prozessmonitoring und Prozesskontrolle. Das entspricht dem Ruf der FDA (Food and Drugs Administration) und EMA (European Medicines Agency), welche empfehlen, PAT und QbD (Quality by Design) Prinzipien in der biopharmazeutischen Industrie anzuwenden, um eine konsistente Qualität der erzeugten Arzneimittel zu erzielen, die der Sicherheit der Patienten dient.

Keywords: CHO Zellkultur Metabolismus, PAT, Modellieren, Prozessmonitoring und Prozesskontrolle

TABLE OF CONTENTS

1 Introduction

1.1 Problem Statement

In the biopharmaceutical industry, a variety of microbial and mammalian cell lines and clones are used to produce recombinant proteins at different scales. Due to the high level of variability in biological processes, relatively simple open or closed process control loops are the de-facto industrial standard [1][2]. In the simplest case, historical feed rates or feeding frequencies are tested in a previous process and applied with minor adaptations for the next. In more advanced scenarios, feed rates are based on real-time process feedback (closed loop) from i.e. a soft sensor [3][4][5] to estimate biomass and control the feed pumps accordingly. Table 1 summarizes the Measurement, Modelling, Monitoring and Control (M^3C) targets of this thesis, as suggested by the M^3C Working Group. The M^3C is a brand name of the European Federation of Biotechnology (EFB) and the European Society of Biochemical Engineering Science (ESBES) [6][7]. PAT (Process Analytical Technology) technology was applied to improve process understanding, with the aim of increasing the quality and robustness of bioprocesses, which are among the central themes in QbD (Quality by Design) [8]. Ideally, PAT technology could be employed throughout all scales, i.e. monitoring viable biomass in real time to adjust important control variables, such as the feed rate in fed-batch mode. However, this is easier said than done, and a frequent reason for this is that industrial processes undergo a significant scale up from lab to commercial scale. Especially the estimation of viable biomass and its metabolic state cause challenges in the benchmarking and implementation of new technologies [9], some of which were addressed in different peer-reviewed publications that comprise this thesis.

Table 1: Overview of M^{3}C methodology applied in this thesis.

Measurement	Modeling	Monitoring	Control	Benefits	Ref
Capacitance	Statistical model	Biomass	Stoichiometric feedrate	PAT tools were used to construct real-time, statistical biomass models. The models were interchangeable between scales and cell lines, making them universally applicable for the CHO cell fed-batch processes observed in our labs.	[10]
Capacitance, Metabolic Analyzer	Statistical model	Biomass, GLC, LAC, GLN, GLU, NH4, IgG	Low GLC setpoint to control the metabolic state in an industrial fed-batch	Application of the universal biomass model to implement robust process control. Reduced LAC accumulation during fed-batch operation correlated with high productivity, indicating a link between metabolic state and product formation.	[11]
Transcriptome and isozyme composition of different cell lines, clones and tissues	Complex mechanistic model	External and internal metabolites, fluxes and external/internal pH	Cell line engineering and process control parameters to control the metabolic state	Mechanistic modelling revealed potent engineering targets to reduce LAC accumulation by improved clone design. *In-silico* simulation allowed the exploration of possible performance of control conditions, in particular to investigate the effect of pH on metabolism.	[12]
Biomass, GLC and metabolic state estimation with high errors	Simple mechanistic or statistical models	Biomass, GLC, LAC, metabolic state	GLC setpoint, metabolic state	A generally applicable control strategy for novel sensor technologies that may come with high errors. *In-silico* DoEs may be used to select setpoints that lead to both low sampling requirements and low LAC accumulation in industrial fed-batch processes.	[13]

Capacitance probes proved to be the most promising technology for scale up of mammalian CHO (Chinese Hamster Ovary) cell processes. The challenge of estimating biomass across scales and clones could be solved using transfer learning. Unlike machine learning techniques, which try to learn each task from scratch, transfer learning tries to transfer the knowledge from a previous task to a target task, especially when the target task has smaller available data sets [14] or setup/conditions are changed [15], which was both the case here [10]. The biomass model was validated in wet lab experiments in our labs and used successfully to control a dynamic fed-batch process to reduce lactic acid build-up and stabilize the pH profile [11].

While statistical models are very well suited for process control, they might not be very intuitive to study the root cause why clones behave differently. Therefore, mechanistic models of the glycolytic pathway of different tissues, cell lines and clones of CHO *spp.* could help to answer some of these questions. It is clear that the genetic fingerprint of different clones has an

important influence on metabolism, which drives the cells to produce or consume lactic acid, and turn them into low or high producing cell factories [16]. However, when process conditions (i.e. pH and substrate availability) and the genetic background are considered together instead of in isolation, then improved processing conditions to reduce variability may be already found *in-silico*, before a potential high producer is assessed in wet lab experiments [12].

Novel technologies are changing the odds completely and bring new opportunities in the landscape of the production of biopharmaceuticals. These sensors may be faster, in-situ sterilised, cheaper, smaller or otherwise better than current technologies. However, they may also come with a higher measurement error, which could potentially delay their immediate application in industry. To address this challenge, a sensor strategy that accepts high uncertainty and errors of various input signals was developed [13]. The control strategy was tested on process development and industrial datasets and over 500 *in-silico* DoEs (Design of Experiment) led to the best operating conditions for different clones and process scenarios. Our strategy incorporated the uncertainty of especially biomass and metabolic state in its design, which helped to find operating conditions *in-silico* to maximize process robustness by holding glucose concentration setpoints within the desired target range at all times.

1.2 Goals

The goal of this thesis was to establish online monitoring and control in mammalian cell culture processes that were provided by our industrial partner Böhringer Ingelheim (BI). Our working hypothesis is that both todays and future PAT tools together with novel, real-time based control strategies can be readily implemented in modern industrial process formats. The proposed methodology would be applicable also to new processes so that they would become more robust and reliable by such improved process control. For this reason, predictive models for cell physiology and sensor signals were established based on historical data analysis from BI to optimize feed control and to enable predictive run forecast. Our scientific contributions are novel mechanistic and *in-silico* models, under dynamic fed-batch conditions that help to gain

novel insights into cell metabolism and process understanding. The established methods aim particularly at improving process transfer and scale up through better process monitoring and control of the industrial platform process, i.e. by estimating biomass for online feed control. This is in line with recommendations from the FDA (Food and Drugs Administration) and EMA (European Medicines Agency) to implement PAT & QbD approaches in the pharmaceutical manufacturing environment [17][18][19][20]. As more in-line and offline data is captured to help understand potential challenges in the process, a part of our studies featured in-silico experiments for use in truly novel applications. For instance, simulations using transcriptomic data input may be an efficient way to complement screening high producers during cell line development prior to actual bioreactor experiments. In another example, novel and not yet publicly available in-silico inline biosensor data was used to test a very robust control strategy, to be used when the sensors are commercially available. By delivering solutions to both current and future challenges, we expect an increase of productivity and improvement of robustness in the novel process formats, as control strategies are based on metabolic demand rather than predefined empirical trajectories. This will go hand in hand with driving more rapid process development timelines as the developed tools are generally transferrable between new processes with limited amounts of process data.

1.3 Methodology

PAT Tools

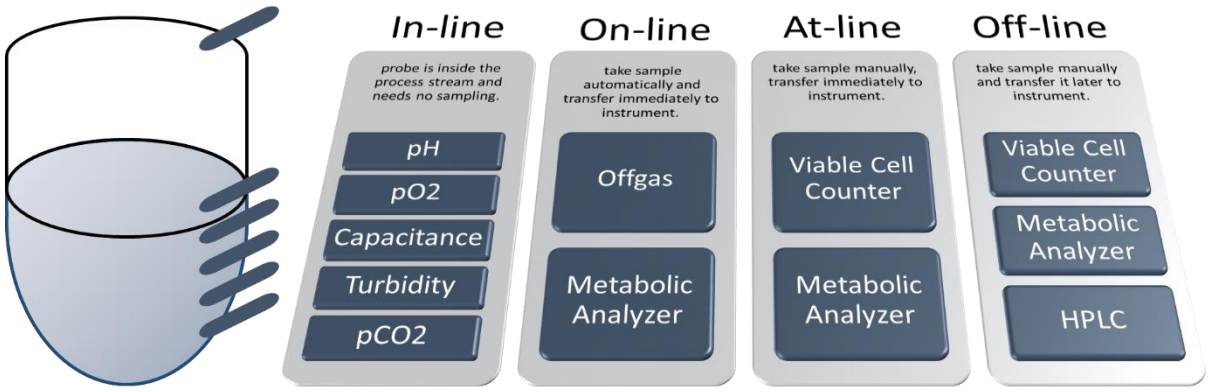


Figure 1: Available PAT Tools in this project.

An industrial mammalian CHO cell bioprocess was transferred from BI to VUT (Vienna University of Technology). Various online, at-line, in-line and off-line PAT tools *(Figure 1)* were tested at VUT and BI. Only the tools that worked best and justified the effort of installation in industry are considered in this contribution. A hard limitation on deciding which PAT tools to use simultaneously was the port space on our and the commercial scale fermenters. Only a careful selection of equipment could be used during the experiments (with the exception of offgas), therefore the risks and benefits had to be evaluated. In this project, we used combinations of pH, pO_2, pCO_2, capacitance, metabolic analyser membrane, inoculation/sampling, feed and inlet gas tubes. Offgas could have been a very valuable measurement taken in addition in our processes, as it would allow calculation of OUR (oxygen uptake rate) and therefore indirectly cell count without invasive technology. However, it turned out to be a very unreliable signal for the industrial platform process that we scaled down in our labs. Pure oxygen, process cascades with non-constant stirrer speeds and pH control using CO_2 together with an overall strong CO_2 buffer background made offgas analysis very challenging at benchtop scale. Regarding the other available sensors, we often had to choose between the pCO_2 and the capacitance probe. We decided to go forwarding using the

capacitance probe, as the pCO_2 probe had a tendency to drift, and because almost real-time monitored biomass was a very important signal that unlocked metabolic control in our processes.

Data analysis

The key to convert data to knowledge lies in structured data analysis. At the beginning, simple visualizations of individual or groups of variables plotted usually against time are tools to grasp any given dataset. Are there outliers, missing values, inconsistencies or artifacts? Which parts of the information require trimming and can all these actions be automated? Typically, simple excel sheets are used for small scale data analyses. However, bioprocesses record data every minute (or second), and excel sheets soon reach the limit of approximately 1 million rows when historical batch data is aggregated. Therefore, other platforms are needed that do not have such limitations. Matlab (Mathworks, Natick, MA, USA) was used to read, process and clean the data in a batch-wise manner. Its extensive library contained several useful algorithms that made routine tasks such as outlier detection (‘rlowess’ [21]), interpolation (‘pchip’ [22]) and curve fitting (‘sgolay’ [23]) much easier (Figures 2A, 2B and 2C). The fact that we were dealing with mostly non-linear curve functions with few data points (i.e. n=10 to 20) puts a big emphasis on robust algorithms that could cope with such challenges automatically in scripts. They were indispensable for obtaining important mechanistic features in our dataset such as specific metabolic rates and yields in our culture that would fuel further modelling efforts.

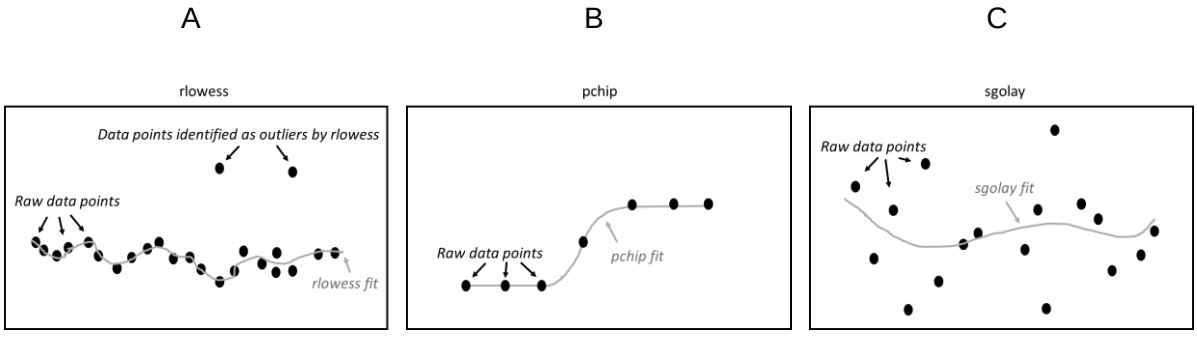


Figure 2: Data processing algorithms that work well for small datasets as encountered in bioprocesses. Robust outlier detection, interpolation of missing values and curve fitting algorithms were important for automatic processing steps during data collection to turn raw data to knowledge. Figures adapted from [21][22][23].

For visualisations of the processed results we used Spotfire (TIBCO Software Inc., CA, USA), a program whose strength is a very strong filtering function to drill down and distil a dataset to its essence, which then can be exported for further analysis. With a tidy, tabular dataset that has meaningful features, modelling and predictive data analysis is possible. The first models were built using Datalab software [24] (kindly provided by Prof. Lohninger, Vienna University of Technology, Vienna, Austria). Other commercial software solutions like Modde and Simca (MKS Data Analytics Solutions, Umeå, Sweden), among many others, allowed a very fast extraction and visualisation of process knowledge from multivariate datasets.

Modelling

Mammalian cell culture exhibits a highly non-linear behaviour which makes it very challenging to represent fully and accurately with simple mathematical models. Cell growth and factors affecting it seem to be elusive, as biological systems, in particular mammalian systems, may react uniquely and time-dependently to what seem to have been identical process conditions or perturbations. It is clear that a simple model will always lack essential parts that represent all possible metabolic behaviours. However, for an aspiring modeller, it may not be so clear that a very well formulated, complex model may be lacking in unexpected areas. For instance, acquiring precise data for variables in a complex model from external sources such as publications that were not conducted with the particular model organism in the own lab, under conditions that almost certainly had an impact on those variables (i.e. medium components that influence cellular substrate transporter affinities), is problematic. It has to be therefore kept in mind that with rising model complexity, the requirements on the amount and precision of data will rise in equal proportion if the model should be predictive and accurate.

Within this project, two areas of modelling were explored in more detail:

i) Multivariate statistical models, which estimate biomass [10] in different processes and clones, allow control of the process in real time [11] using PAT tools.

ii) Complex mechanistic models [12], that combined genetic information of different clones and tissues of CHO cells with industrially relevant process control scenarios. Simple mechanistic models could be used in combination with prior knowledge to simulate likely process behaviour even under large uncertainty [13].

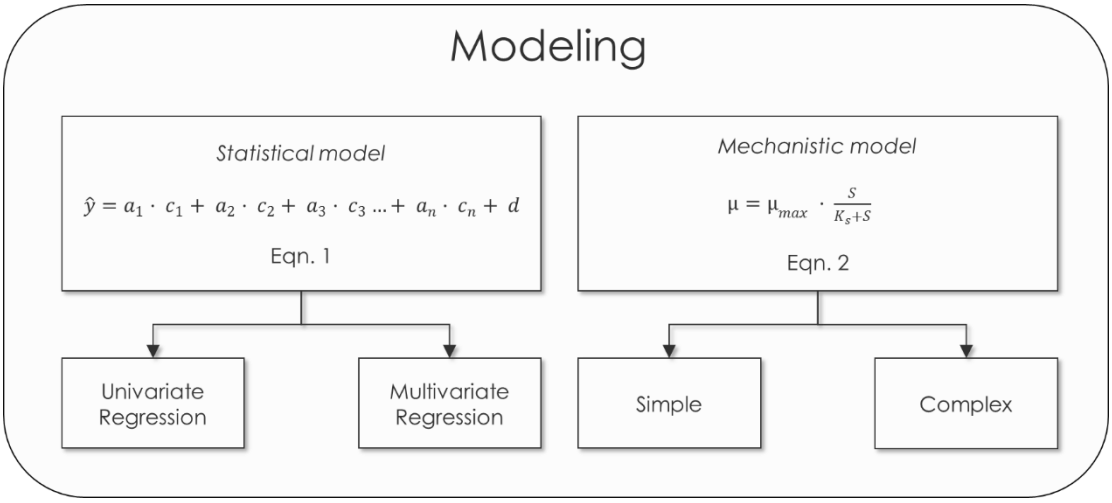


Figure 3: General modelling techniques explored in this thesis. Where possible, simple and univariate techniques are applied. Where necessary, more complex models were employed that could address various challenges in more detail. Multivariate regression models (PLS-R) were more accurate than univariate regression models, while complex mechanistic models helped to gain fundamental understanding on cell metabolism, which simple models could not represent adequately.

In Figure 2, Eqn. 1, $\hat{y}$ represents an estimated variable (i.e. growth or cell count), where $a_{1\ldots n}$ are measured parameters (i.e. capacitance), while $c_{1\ldots n}$ and d are the regression coefficients respectively the intercept of the statistical model. Eqn. 2, μ represents growth rate, which is mechanistically linked to substrate availability S and an empirical coefficient K_s that describes the cell's ability to take up substrate within its maximum possible growth rate μ_{max}. More complex versions of the Monod-Wyman-Changeux Model (often simply referred to as Monod model [25][26]) may add more substrate or inhibition terms, among others, to the base equation. Both Eqn.1 and 2 could be used to model the same variable, and sometimes one

model outperforms the other. While mechanistic models rely on parameters that need to be fitted to a particular mechanistic relationship of parameters in a given dataset, statistical models are fitted to a suggested relationship of multiple factors, in this example in the form of simple polynomials. These could be then used in an exploratory or predictive way, for instance by establishing which relationships may exist, or to estimate a variable that is otherwise hard to measure. We found both routes very attractive to gain more insights in our processes and presented our approach at various conferences [27][28]. Data-driven models may be very powerful as they grow more accurate with even more data and were especially well-suited for process control applications using PAT tools. Both approaches were implemented with success and were important at different stages in the project to predict or estimate variables which are otherwise difficult to monitor in real time [9].

1.4 A short summary of the scientific publications

Universal Capacitance Model for Real-Time Biomass in Cell Culture (Chapter 2.1) deals with the transformation of a capacitance probe's signal into estimated biomass using a multivariate PLS-R model. The aim and novelty of this work was a methodology to construct the model in such a way that it was transferrable between different cell lines and processes, hence universally applicable for a standard CHO platform process. The mean error of our models was in the range of 20% for the whole process time, which made it very suitable for monitoring and control purposes. A simple linear factor was found to be responsible for the transferability of the model, indicating a link to the phenotype or physiology of different cell lines and clones.

Metabolic Control in Mammalian Fed-Batch Cell Cultures for Reduced Lactic Acid Accumulation and Improved Process Robustness (Chapter 2.2) uses the previously developed model to monitor and control a mammalian cell culture fed-batch in real time. The biomass estimation was so good that glucose-limiting conditions could be effectively held in a dynamic

fed-batch, reducing lactic acid build-up by controlling the specific glucose uptake rate in real time within a designed range and increasing productivity at lab scale. Multivariate data analysis (MVDA) was used to rank the importance of process parameters that need to be tightly controlled. The suggested control strategy shows that process-independent variables, such as specific glucose uptake and lactic uptake/production rates, could be tightly and reliably controlled, which resulted in a strong increase of robustness and productivity of the process.

Improved metabolic process control by analysis of genetic clone background in mammalian cell culture (Chapter 2.3) explores the mechanistic reasons behind different clone behaviour regarding lactic acid production as observed in our historical datasets. We wished to explore the inherent reasons for high or lactic acid production in CHO cell culture, and a metabolic model kindly provided by the Hu group was coupled with transcriptomic information from different clones and cell lines. The extended new model involved the different compositions of isozymes involved in glycolysis, and was capable of simulating the transient response to pH, substrate and other process shifts that were relevant for process control purposes. Our simulations revealed an important link between the external and internal pH, which is also reported by other researchers [29][30], where key enzymes in glycolysis are influenced [31][32] and affect cell metabolism as function of the genetic background of the studied clone or cell line. Some enzymes, such as HK (Hexokinase), PFK (phosphofructokinase), PFKFB (6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase), PK (pyruvate kinase) and LDH (lactate dehydrogenase) among others, hold a key role in the lactic acid metabolism [33][34][35][36][37]. Our simulations under industrially relevant process conditions showed that these enzymes and their isoforms (i.e. HK-IV [38] and LDH-C [39][40]) may be promising targets for epigenetic engineering and cell line development activities for more robust cell culture processes. Novel clones may be engineered by design, i.e. by screening for more variants which are more tolerant towards high glucose concentrations and reduced lactic acid production in industrial processes.

A robust feeding strategy to maintain set-point glucose in mammalian fed-batch cultures when input parameters have a large error (Chapter 2.4) builds on the extended modelling activities and distils the most important aspects of various clone behaviour into a simpler model representation. The clone-specific behaviour was thus coupled to an *in-silico* bioprocess simulation and results were, among others, ideal feeding trajectories and sampling times to hold glucose concentrations within a desired control range. This control strategy was tested for lab scale processes featuring a process shift, on an industrial dataset at commercial scale (12000L) and on several synthetic conditions that were obtained in a DoE, by turning critical assumptions into variables to verify adequate robustness of the control strategy. In brief, the developed algorithm was capable of detecting and slightly correcting when a process was going out of specifications in real time by using uncertainty trajectories. This study demonstrates that a robust and sufficient level of control could be demonstrated even with high errors for biomass, glucose or metabolic state estimation, as could be soon the case when novel sensor technologies are implemented that come with a high error.

1.5 The scientific contribution to the field of study

In the beginning of our journey to employ PAT tools for bioprocess monitoring and control, we expected the novel sensors to be well-developed, so that the sensor signal itself could be used directly and immediately to monitor and control the process. However, almost always some sort of data treatment was required, which led to the development of soft sensors. The soft sensor we developed was an estimation of biomass in CHO cell culture, which was adequate to monitor and control one of the most dynamic and complex process formats in fermentations: mammalian cell culture fed-batches. Our scientific contribution to the field lies in the way how the soft sensor was constructed. Extreme care must be taken when data-driven models are employed in predictions, as statistical models tend to overfit. If this was the case, it would be highly specific to the training dataset, but less transferable and applicable to new cases.

However, our models were built not only with data from different scales and validated on different clones, but also built with different clones at different scales. As a consequence, the test set was very robust from the start. This level of diversity was required to build a highly transferable biomass model that we could validate on a total of 4 clones and 2 scales, which was a quite remarkable accomplishment for a statistical model.

With a working soft senor to estimate biomass, we could investigate in detail how exactly glucose and pH interact to drive metabolic lactic acid production or consumption. Although literature states that higher (more alkaline) pH positively influences cell growth, and simultaneously always increases lactic acid production, by decoupling both effects we could demonstrate that this does not have to be the case. A very high pH setpoint could be selected, which may have influenced the internal pH and with it many internal enzymes [29], to enable fast cell growth. At the same time the specific glucose uptake rate could be controlled within the range of error for biomass and metabolic state estimation in a glucose limited fed-batch mode. With very precise knowledge of biomass concentration and a closed-loop control of the metabolic state, we witnessed the lowest lactic acid levels and the highest productivity of this process in our labs, which is consistent with the findings of other researchers [34][41][42][43].

A research exchange at the Hu group (CEMS, MN, USA) was undertaken to investigate and explore lactic acid behaviour. A mechanistic model that described a part of the central glycolytic pathway was kindly provided by the host lab. It was extended to include typical industrial process conditions such as external and internal pH gradients, different isozyme compositions and expression levels. Then, the model was used *in-silico* to simulate lactic acid and pH evolution with different model tissues and clones of CHO *spp.* origin [12]. Many factors, including process conditions (GLC, LAC, pH), different isozymes (HK, PFKFB, PFK, PK, and LDH), the energy pool (NAD/NADH), among others, were contributing and driving the cell's metabolic behaviour [35][37]. In conclusion, it was shown that not just one, but several strategies can be employed synergistically to control the cell metabolism by design. Our contribution integrated protein-specific substrate specificity, genetic distribution of isoforms

and process control by pH shifts or glucose pulsing in a systems approach to make a qualitative statement about a screened clone's potential future metabolic behaviour.

With more in-depth knowledge on how to control lactic acid concentrations, we returned one more time to use these insights in combination with PAT tools for process control. The glucose consumption rate is the number one parameter that must be controlled within tight ranges because it impacts on how substrate is utilized [44][45], and whether precursors turn into lactic acid or into product. Our complex mechanistic model was reformulated as a minimalistic mechanistic model and fed with process development and manufacturing data. It is unlikely that any sensor or estimation method will ever be so accurate that there is absolutely no error. Therefore, we implemented errors on biomass estimation, the metabolic state and other critical parameters for a more robust control strategy [13]. The output was what we called uncertainty trajectories, which led to an automatic, slight adaptation of the feed rate to stay within a defined range around the glucose setpoint. Process development and manufacturing data was used to challenge the algorithm by permuting critical assumptions that could affect its robustness, such as the selected setpoint range, error on measurements and clone-specific substrate affinity resp. lactic acid behaviour. The main scientific output was a glucose setpoint control strategy for mammalian cell culture fed-batch modes, which allowed sufficient control, even if the available sensors and assumptions were not very accurate (>50% error). We are therefore confident that even novel sensor technologies [46] with potentially high errors in full scale could be used for process control purposes, speeding up development activities in a perpetually changing industrial environment.

1.6 Author contributions

The author contributions are listed together with the major scientific focus of each contribution, which is depicted below as shaded area in four major topics: cell metabolism, data analysis, modelling, and process control. The individual chapters of this thesis are publicly available at the individual journal publishers at the time of writing. Chapter 2.3 is publicly available by the Marshall Plan Foundation, which holds the right of publication as specified in the Marshall Plan Foundation Grant Agreement. The other studies are peer-reviewed and DOI numbers are available: Ch.1: 10.3390/s150922128, Ch.2: 10.3390/bioengineering3010005 and Ch.4: 10.1002/btpr.2438.

Table 2: Author contributions and area of scientific focus in the publications.

Manuscript	Ch.	Author contributions	Mammalian CHO cell metabolism	Soft sensors, data analysis and MVDA	Statistical and Mechanistic Modelling	Process Control using PAT Technology
Universal Capacitance Model for Real-Time Biomass in Cell Culture, Sensors	*2.1*	VK designed and performed a part of the experiments, analyzed the data, developed the universal models, performed statistical analysis, prepared tables, figures, additional files, drafted and wrote the manuscript.		■	■	
Metabolic Control in Mammalian Fed-Batch Cell Cultures for Reduced Lactic Acid Accumulation and Improved Process Robustness	*2.2*	VK designed, planned and performed the experiments at Vienna University of Technology, analyzed the data, performed statistical analysis, prepared tables, Figures, additional files, drafted and wrote the manuscript.	■	■	■	■
Improved metabolic process control by analysis of genetic clone background in mammalian cell culture	*2.3*	VK designed and performed the simulations, wrote the manuscript, prepared tables, figures and adapted the source code for the simulations.	■		■	
A robust feeding strategy to maintain set-point glucose in mammalian fed-batch cultures when input parameters have a large error	*2.4*	VK designed, planned and performed the experiments at Vienna University of Technology, built the model and designed the algorithm, set up, analyzed and interpreted the DoE, analyzed the results, performed statistical analysis, prepared Tables, Figures, additional files, drafted and wrote the manuscript.	■	■	■	■

2 Scientific publications

2.1 Chapter 1

Sensors **2015**, *15*, 22128-22150; doi:10.3390/s150922128

OPEN ACCESS

sensors

ISSN 1424-8220

www.mdpi.com/journal/sensors

Article

Universal Capacitance Model for Real-Time Biomass in Cell Culture

Viktor Konakovsky [1], Ali Civan Yagtu [1], Christoph Clemens [2], Markus Michael Müller [2], Martina Berger [2], Stefan Schlatter [2] and Christoph Herwig [1,*]

[1] Institute of Chemical Engineering, Division of Biochemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A 166-4, 1060 Vienna, Austria; E-Mails: vkonakovtuwien@gmail.com (V.K.); civan.yagtu.tuwien@gmail.com (A.C.Y.)

[2] Boehringer Ingelheim Pharma GmbH & Co. KG Department Bioprocess Development, 88400 Biberach, Germany; E-Mails: christoph.clemens@boehringer-ingelheim.com (C.C.); markus_michael.mueller@boehringer-ingelheim.com (M.M.M.); martina.berger@boehringer-ingelheim.com (M.B.); stefan.schlatter@boehringer-ingelheim.com (S.S.)

* Author to whom correspondence should be addressed; E-Mail: christoph.herwig@tuwien.ac.at; Tel.: +43-1-58801-166-400; Fax: +43-1-58801-166-980.

Academic Editor: Alexander Star

Received: 17 April 2015 / Accepted: 25 August 2015 / Published: 2 September 2015

Abstract: Capacitance probes have the potential to revolutionize bioprocess control due to their safe and robust use and ability to detect even the smallest capacitors in the form of biological cells. Several techniques have evolved to model biomass statistically, however, there are problems with model transfer between cell lines and process conditions. Errors of transferred models in the declining phase of the culture range for linear models around +100% or worse, causing unnecessary delays with test runs during bioprocess development. The goal of this work was to develop one single universal model which can be adapted by considering a potentially mechanistic factor to estimate biomass in yet untested clones and scales. The novelty of this work is a methodology to select sensitive frequencies to build a statistical model which can be shared among fermentations with an error between 9% and 38% (mean error around 20%) for the whole process, including the declining phase. A simple linear factor was found to be responsible for the transferability of biomass models between cell lines, indicating a link to their phenotype or physiology.

Keywords: CHO cell culture; capacitance; fed batch; PLS; statistical model

1. Introduction

1.1. Problem Statement

One of the most important parameters in microbial or mammalian cell culture is the viable cell concentration (VCC). It is permanently subject to change in a typical batch or fed-batch process. VCC is usually measured offline by a cell counting device, but optimally, VCC could be also modeled without taking sample, *i.e.*, by an inline capacitance probe. The measurement principle of a capacitance probe is frequency-dependent polarization of dielectric material which enables the detection of the living spherical cells in form of capacitance [1–4]. This task is made more difficult when particles of varying shape and size are co-measured, which is the case when cells die and fragment. However, multivariate approaches are useful to filter out the noise [5,6] and can be used to model VCC during the whole process time instead of stopping when viability drops.

1.2. State of the Art

The models which describe the relationship between capacitance and VCC, are often based on linear regression, multiple linear regression, Cole-Cole and PLS. Applications of permittivity measurements with single and multi-frequency measurements were well summarized by Yardley *et al.* [7]. Noll and Biselli used linear regression to set a particular feed rate based on a constant glutamine consumption in continuous cell culture [8]. Zeiser [9] and Ansorge *et al.* [10] used multi-frequency permittivity measurements to monitor process events such as timepoint of infection and virus release in Sf-9 cells (with baculovirus), and later also in HEK cells (with lentivirus) [11]. While multivariate frequency measurements are more informative than classical single or dual frequency measurements, they are also much harder to interpret [1]. Opel *et al.* [5] described various methods to correlate VCC to capacitance in batch and fed-batch in great detail; in brief, linear models required frequent recalibration, Cole-Cole models were reliable during the cell growth phase, but not in other phases, and PLS models were accurate during the whole process phase, but required robust calibration datasets and complex analysis. There seems to be a general lack of motivation to publish data of the predictive accuracy of any capacitance model for a long-term mammalian cell cultivation. We hypothesize that one of the reasons for this may lie in the difficulty of standard models to describe the declining phase of a culture sufficiently well, as already mentioned by Cannizzaro [12]. Once the monitoring of cell concentration is accurate enough, control, in the form of closed-loop feeding strategies, might be employed to push cell count and titers in the process [13–17].

Conclusively, multivariate frequency models which were built with PLS regression from capacitance measurements had one particular benefit over all other models: They allowed the modeling of VCC also in the declining phase, implying that recalibrations may not be required. Instead, the transformed signal can be used for the whole process phase of a dynamic fed-batch.

1.3. Novelty of This Approach

PLS models of multivariate signals usually result in very low errors when created offline, *after* the process is over. These low estimation errors can become unacceptably high if the same model is applied as it is for the next process. The knowledge of what to include into the model before its construction is very important, as it will determine both reusability, transferability and with it the overall benefit of the signal.

Capacitances measured at high frequencies correlate with particles which are smaller than the viable cells [12,18]. In a statistical model, such as PLS, every single frequency is mean centered or standardized before calculating the coefficients and therefore has a certain weight in the analysis [19,20]. Particles (cell debris) may be detected and interpreted as biomass if those high frequencies are included. Capacitance maps of a multi-frequency scan can help to identify these variables, which must be excluded in order to generate a robust model. Afterwards, the same PLS model can be re-used for estimating VCC of a new clone by just adapting the model's slope. This made the constructed PLS models so remarkably universal that they could be successfully applied from one lab in Germany to another one in Austria.

Conclusively, we found the combination of two points to be decisive for a transferable model based on capacitance: (i) variable selection *before* and (ii) slope adaptation *after* model construction. The knowledge of VCC in the system without the need of sampling unlocks radical new control strategies and concepts for improving cell culture bioprocesses in both process development and manufacturing.

1.4. Goal

The goal of this work is to establish a *transferable* VCC estimation model in CHO cell lines regardless of clone, setup or scale.

1.5. Roadmap and Workflow

We propose a novel method to use capacitance spectra for estimating biomass without further operator intervention in mammalian fed-batches. Typically, linear regression methods with just one frequency subtracted with the highest frequency (so-called "dual mode" to remove the influence of medium, *etc.* on measurement) are often enough for many short-term cultivation formats such as batches or short fed-batches while viability is still very high. However, continuous cultures and perfusion processes, where interfering cell debris is flushed out and cells are kept in a highly viable state for an extended period can also be described well with linear models [12,21]. In this work, we have found that PLS models are the ones which fit best to fed-batch formats, where viable cells undergo a dynamic growth and decline phase. Usually, useful statistical models depend strongly on high-quality historical datasets. Our approach was the transfer of *already existing* models instead of constructing new ones. Finally, we demonstrate that a transfer of the same model is not only possible between different clones but also between different scales. For this reason a suitable acceptance criterion was tested. Hence, this is a universal model for biomass estimation relying solely on real time information (Figure 1).

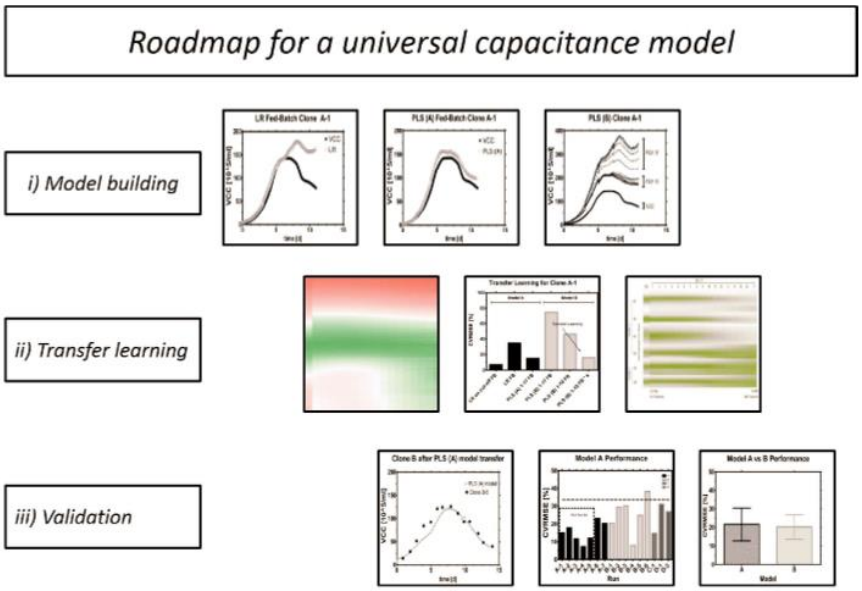


Figure 1. Roadmap.

2. Experimental Section

2.1. Process Setup

Different CHO cell lines (A, B, C, D) producing different mAbs were cultivated in shake-flasks (Corning Incorporated, Corning, AR, USA) in incubators (Minitron, Infors, Bottmingen, Switzerland) with 5% partial CO_2 pressure at physiological temperature (35–37 °C) on orbital shakers at 120 rpm (orbit 50 mm). Passaging was performed every third day in proprietary serum-free media and the bioreactors were inoculated at an initial seed density of (3–10) × 10^5 cells/mL. At Vienna Technical University, different bioreactors (2.0 L operating volume, Infors HT, Bottmingen, Switzerland) were used than at Boehringer Ingelheim (Ingelheim, Germany) (2.0–80 L operating volume, no information on manufacturer).

2.1.1. Data Source

Mammalian fed-batch fermentations using a capacitance probe (Biomass Sensor, Hamilton Bonaduz AG, Bonaduz, GR, Switzerland) in scanning mode were used to construct various linear and multivariate models. The bioprocesses were subject to variations in scale (2–80 L), internal bioreactor architecture (including different impeller types, H/D ratios and glass or stainless steel shell), setpoints of pH and pO_2 (7.0 ± 0.2 resp. 5%–60%) (Hamilton Bonaduz AG, Bonaduz, GR, Switzerland), stirrer cascades, bolus additions of concentrated feed solutions, media and feed composition (both chemically defined and

serum-free, proprietary formulations) and feed strategy. Temperature was held constant in all processes at 37 °C.

Capacitance measurements were recorded in mammalian bioreactor runs varying in scale, cell line and bioreactor geometry between sites. A summary of different setups can be seen in Table 1. On the top row, the individual experiments are discriminated against scale in columns, the second hierarchy is the cell clone in columns. The experiments are listed in rows. The first model (PLS-A) was created from the largest dataset using run A1–A5 and validated on runs A6–A7. The second model (PLS-B) was created from B1–B3 and validated on B4–B6. After model transfer, PLS-A could be used for Clone B, C and D and Clone B could be used for Clone A, C and D, regardless of scale (see Results and Discussion section as well as supplement).

Table 1. Crosstable fermentation conditions.

Run	Scale 1 (80 L)			Scale 2 (2 L)		Chapter
	Clone A	*Clone B*	*Clone C*	*Clone B*	*Clone D*	
A1	x					**Finding the best model**
A2	x					
A3	x					
A4	x					
A5	x					
A6	x					
A7	x					
B1		x				**Variable selection**
B2		x				
B3				x		
B4				x		
B5				x		**Transfer learning**
B6				x		
C1			x			**Validation**
D1					x	
D2					x	

2.1.2. Media

Media for the fed-batches are proprietary in composition and subject to variations in starting levels of metabolites, growth factors, *etc.* All components were serum-free and chemically defined.

2.1.3. Cell Lines

All cultures were engineered CHO cells. Suspension of Clone B cells for experiments at the VUT were kindly provided by Boehringer Ingelheim (Ingelheim, Germany), while Clones A, C and D were used for model building and validation.

2.1.4. Analytics

Process information was logged using the process management system Lucullus (PIMS, Lucullus, Biospectra, Schlieren, Switzerland). Capacitance spectra were recorded by exposing the medium broth via inline probe to an excitation frequency ranging from 0.3 MHz to 10 MHz every minute. The full capacitance spectrum (all 17 frequencies) was recorded but not all capacitance values were used to construct linear and multivariate models. Multivariate modeling was performed using Datalab software [21] (kindly provided by Prof. Lohninger, Vienna University of Technology, Vienna, Austria). The concentration of total cells, viable cells and viability was measured using an image-based white/dark classification algorithm after automatic trypan blue staining integrated in the cell counter (Cedex HiRes, Roche, Basel, Switzerland). Main metabolite concentrations (Glucose, Glutamine, Lactic acid, Glutamic acid, NH_4, IgG) were measured on-line and off-line using a photometric robot (CubianXC, Optocell, Bielefeld, Germany). Amino acid concentrations were measured off-line by HPLC using pre-column OPA-derivatization (Agilent, 1100 HPLC, Santa Clara, CA, USA).

2.1.5. Multivariate Data Analysis

The SIMPLS algorithm integrated in Datalab software was used to calculate all regression coefficients for any given model. These coefficients are finally multiplied with all or only a selection of available independent variables (capacitance measured at certain frequencies) to estimate the dependent variable (VCC). VCC (~20–80 data points) was interpolated offline but also in real-time when required, to match the very frequently measured variable capacitance (~500 data points) using the *spline* function (MATLAB) which preserves the curvy character of VCC in-between offline measurements. In this contribution, three Principal Components (PC) held over 99% of the whole variance while the remaining one percent held mostly noise and could be omitted from the analysis. All PLS models with more than three predictor variables (frequencies) were built with three PCs to predict VCC to make the models comparable. For the special cases that three PCs could not be used because fewer signals were used as input, the number of PCs had to be reduced to two resp. one (this was only the case where one or two frequencies were used for model construction). A more thorough explanation of constructing such a multivariate model [22–25] and its interpretation [19,26,27] is given elsewhere.

In total, 34 PLS models were constructed from Clone A data using standardization and mean centering, and 34 PLS models from Clone B data. From these 68 PLS models, 64 could be constructed with three PCs which made the models comparable. All models are also provided in the supplement and we want to invite the reader to test them with their own data.

$$\hat{y} = a_1 \cdot c_1 + a_2 \cdot c_2 + a_3 \cdot c_3 \ldots + a_{17} \cdot c_{17} + d \quad (1)$$

In Equation (1), $\hat{y}$ represents the estimated VCC in (10^5/mL) while $a_{1\ldots17}$ are the individual capacitance values in (pF/cm) at a certain frequency (*i.e.*, 100 (pF/cm) at frequency 0.3 MHz) and $c_{1\ldots17}$ are the regression coefficients for this frequency in (*i.e.*, −1.5 (cm × 10^5)/(pF·mL)), calculated by PLS-Regression. The intercept d of the model is given in the same units as VCC (10^5/mL).

2.2. Acceptance Criteria and Control Specifications

The model's root mean square error (RMSE) and its coefficient of variation (CVRMSE) were calculated for both the exponential phase and the whole process phase.

$$RMSE = \sqrt{\frac{\sum (y - \hat{y})}{n-1}} \quad (2)$$

$$CVRMSE = \frac{RMSE}{\bar{y}} \quad (3)$$

In Equations (2) and (3), y represents the measured VCC in (10^5/mL), $\hat{y}$ the estimated VCC in (10^5/mL), and n is the number of observations and $\bar{y}$ represents the mean of the estimated VCC in (10^5/mL) multiplied with 100 to obtain the CVRMSE in (%). All calculations are also available with examples in the supplement.

A robust model acceptance criterion that is still practical for purposes beyond process monitoring is simply not yet defined. Therefore, the following criterion is suggested to be acceptable from a process engineer's point of view but will have to prove itself in practice.

CVRMSE

The CVRMSE does optimally not exceed 25% for the exponential growth phase and 33% for the whole process phase. An absolute upper limit to reject the model is defined at 50%. Model estimations which deviate between 33% and 50% CVRMSE are not rejected but it should be questioned if such deviations work for the given problem statement. Often, gross errors which include handling, liquid level, extreme process events, contamination, massive changes in large or small particle load in the broth and changes associated with cell morphology or membrane potential as well as flawed reference measurements (*i.e.*, after dilution or wrong classification between dead and viable cells) are the prime reasons for offsets [3,5,10].

The herewith suggested model acceptance may seem rather broad. However, the goal of this contribution is not running behind theoretical estimations which sound great on the paper but are absolutely unrealistic if any seemingly insignificant parameter during the process is changed. This contribution compares fermentations from process development where different media and feed formulations, reactor geometry in the same scale and between different scales and clones, process set-points and fermentation conditions with and without sudden feed additions and various process events interfere with regular fermentation profiles according to the book. Capacitance measurements are furthermore also sensitive towards other capacitors and signals affecting the measurement such as metallic probes, stirrers or different grounding in various scales and set-ups. The required offline analysis was performed by various operators working at different sites and although the protocols were the same, the same analysis can yield slightly different results when the devices or time between sampling and measurement are subject to variation.

It is absolutely necessary that this novel method holds its promise to perform well under authentic process conditions with often orthogonal process conditions as well as in standard processes, which are easier to maintain and evaluate. Other research groups used the capacitance signal with linear regression

for control purposes [8], and modeled VCC with PLS using multiple frequencies *after* the process was completed, but a process where PLS with multiple frequencies was used to control a process has not been yet reported. We want to stress that using linear models requires offline calibration while PLS models could be calibrated once and used for the whole process time, if they are robust enough for the specified control purpose. Most contributions analyzed VCC with PLS models in hindsight and found low errors for this method, which might not tell the whole story, as the analysis method and selection of validation has a significant impact on the final error. What is often missing and therefore misleading is a realistic description of the performance of both linear and PLS models.

Our data suggests that the expected CVRMSE from LR models can be twice as high as CVRMSE in PLS models. The average CVRMSE from PLS models in this contribution was 21% ± 9% CVRMSE. Therefore, we are confident that the suggested criterion of maximum 33% CVRMSE is not only realistic, but also very suitable for this model and will find many applications in instances where VCC may be required as an input.

3. Results and Discussion

3.1. From Signal to Model

A typical capacitance spectrum is shown in Figure 2. A trend can be seen when capacitance measured at frequencies between 0.3 and 10 MHz is plotted together with viable cell count (VCC) and total cell count (TCC) over cultivation time. However, the trend of total cell count does not match as well as the one from VCC after cells start dying, which can be explained by the measurement principle of this probe [28]. Viable cells are able to sustain an intact cell membrane and build up capacitance when charged by a specific frequency, while dead cells, which have lost the ability to keep charged ions in their cytoplasm, do not contribute to capacitance in the same frequency range. This means that in theory, the probe is suited to detect viable cells very selectively.

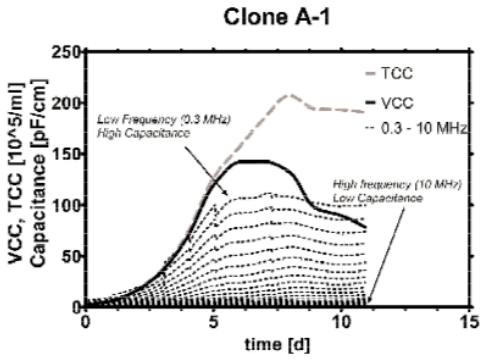


Figure 2. Signals over time: In this plot, VCC (solid line) and TCC (dashed line) are plotted together with capacitance values induced by varying frequencies (dotted lines). The lowest frequency (top dotted line) induces the numerically highest capacitance and reaches maxima around the time when VCC and TCC reach their maxima.

3.2. Finding the Best Model

As the first step, linear models were constructed for one Clone (A), scale (1) and fed-batch (A-1). Various phases were compared (highly viable fed-batch, full fed-batch in Figure 3) and discrepancies in performance found (see supplement for more data). As univariate models were not able to describe the declining viable cell concentration sufficiently, multivariate models were employed to find yet overseen correlations by weighting the coefficients of capacitance signals differently. As seen below, the multivariate model was best suited to predict VCC in a typical fed-batch cultivation.

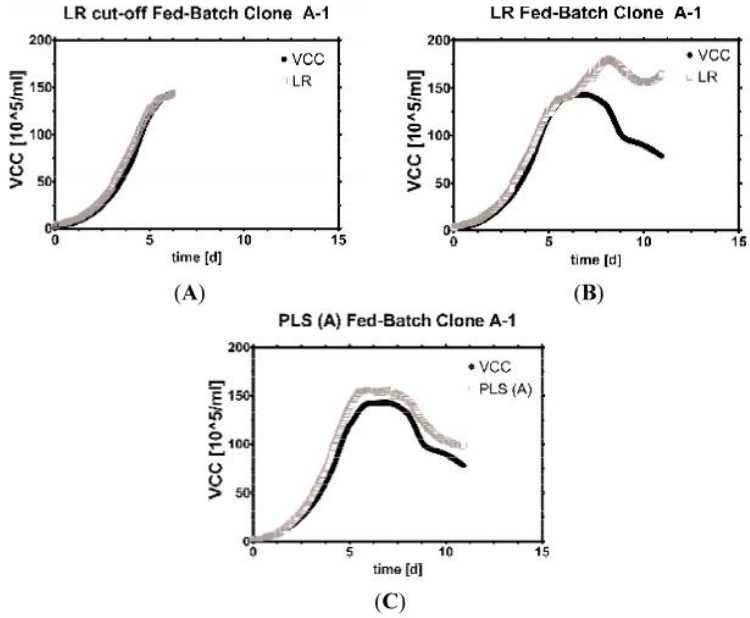


Figure 3. Model development. Fit *versus* measurement: (**A**) (cut-off) Fed-Batch while viability was high with a univariate model; (**B**) (full) Fed-batch with a univariate model; (**C**) (full) Fed-batch with a multivariate model.

3.2.1. Linear Model

Linear models can be employed in all batch-type and even short fed-batch processes, *i.e.*, in virus production to seed trains for inoculation of production scale bioreactors [29,30]. Which frequency works best in describing VCC is somewhat disputable. Yardley [7] has summarized frequently used frequencies for mammalian cultivations which range from 0.5 to 0.8 MHz and a potential extension of up to 3 MHz. Our own observation is that in standard short-term fermentations, even the highest frequencies (~10 MHz) correlate excellently with VCC ($R^2 > 0.99$, see supplement). Therefore, the standard frequency provided by the manufacturer was chosen (around 1 MHz) for LR (Figure 3A).

The trouble with linearly estimated VCC begins when cells leave the exponential growth phase (Figure 3B) in fed-batch cultivations. To be more precise, the calibrated probe mostly overestimates viable cell count, sometimes by a constant offset, other times accompanied with one or more spikes (Figure 3C). Thus, a linear calibration model can always be used before VCC peaks and only with great care or recalibrations later on. Whether VCC rises or falls is exactly the information one seeks in a fed-batch, and is very hard to answer in real time with only one single frequency, no matter which one was chosen.

Some authors therefore proposed to make use of an adapting frequency which is dependent on several cell-specific factors and can change during the course of a fermentation. In brief, this so-called critical frequency *fc* shall explain the deviations in a mechanistic way by pricing-in influences ranging from cell size change, conductivity and others [7,31,32], thus extending the regular linear model by several terms and parameters for which a mathematical fit in a new model (Cole-Cole) is required. Although the concept is without a doubt plausible, all fitted coefficients require calculation, pre-smoothing steps of the data and might differ under different process conditions, reactor geometries, physico-chemical conditions and so on. The factors might therefore require readjustments after every change of the bioprocess which might be more problematic in bioprocess development runs than in established process formats. This dataset was recorded in development format and therefore subject to severe disturbances where the spectra were very difficult to interpret. Because of transferability concerns, the development of mechanistic models is acknowledged [10,33–35] but not pursued in more detail.

3.2.2. Multivariate Model

By using all frequencies instead of just one, these challenges are accepted by employing statistical multivariate models such as Multiple Linear Regression (MLR) or Partial Least Squares Regression (PLS or PLSR) to predict the target variable VCC more accurately. Other possible methods would encompass Principal Component Regression (PCR) [25], or Artificial Neural Networks (ANN) [36]. MLR (and also PCR) models tend to overfit any given dataset if enough data are available [20], while constructing an ANN was out of scope of this contribution, which should demonstrate the transferability of one designated model. Therefore, a PLS model was constructed with five runs of cell line Clone A and fit to the remaining two runs. The fit had an $R^2 > 0.95$ and was useful for estimating VCC under the same conditions (Figure 3A,B). Similar results were found for Clone B constructed in two different scales as seen later on in this contribution.

3.3. Multivariate Variable Selection

3.3.1. Scaling

Two methods were used to process capacitance data: Standardization (SD) and Mean Centering (MC). The difference is explained well in literature [37,38]. In general, MC seems to be used more frequently to construct PLS models.

The interpretation of coefficients derived from a statistical model could lead to an over-interpretation of the results. All capacitance signals alone correlate naturally positively with VCC, but due to the model fit some variables are finally attributed to a negative VCC trend which seems to not comply with the

measurement principle. This may also be the reason why PLS models are readily created, but in contrast to linear regression the coefficients are actually never explicitly shown [5,12,39].

The final weighting procedure in PLS represents and considers also other effects in the course of the fermentations which are not immediately recognized. For instance, in our experiments, almost all multivariate models of this contribution correlate capacitance at a high frequency with a decline in VCC. When the highest frequency is removed from the model, the next-highest frequency takes the former's role until in a linear model only one frequency is left (as included in the supplements).

It is hypothesized that this is probably caused by late-stage discrimination between VCC and charged, small particles, namely cell fragments and partially organelles, which store electric charge in form of capacitance and perhaps the main reason why this kind of model performs so well. Cannizzaro used the non-viable part to build a PLS calibration model from only dead cells [12]. The quantification of particle count proved to be so extremely difficult (CVRMSE between 33% and 141% in the calibration set and 77%–192% in the validation) that the information is of rather qualitative than quantitative nature. The main question, whether the model coefficients for only viable cell count are so robust that they are also valid for other systems becomes apparent when a transfer between clones and scales is demonstrated.

3.3.2. Capacitance Maps

Ideally, there are no other dielectric materials except the viable cells, but in practice, a fermentation broth contains also dead, apoptotic, enlarged or fragmented cells and particles [8]. A heat map is constructed, where the individual capacitance values are represented as colors. These herewith called "capacitance maps" can give the operator a visual information of the distribution of dielectric material before a model is constructed.

Clone-specific capacitance maps were used to elucidate which frequencies are suitable for the model. The capacitance map displays VCC on the very left column, the time course of the fermentation follows from top to bottom in every column. The numerically largest signals can be extracted from low frequencies (on the left, starting with 0.3 MHz), while very high frequencies (on the right, 10 MHz) were numerically small, carrying almost no information except mostly noise. However, if the values are normalized, an interesting picture is obtained; different reactor geometries and cultivation conditions were found to have a significant impact on the capacitance fingerprint profile as can be seen in Figure 4. More data are available in the supplement. The new, potentially transferrable PLS models were constructed using only frequencies between 0.3 and 2.16 MHz. These new specifications were further tested in different clones, setups and scales.

3.4. Transfer Learning

A change leading to new measurement conditions, be it through a slight modification in substrate levels [1,4] or other modifications of the environment [34,40–42], leads to a common problem in chemometrics: The direct application of the model is no longer possible [43]. Finding a way to model the target variable under the new measurement conditions using either new calibration data or a new calibration model by transfer of knowledge that has already been learned is called transfer learning [44]. In this contribution, the concept of transfer learning was used by applying previous knowledge from simple linear regression to a more complex PLS model. To our knowledge, no

multivariate biomass model was transformed successfully in such a way before. It is hypothesized that the new, multivariate signal behaves similarly to the univariate capacitance signal, but without the noise in the data. These multivariate models are potentially more useful than univariate models and in addition universally shareable.

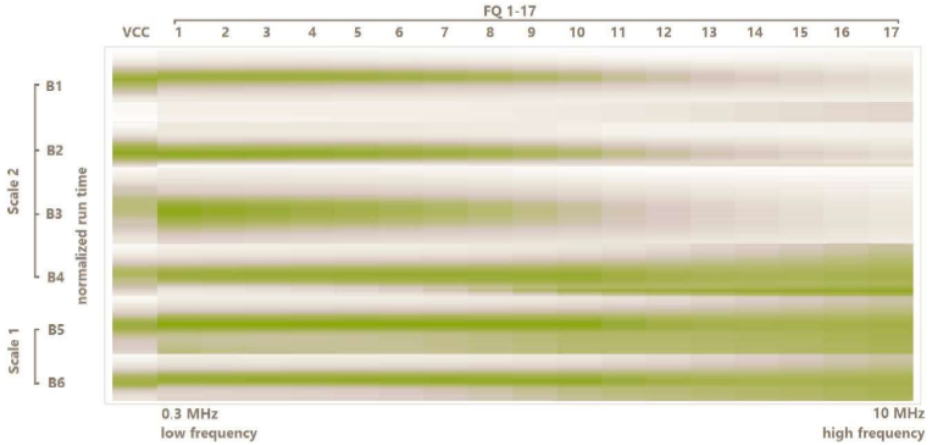


Figure 4. Normalized capacitance map: Normalized VCC are located on the very left, normalized capacitances in all remaining columns. High (left) and low (right) capacitance at different frequencies (FQ 1–17, from 0.3 to 10 MHz). The colors are relative to all runs, which means the maxima of all runs are green and the minima are white. Different scales are indicated. Frequencies which would later negatively influence the estimated VCC in the model can be thus identified visually. The region of interesting frequencies for model transfer encompasses frequencies 1:10.

3.4.1. Direct Model Transfer

A cross-validation of PLS models was not performed because it is clear that any coefficients eventually will be found which fit to the given data. The effect of overfitting coefficients can be seen in Figure 5. Such coefficients might show their best performance for very similar runs under the same conditions (same clone or scale) but will not be able to estimate VCC under different conditions (*i.e.*, for a different clone in a different scale). This leads to the lengthy development of a plethora of small, local models. The extent of variation is visualized by using several PLS (B) models for one run (A-1). All Clone B models overestimate Clone A's VCC but the fit gets better as certain high frequencies are excluded (Figure 5). After the partial success with the first direct transfer of Clone A to B and B to A models, various PLS models (A) and (B) as well as processing techniques were tested on run A-1 (see supplement).

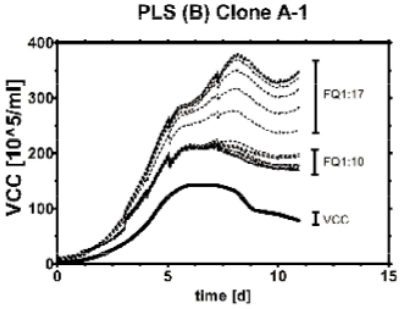


Figure 5. Testing frequency range effect on transferability. Model B is directly used on Clone A data with decreasing frequency input before PLS (B) model construction (1:17 all 17 frequencies, 1:10 frequency 1–10, *etc.*).

3.4.2. Attenuation Factor κ

After considering only selected frequencies, the fit suffers from less noise. Linear regression analysis of PLS estimation *vs*. measured VCC reveals that, *i.e.*, 1 unit estimated VCC corresponds to 0.69 units of real VCC with a very high R^2 value (see supplement).

When this statistically linear factor was considered and applied by multiplication on the same dataset, R^2 stayed the same, only the ratio of estimated *vs*. real VCC (the slope κ in Equation (4)) changed to one unit estimated VCC per one unit real VCC (Figure 6B,C). A PLS model constructed from a different cell line can be therefore readily re-used for any unknown cell line by knowledge of this simple linear and dimensionless factor termed Kappa (κ [-]), which is calculable online and offline due to its simple formula, where $\hat{y}$ is the estimated VCC, $\hat{x}$ the estimated VCC from a prior model, and d the intercept, all in (10^5/mL).

$$\hat{y} = \kappa \cdot \hat{x} + d \tag{4}$$

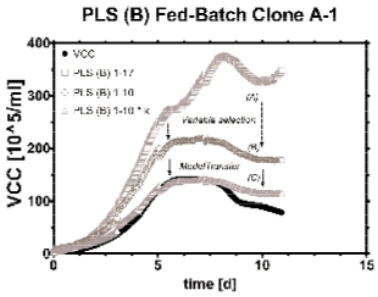


Figure 6. Transfer learning: (**A**) A PLS model (from Clone B) applied as it is for Clone A runs; (**B**) PLS (B) adapted by excluding the highest seven frequencies from the model; (**C**) PLS (B) modified with a linear factor. All runs were fed-batches with decreasing viability.

Because κ is calculable, it can be readily obtained autonomously already during the first bioreactor run of the new clone. This may be of importance for researchers or contract manufacturers with very tight timelines where no runs can be afforded for the establishments of "standard fed-batch profiles" and also where material is difficult to acquire or resources of any kind are a limiting factor. In our experience, the slope is available early in the process, when the culture is still in an exponential growth phase, but sometimes early process events such as pH regulation or aeration as well as drifts may lead to an inaccurate determination. Therefore, it is advised to start such an algorithm after a short lag time of a few hours (see supplement).

This methodology has been conducted also vice versa by using a PLS model from the Clone A training data and applying the same calculation routine to Clone B fed-batches. The results are available and described later on in this paper as well as in the supplements. A step-by-step transfer of model B to model A runs was investigated under the previously defined model acceptance criteria. The transferred Clone B model performed well with roughly 20% CVRMSE (Figure 7) considering that a linear regression calibration of the same clone and model (A) was determined with roughly double CVRMSE (~40%).

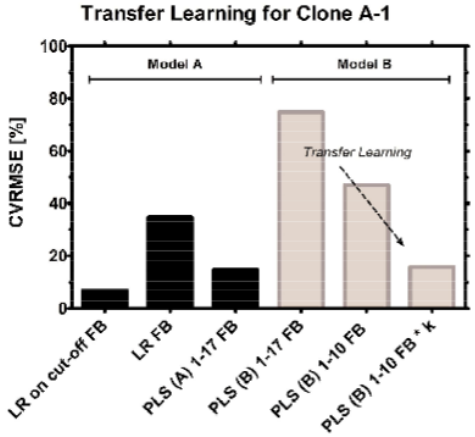


Figure 7. Model selection. From left to right: LR for highly viable fed-batch, LR for the whole FB, PLS regression for FB. Then, a Clone B model was employed to estimate Clone A: PLS model used as it is, PLS model with selected frequencies (1–10), PLS model with selected frequencies (1–10) and a linear factor (k).

If clone A was a new clone and used the first time in a bioreactor setting, naturally no runs would be available for constructing a PLS model for this cell line and standard linear regression would be the only available method for estimating VCC. PLS models from historical runs however may be useful for early biomass estimation where few or no data are available.

Exemplarily, the power of the model transfer is shown on two examples: A PLS (B) model is transferred so that Clone A-7 data is estimated, and the same was done with a PLS (A) model for Clone B-5 data, the errors are shown later in Figure 8.

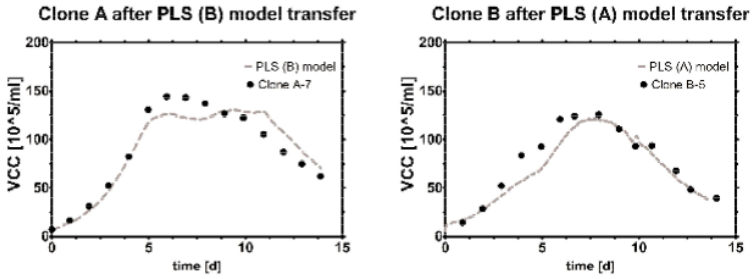


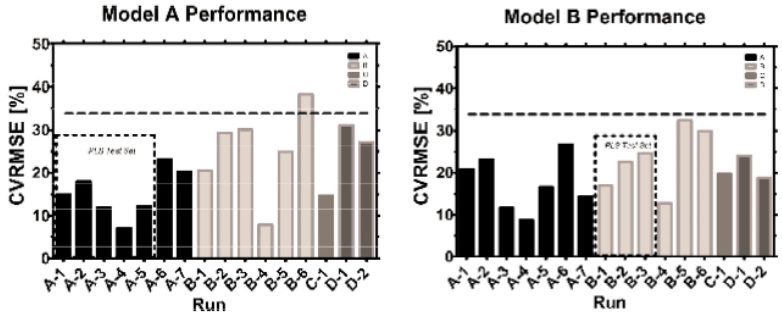
Figure 8. Model transfer. Estimation of VCC *versus* offline measurements with (**A**) PLS (B) model used on Clone A data; (**B**) PLS (A) model used on Clone B data. The death phase is captured sufficiently well with both models. The PLS (A) model seems to capture the decline phase slightly better in this particular run (see supplement for more runs), possibly because it was built with more data than the other model. Transferable, real-time resolved knowledge of VCC with this level of accuracy is believed to contribute significantly to the field of bioprocess monitoring and control.

3.4.3. Reasons for Transferability

The difference in the slope VCC *vs.* capacitance was observed by many research groups in the past [1,2,10,45]. The most common interpretations range from a change in size [5], a difference in internal or membrane conductivity [30] to the presence of mitochondria or other organelles [21]. Some researchers have observed that size alone cannot explain the difference in slope [8,12] of biological objects, or only after extensive offline de-correlation [35], which is not in the spirit of automatic process control. Therefore it is hypothesized that each clone exhibits a certain unique capacitance response to frequencies excitation in the right range.

3.5. Validation and Comparison with Literature

Model A was validated with Clone B, C and D. Model B was validated with Clone A, C and D. For Clone A and B, PLS models could be constructed, while Clones C and D did not yield enough data. Model B contained only three runs from mixed locations and scales from Clone B while model A contained five runs from scale 1, indicated by a dotted black square. All remaining runs (with varying scales and clones) were used as validation datasets. A comparison of both models shows that almost all runs were below 33% CVRMSE, indicating a satisfying capability to predict VCC in a fed-batch with declining viability (Figure 9).



(**A**) (**B**)

Figure 9. Model (**A**) *vs.* model (**B**) performance: The test sets for the PLS models are indicated, while all other runs served as model validation sets. A horizontal dashed line indicates 33% CVRMSE. Small differences in final CVRMSE calculation are possible when minute-by-minute records of capacitance spectra are aligned with the daily offline VCC measurements by various techniques [46–48].

3.5.1. Scale to Scale Transferability

Clone A was satisfyingly well predicted in either model with varying degrees of accuracy. In A7, a sudden burst of gas bubbles caused a slight drop in capacitances which was better predicted by model B, thus leading to a better final CVRMSE. Clone B runs were partially better predicted by model A (B4, B5), but one run (B6) also trespassed the acceptance criterion. However, in B6 the cells were treated with a detergent followed by a strong formation of cell debris which made it difficult to make use of frequency scanning in general. B5 experienced a strong morphological change in cell size, which caused the designated model to constantly overestimate larger cells. Clone C prediction was slightly improved when a model in the same scale was employed (model A). Model B, a mixed scale model, could be used to predict VCC sufficiently well. The difference in CVRMSE was rather small (5%) and the total CVRMSE did not exceed 20%, indicating sufficient accuracy in fed-batches. Clone D performed in general slightly better (5%–10% CVRMSE) with model B than model A. Both model predictions were sufficiently accurate for the previously defined acceptance criteria.

Conclusively, scale had a certain influence on model performance, but the influence was rather low in the given datasets of 16 runs.

3.5.2. Clone to Clone Transferability

Four clones were analyzed via capacitance profiling. The emitted frequencies evoked slightly different amounts of capacitance, leading to a physiological or morphological link between the measurement and individual clone. By knowing the proportion of capacitance per clone, linear factors can be considered in the model generation. However, such proportions may not only depend on the clone itself. Morphological differences might be associated with differences in the physico-chemical environment of platform media and processes. Clone B was subject to very orthogonal process

conditions leading to various physiological and morphological responses. The worst estimation of VCC thus happened in Clone B runs, where cell size increase correlated with overestimation of VCC (B5) and massive cell shrinkage correlated with underestimation (B6). Knowledge of the exact surface size of individual cells in real time together with a minimum of interfering particles might therefore lead to an improved model performance in any cell line.

In brief, either model predicted VCC well, no matter if the clone was A, B, C or D and it is possible that new CHO clones, and perhaps even other cell lines, follow this pattern.

3.5.3. Internal Model Comparison

Average CVRMSE of both models were compared and plotted together with their corresponding standard deviation (σ). Because of the large variance (roughly 20% ± 9% σ for average) in both datasets no significant difference between the methods was found which emphasizes their interchangeable character and universal applicability (Figure 10). However, the hybrid scale model B had slightly less variance which might indicate slightly more consistency in VCC estimation.

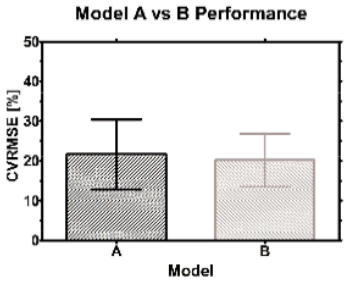


Figure 10. Model performance: Bars indicate the average CVRMSE found for model A and B, while error bars represent the standard deviation (SD) in the sample.

The hybrid model might prove to be more robust with more runs, because the influence of scale may be further integrated in the statistical model. On the other hand, data from only one scale but orthogonal process conditions might be equally suited to estimate VCC in other scales or clones. Since statistical models rely on numbers, time and more data from either approach will tell which strategy works best in the end.

3.5.4. External Model Comparison

A direct comparison between this contribution and the scientific work from other research groups is displayed below (Table 2). Due to the application of different performance criteria, post-processing steps, shorter observation windows (batch only) or other target variables than VCC (PCV) some data are not comparable. Although some authors provided the coefficient of determination (R^2), caution is advised when interpreting it outside of the specified range (which is mostly the highly viable phase in cultures). Somehow, numeric CVRMSE values were hard to come by for mechanistic models in particular, which becomes apparent because all PLS and linear models had them included.

Table 2. Literature model performance.

Author	Year	CVRMSE	Comments	Ref
Noll	1998	n.a.	Linear model, $R^2 = 0.99$, from a calibration curve with a serial dilution of a defined cell concentration	[8]
Cannizzaro	2003	9%–22% Batch phase, 24%–36% perfusion fed-batch	PLS model, 1 and 2 principal components, only one run available for validation (2 runs available), perfusion process with high viability	[12]
Ansorge	2007	n.a.	Linear model, 20% change in cell size corresponds to the third power (80% variance) in permittivity signal; $R^2 = 0.99$ provided for batch phase	[10]
Ansorge	2010	n.a.	No numeric performance parameters from the Cole-Cole model available. Linear model parameters: $R^2 = 0.74$–0.89, capacitance *vs.* packed cell volume (PCV), two different clones in a fed-batch, only samples with viability >70% taken into account	[1]
Opel	2010	7%–23% Mixed results, batch and fed-batch	PLS model, Result from cross validation with 5 principal components using 5 batches and 5 fed-batches as data source. Relative error is based on viable packed cell volume (vPCV)	[5]
Heinrich	2011	n.a.	No numeric performance parameters from the Cole-Cole model available. Linear model parameters: $R^2>0.98$ for highly viable cells in perfusion	[45]
Parta	2013	5%–45% without smoothing, 9%–15% with smoothing, all fed-batch	Three principal components, using always 1 out of 6 fed-batches for validation, with and without Savitzky-Golay smoothing to compensate extreme outliers	[39]
This contribution	2014	7%–38% model A 9%–32% model B all fed-batch	Three principal components, one model (A or B) predicts VCC for 4 different clones and two different scales in a total of 16 orthogonal fed-batches	[-]

4. Conclusions and Outlook

This contribution proposes a roadmap for creating and validating a universal PLS model for biomass estimation in mammalian fed-batch cell cultivations with capacitance spectra. Transferable models were established and tested on a total of 16 fed-batches, comprising four clones and two different scales under highly orthogonal process conditions typical in process development.

The herewith presented novel approach is able to estimate viable cell concentration in real time completely without further off-line analytics after exactly one biomass measurement at inoculation for offset correction. Viable cell concentration from new clones, which were not tested in a bioreactor setting yet, can be estimated with these statistical models by calculation of a linear attenuation factor κ, which can be established while the process is still running. Because no performance criteria were defined in literature yet, a new acceptance criterion is presented, which is the CVRMSE of the monitored fed-batch run. CVRMSE itself does not require further pre- or post-treatment of the continuously recorded data, but an improvement in accuracy is expected when measured VCC is splined in-between measurements and compared to the continuously estimated model VCC [46].

A minimum of one dataset is required for creation and cross-validation of a PLS model, however, validation on the same test set tends to be too optimistic. Results from the model on new datasets are

expected to improve when the PLS model is trained and validated with more than just one dataset, in this example at least five for model A and at least three for model B. A quickly growing database of fermentations with little process variation in production scale is expected to make one universal PLS model even more accurate and powerful, and will open up new avenues in bioprocess development.

Acknowledgments

We sincerely acknowledge Patrick Wechselberger for providing technical support in this work and Hans Lohninger from the Vienna University of Technology for providing us with a full version of Datalab [21] which we used to construct the PLS models. We also want to thank our project partners from Boehringer Ingelheim for sharing their data from different scales, cell lines and process formats without which this contribution in its current form would not have been possible.

Author Contributions

V.K. designed and performed a part of the experiments, analyzed the data, developed the universal models, performed statistical analysis, prepared tables, figures, additional files, drafted and wrote the manuscript. A.C.Y. helped in some experiments and in data analysis, created various PLS models, and summarized literature on PLS models and capacitance measurements in cell culture. C.C., M.B. and M.M. contributed the largest part of the data, helped with critical discussions and valuable support. C.H. coordinated and supervised research and helped drafting the manuscript. C.H. and S.S. conceived the study. All authors read and approved the final manuscript. Partial financial support from the CATMAT Doctorate College of the Vienna University of Technology is gratefully acknowledged.

Conflicts of Interest

The authors declare no conflict of interest.

Abbreviations

$a_1 - a_{17}$.....Capacitance 1–17 (pF/cm)
ANN.....Artificial Neural Networks
B.....Batch
FB.....Fed-Batch
$c_1 - c_{17}$……..Coefficient 1–17 [-]
CHO.....Chinese Hamster Ovary
CVRMSE.....Coefficient of Variation of RMSE (Root Mean Square Error)
d.....Offset (cells/mL)
fc.....Capacitance at the critical frequency (pF/cm)
FQ.....Frequencies
κ.....Linear factor [-]
LR.....Linear Regression
mab.....Monoclonal antibody
MC.....Mean Centering

MLR.....Multiple Linear Regression
n.....Number of measurements
PC.....Principal Component
PCR.....Principal Component Regression
PLS.....Partial Least Squares
PLS-R.....Partial Least Squares Regression
R^2.....Regression Coefficient [-]
RMSE.....Root Mean Square Error (cells/mL)
SD.....Standardization
TCC.....Total Cell Concentration (cells/mL)
VCC.....Viable Cell Concentration (cells/mL)
$\hat{x}$ .....Estimated VCC from a prior model (cells/mL)
y.....Measured VCC (cells/mL)
$\overline{y}$ …….Average VCC (cells/mL)
$\hat{y}$ .....Estimated VCC (cells/mL)

References and Notes

1. Ansorge, S.; Esteban, G.; Schmid, G. On-line monitoring of responses to nutrient feed additions by multi-frequency permittivity measurements in fed-batch cultivations of CHO cells. *Cytotechnology* **2010**, *62*, 121–132.
2. Yardley, J.E.; Todd, R.; Nicholson, D.J.; Barrett, J.; Kell, D.B.; Davey, C.L. Correction of the influence of baseline artefacts and electrode polarization on dielectric spectra. *Bioelectrochemistry,* **2000**, 51, 53–65.
3. Dabros, M.; Dennewald, D.; Currie, D.J.; Lee, M.H.; Todd, R.W.; Marison, I.W.; Stockar, U. Cole-Cole, linear and multivariate modeling of capacitance data for on-line monitoring of biomass. *Bioprocess Biosyst. Eng.* **2008**, *32*, 161–173.
4. Harris, C.M.; Todd, R.W.; Bungard, S.J.; Lovitt, R.W.; Morris, J.G.; Kell, D.B. Dielectric permittivity of microbial suspensions at radio frequencies: A novel method for the real-time estimation of microbial biomass. *Enzyme Microb. Technol.* **1987**, *9*, 181–186.
5. Opel, C.F.; Li, J.; Amanullah, A. Quantitative modeling of viable cell density, cell size, intracellular conductivity, and membrane capacitance in batch and fed-batch CHO processes using dielectric spectroscopy. *Biotechnol. Prog.* **2010**, *26*, 1187–1199.
6. Favre, E.; Voumard, P.; von Stockar, U.; Péringer, P. A capacitance probe to characterize gas bubbles in stirred tank reactors. *Chem. Eng. J.* **1993**, *52*, 1–7.
7. Yardley, J.E.; Kell, D.B.; Barrett, J.; Davey, C.L. On-line, real-time measurements of cellular biomass using dielectric spectroscopy. *Biotechnol. Genet. Eng. Rev.* **2000**, *17*, 3–35.
8. Noll, T.; Biselli, M. Dielectric spectroscopy in the cultivation of suspended and immobilized hybridoma cells. *J. Biotechnol.* **1998**, *63*, 187–198.
9. Zeiser, A.; Bédard, C.; Voyer, R.; Jardin, B.; Tom, R.; Kamen, A.A. On-line monitoring of the progress of infection in .Sf-9 insect cell cultures using relative permittivity measurements. *Biotechnol. Bioeng.* **1999**, *63*, 122–126.

10. Ansorge, S.; Esteban, G.; Schmid, G. On-line monitoring of infected Sf-9 insect cell cultures by scanning permittivity measurements and comparison with off-line biovolume measurements. *Cytotechnology* **2007**, *55*, 115–124.
11. Ansorge, S.; Lanthier, S.; Transfiguracion, J.; Henry, O.; Kamen, A. Monitoring lentiviral vector production kinetics using online permittivity measurements. *Biochem. Eng. J.* **2011**, *54*, 16–25.
12. Cannizzaro, C.; Gügerli, R.; Marison, I.; von Stockar, U. On-line biomass monitoring of CHO perfusion culture with scanning dielectric spectroscopy. *Biotechnol. Bioeng.* **2003**, *84*, 597–610.
13. Niklas, J.; Heinzle, E. Metabolic flux analysis in systems biology of mammalian cells. In *Genomics and Systems Biology of Mammalian Cell Culture*; Hu, W.S., Zeng, A.-P., Eds.; Springer: Berlin, Germany, 2012; pp. 109–132.
14. Chen, Z.; Chen, Y.; Chen, J.; Shen, C. Effects of ammonium and lactate on hybridoma cell growth and metabolism. *Chin. J. Biotechnol.* **1992**, *8*, 255–261.
15. Cruz, H.J.; Moreira, J.L.; Carrondo, M.J. Metabolic shifts by nutrient manipulation in continuous cultures of BHK cells. *Biotechnol. Bioeng.* **1999**, *66*, 104–113.
16. Templeton, N.; Dean, J.; Reddy, P.; Young, J.D. Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture. *Biotechnol. Bioeng.* **2013**, *110*, 2013–2024.
17. Zeng, A.-P.; Hu, W.-S.; Deckwer, W.-D. Variation of stoichiometric ratios and their correlation for monitoring and control of animal cell cultures. *Biotechnol. Prog.* **1998**, *14*, 434–441.
18. Ivorra, A. Bioimpedance monitoring for physicians: An overview. *Cent. Nac. Microelectròn. Biomed. Appl. Gr.* **2002**, *1*, 1–35.
19. Garthwaite, P.H. An Interpretation of partial least squares. *J. Am. Stat. Assoc.* **1994**, *89*, 122–127.
20. Wise, B.M. *Properties of Partial Least Squares (PLS) Regression, and Differences between Algorithms*; Technical Report; Eigenvector Research, Inc.: Manson, WA, USA, 2015.
21. Lohninger, H.H. Datalab 3.5, A Programme for Statistical Analysis. 2000. Available online: http://datalab.epina.at/ (accessed on 28 August 2015).
22. Rathore, A.S.; Mhatre, R. *Quality by Design for Biopharmaceuticals: Principles and Case Studies*; Wiley-Interscience: Hoboken, NJ, USA, 2011.
23. Rosipal, R.; Krämer, N. Overview and recent advances in partial least squares. In *Subspace, Latent Structure and Feature Selection*; Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J., Eds.; Springer: Berlin, Germany, 2006; pp. 34–51.
24. Haenlein, M.; Kaplan, A.M. A beginner's guide to partial least squares analysis. *Underst. Stat.* **2004**, *3*, 283–297.
25. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17.
26. Tobias, R.D. An introduction to partial least squares regression. In Proceedings of the 20th Annual SAS Users Group International Conference, Orlando, FL, USA, 2–5 April 1995.
27. Mehmood, T.; Liland, K.H.; Snipen, L.; Sæbø, S. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62–69.
28. Pinto, R.C.V.; Marinho, P.A.N.; Oliveira, A.B.; Esteban, G.; Melo, P.A.; Medronho, R.A.; Castilho, L.R. Biomass monitoring and cho cell culture optimization using capacitance spectroscopy. In *Cells and Culture*; Noll, T., Ed.; Springer: Dordrecht, The Netherlands, 2010; pp. 343–348.

29. Wong, J.; Implementation of Capacitance Probes for Continuous Viable Cell Density Measurements for 2K Manufacturing Fed-Batch Processes at Biogen Idec. Available online: http://www.infoscience.com/JPAC/ManScDB/JPACDBEntries/1394130144.pdf (accessed on 28 August 2015)
30. Justice, C.; Brix, A.; Freimark, D.; Kraume, M.; Pfromm, P.; Eichenmueller, B.; Czermak, P. Process control in cell culture technology using dielectric spectroscopy. *Biotechnol. Adv.* **2011**, *29*, 391–401.
31. Beving, H.; Eriksson, L.E.; Davey, C.L.; Kell, D.B. Dielectric properties of human blood and erythrocytes at radio frequencies (0.2–10 MHz); dependence on cell volume fraction and medium composition. *Eur. Biophys. J.* **1994**, *23*, 207–215.
32. Gerckel, I.; Garcia, A.; Degouys, V.; Dubois, D.; Fabry, L.; Miller, A.O.A. Dielectric spectroscopy of mammalian cells. *Cytotechnology* **1993**, *13*, 185–193.
33. Kell, D.B.; Woodward, A.M.; Davies, E.A.; Todd, R.W.; Evans, M.F.; Rowland, J.J. Nonlinear dielectric spectroscopy of biological systems: Principles ans applications. In *Nonlinear Dielectric Phenomena in Complex Liquids*; Rzoska, S.J., Zhelezny, V.P., Eds.; Springer: Dordrecht, The Netherlands, 2005; pp. 335–344.
34. Markx, G.H.; Kell, D.B. Use of dielectric permittivity for the control of the biomass level during biotransformations of toxic substrates in continuous culture. *Biotechnol. Prog.* **1995**, *11*, 64–70.
35. Ansorge, S.; Esteban, G.; Schmid, G. Multifrequency permittivity measurements enable on-line monitoring of changes in intracellular conductivity due to nutrient limitations during batch cultivations of CHO cells. *Biotechnol. Prog.* **2010**, *26*, 272–283.
36. David, J.; Nicholson, D.B.K. Deconvolution of the dielectric spectra of microbial cell suspensions using multivariate calibration and artificial neural networks. *Bioelectrochem. Bioenerg.* **1996**, *39*, 185–193.
37. Lohninger, H. *Teach/Me—Data Analysis*, 1st ed.; Springer: Berlin, Germany, 1999.
38. Beebe, K.R.; Pell, R.J.; Seasholtz, M.B. *Chemometrics: A Practical Guide*; Wiley: Hoboken, NJ, USA, 1998.
39. Párta, L.; Zalai, D.; Borbély, S.; Putics, A. Application of dielectric spectroscopy for monitoring high cell density in monoclonal antibody producing CHO cell cultivations. *Bioprocess Biosyst. Eng.* **2014**, *37*, 311–323.
40. El Wajgali, A.; Esteban, G.; Fournier, F.; Pinton, H.; Marc, A. Impact of microcarrier coverage on using permittivity for on-line monitoring high adherent Vero cell densities in perfusion bioreactors. *Biochem. Eng. J.* **2013**, *70*, 173–179.
41. Zeiser, A.; Elias, C.B.; Voyer, R.; Jardin, B.; Kamen, A.A. On-line monitoring of physiological parameters of insect cell cultures during the growth and infection process. *Biotechnol. Prog.* **2000**, *16*, 803–808.
42. Davey, C.L.; Markx, G.H.; Kell, D.B. Substitution and spreadsheet methods for analysing dielectric spectra of biological systems. *Eur. Biophys. J.* **1990**, *18*, 255–265.
43. Natschläger, T.; Zauner, B. Fused Stage-Wise Lasso—A Waveband Selection Algorithm for Spectroscopy. Available online: http://www.scch.at/de/publikationen/publication_id/802 (accessed on 19 August 2014).

44. Torrey, L.; Shavlik, J. Transfer Learning. Available online: ftp://ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf (accessed on 28 August 2015).
45. Heinrich, C.; Beckmann, T.; Büntemeyer, H.; Noll, T. Utilization of multifrequency permittivity measurements in addition to biomass monitoring. *BMC Proc.* **2011**, *5*, doi:10.1186/1753-6561-5-S8-O10.
46. MATLAB, Inc. Matlab 1-D Data Interpolation with Interp1. Available online: http://de.mathworks.com/help/matlab/ref/interp1.html (accessed on 28 August 2015).
47. Motulsky, H. *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, 1st ed.; Oxford University Press: Oxford, UK, 2004.
48. Motulsky, H. *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*, 3rd ed.; Oxford University Press: New York, NY, USA, 2013.

Note: This contribution is provided together with sample data for frequencies and VCC, which were modified with regard to viable cell concentration by a linear factor between 0.5 and 2.0. The reader is invited to use the supplied excel sheet to estimate VCC with their own data or review historical bioreactor runs with one of our models.

2.2 Chapter 2

bioengineering

MDPI

Article

Metabolic Control in Mammalian Fed-Batch Cell Cultures for Reduced Lactic Acid Accumulation and Improved Process Robustness

Viktor Konakovsky [1], Christoph Clemens [2], Markus Michael Müller [2], Jan Bechmann [2], Martina Berger [2], Stefan Schlatter [2] and Christoph Herwig [1,*]

Received: 18 July 2015; Accepted: 4 January 2016; Published: 11 January 2016
Academic Editors: Mark Blenner and Michael D. Lynch

[1] Institute of Chemical Engineering, Division of Biochemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A 166-4, 1060 Vienna, Austria; vkonakovtuwien@gmail.com
[2] Boehringer Ingelheim Pharma GmbH & Co. KG Dep. Bioprocess Development, Biberach, Germany; christoph.clemens@boehringer-ingelheim.com (C.C.); markus_michael.mueller@boehringer-ingelheim.com (M.M.M.); jan.bechmann@boehringer-ingelheim.com (J.B.); martina.berger@boehringer-ingelheim.com (M.B.); stefan.schlatter@boehringer-ingelheim.com (S.S.)
* Correspondence: christoph.herwig@tuwien.ac.at; Tel.: +43-58801-166-400; Fax: +43-58801-166-980

Abstract: Biomass and cell-specific metabolic rates usually change dynamically over time, making the "feed according to need" strategy difficult to realize in a commercial fed-batch process. We here demonstrate a novel feeding strategy which is designed to hold a particular metabolic state in a fed-batch process by adaptive feeding in real time. The feed rate is calculated with a transferable biomass model based on capacitance, which changes the nutrient flow stoichiometrically in real time. A limited glucose environment was used to confine the cell in a particular metabolic state. In order to cope with uncertainty, two strategies were tested to change the adaptive feed rate and prevent starvation while in limitation: (i) inline pH and online glucose concentration measurement or (ii) inline pH alone, which was shown to be sufficient for the problem statement. In this contribution, we achieved *metabolic control* within a defined target range. The direct benefit was two-fold: the lactic acid profile was improved and pH could be kept stable. Multivariate Data Analysis (MVDA) has shown that pH influenced lactic acid production or consumption in historical data sets. We demonstrate that a low pH (around 6.8) is not required for our strategy, as glucose availability is already limiting the flux. On the contrary, we boosted glycolytic flux in glucose limitation by setting the pH to 7.4. This new approach led to a yield of lactic acid/glucose (Y L/G) around zero for the whole process time and high titers in our labs. We hypothesize that a higher carbon flux, resulting from a higher pH, may lead to more cells which produce more product. The relevance of this work aims at feeding mammalian cell cultures safely in limitation with a desired metabolic flux range. This resulted in extremely stable, low glucose levels, very robust pH profiles without acid/base interventions and a metabolic state in which lactic acid was consumed instead of being produced from day 1. With this contribution, we wish to extend the basic repertoire of available process control strategies, which will open up new avenues in automation technology and radically improve process robustness in both process development and manufacturing.

Keywords: CHO cell culture; scale-down; fed batch; automation; Lactic acid control; pH; metabolic control; MVDA; uncertainty; online analyzer

1. Introduction

1.1. Problem Statement

A great amount of modern biopharmaceuticals such as monoclonal Antibodies (mAbs), fusion proteins, bi-specific antibodies, IgG and others are produced today by highly specialized cells in a fed-batch bioprocess. The Chinese hamster ovary (CHO), well known by cell banking institutes and authorities, is often engineered (*i.e.*, GS-CHO) to satisfy a tailored production purpose. Often a large-scale fed-batch or perfusion operation is required to meet the market demand. In such processes, the feeding regime was found to have a great impact on lactate production, which negatively affects the productivity [1]. As a consequence, variations in both quantity and quality during recombinant protein processes affect the profitability of the production plant markedly.

The lactic acid profile of a mammalian cell culture process affects the maximum achievable cell count and final product titer [2,3]. Manufacturing runs with a high and a low lactic acid profile have been linked with productivity of the process by using Multivariate Data Analysis (MVDA) methods [4,5]. It does not come as a big surprise that there is not only one but there are several approaches to decrease lactic acid. Some methods encompass genetic modifications of the clone [6–8], changes in medium composition [9–12], interventions on a genetic level [13,14], feeding strategy [15–18], modification of metabolic state [19–23] or the physico-chemical environment [24–30], to name a few.

As lactic acid is produced, pH falls—which can again impair cell proliferation [31] at critically low pH. Depending on the clone and cell type, the lower critical pH limit is typically between 6.6–6.8 in mammalian cell culture [24]. In several other studies, pH was also associated with interfering with product quality [32]; therefore, the way pH is controlled is still an interesting question today. Ideally, pH should be fixed during the whole process and should never need adjustment by either acidic or alkaline control agents. In practice this is not exactly the case. An increase of osmolality and stress to the cells is often the result of a suboptimal pH control strategy. Operation at low pH is observed to restrict the formation of lactic acid at the cost of inhibiting cellular proliferation while an alkaline pH has the opposite effect [24,25,33]. Therefore, many industrial processes start with high pH and shift to low pH during the culture as will be seen a bit later in this contribution. However, modifying the pH alone is not fast enough to prevent lactic acid build-up, as can be seen in the lactic acid profiles in our data—the metabolic state itself must be modified, and this can be done with an appropriate feeding strategy.

1.2. Feeding Strategies

The choice of an appropriate feeding strategy depends very much on the context; they may be based on well-defined industrial production runs, where little variation is anticipated and a historical run can be used as template for the next run. Alternatively, they are derived by in-line, at-line or on-line signals from the current process in real time, which is often the case in a bioprocess development environment [34–36]. Therefore, a methodology to deliver feed adaptively, accurately and safely in mammalian cell culture development is of very high relevance [37–40]. In order to control lactate production by adaptive feeding, the main understanding is that a limited metabolic state, mainly a carbon limitation, needs to be designed. The methodology is typically divided into data-driven and adaptive feeding approaches.

1.2.1. Data-Driven Feeding

Data driven approaches rely on knowledge gathered from historical data. Drapeau [16], Luan [15,41] and Gagnon *et al.* [42] published methods to control mammalian cell culture in glucose limitation (around 1 mM). As criteria for an increase or decrease of a previously determined feed rate from very similar historical runs, if too much lactic acid is produced, the feed is too high and needs to decrease, whereas when the pH is rising, the feed is too low and needs to increase. This intuitive

and simple method may be limited if buffers are employed, which mask lactic acid production and may lead to periods of starvation [43]. Aehle *et al.* [44] described the control of a mammalian fed-batch cultivation by limiting glutamine instead of glucose availability, thus affecting carbon source utilization and reducing ammonia yields, to control the feed rate. Previous experiments under similar conditions or numerical optimization studies were used to establish a tcOUR (total cumulative oxygen uptake rate) profile. This profile was correlated *ab initio* to the historical viable biomass profile. Future repetitions of the experiment with different initial viable biomass concentrations led to a highly reproducible cell concentration estimation and also to the desired sub-maximal specific growth rate. In order to deliver the correct amount of feed in real time to the variable biomass, the glutamine per biomass yield (Y X/GLN) had to be known and proved to be fairly constant during the six days of cultivation.

Liu *et al.* [45] even suggested the feed-forward control of mammalian cell cultures. They derived a historical growth rate by linear calibration of previous growth curves and used it together with the yield of biomass per glucose (Y X/GLC) to set up a determined feed rate. Validation with three different cell lines led to an improvement of titer as well as IVCC (integrated viable cell concentration). This leads to the question of why data-driven strategies work so extremely well. One reason is that the validation experiments are short so that the verification does not face the difficulty of describing a considerable death phase, which leads to a reduced representation of a typical industrial process. Additionally, different clones in different scales often have different biomass and metabolite profiles. Differences in medium composition significantly diminish the usefulness of pre-determined feed profiles. Typically, even large datasets of one clone in scale 1 cannot predict the behavior of another clone in scale 1 or the same clone in scale 2. Therefore, a sufficient number of experiments are typically conducted *a priori*, which can build up enough confidence that a data-driven approach does not fail, which would be especially devastating in a manufacturing-scale run.

1.2.2. Adaptive Feeding Strategies

Usual orthogonal process development runs are too different to successfully apply a data-driven pre-calculated feed or biomass profile from historical data. For these reasons we need to focus on simple mechanistic relationships which allow automatic feed rate adaptation to the current and not a past process. A look into literature shows that there are several ways to accomplish this by deriving a feed in real time using in-line, at-line and on-line signals.

Zhou and Hu first mention signals which may be used to detect the metabolic state of mammalian cell cultures and derive a feeding profile on-line [46]. The turbidity probe allowed the conversion of optical density (OD) to cell concentration with a simple calibration curve as well as the oxygen uptake rate (OUR) which made it possible to detect a more direct measure for cellular activity. Noll and Biselli [47] used a capacitance probe to deliver feed to a fluidized bed culture. The yield X/GLN was constant during the whole time which allowed keeping glutamine at 0.45 mM with an online massflow dosing system. Dowd *et al.* [48] used capacitance as well to estimate viable cell concentrations in a perfusion process to control the nutrient feed rate. The feeding was adjusted automatically and was used to maximize the productivity. Zhou *et al.* [49] used the OUR measurement to adjust the nutrient feeding rate in a fed-batch culture. This was possible by applying a previously determined yield of glucose per oxygen (GLC/OX) consumption and led to the control of the process at very low glucose and amino acid levels. The strategy was applicable as long as a reliable calculation of the kLa (volumetric oxygen transfer coefficient) allowed determination of OUR. Europa *et al.* [50] used such a setup to show how controlling substrates at low levels not only led to reduced metabolite formation but also to a multiplicity of steady states, in which a more efficient metabolism was observed. Ozturk *et al.* [38] utilized an at-line analyzer to manipulate the perfusion rate of a culture in real time. Li *et al.* [51] proposed several other techniques to control mammalian cell culture, which are based on the cell's specific oxygen uptake rate (qOUR) during the process. In combination with substrate or metabolite measurements, qOUR was used as an indicator to change other process parameters on-line. Among those, feed rate, pH, temperature, stirring speed, and pO_2 were adapted when limitations, such

as reduced qOUR or low substrate, could be detected. Lu [37] compared two methods to control the feed rate: in one instance, an auto sampler method was used to hold a particular glucose concentration, and in the other, a capacitance signal was used as a surrogate for cell growth to control the feed rate. In both cases, the target glucose concentration was non-limited between 4–6 g/L to prevent over- and under-feeding. Variable, specific consumption rates were observed over time, which are typical for operations at such high concentration ranges of substrate. As an online or offline measurement was available, it could be used to correct these deviations. Alternatively, previous experiments were shown to be useful to deduct the evolution of the specific consumption rates, with the drawback that this again makes the process dependent on historical data.

1.3. Challenges in Process Control

The aforementioned contributions describe the state of the art; however, there are still many gaps in knowledge which need to first be mentioned and then attempted to be closed. Many methods are based on some prior knowledge of the process. However, in reality there are many sources for deviations which complicate things considerably. Some sources of complication with which we were confronted in our labs and want to report are listed in Table 1.

Table 1. Selection of observed sources of complications in the development of feeding strategies in our labs.

Source	Influence
Clones	Metabolic needs may differ greatly, leading to the perpetual development of historical feeding profiles. In adaptive feeding regimes, clone-dependent differences of dielectric properties may complicate biomass estimation when capacitance probes are used, while turbidity probes may detect more or less cell debris in the decline phase, depending on which clone was used.
Scales	Especially on-line offgas/kLa–dependent control strategies may become very difficult to transfer because they depend on the aeration and stirrer cascade strategy (*i.e.*, constant or adaptively increasing gas flow to hold pO_2).
Assumptions	Constant yields (*i.e.*, GLC/OX, X/GLC, X/GLN, *etc.*) may change over time, leading to stoichiometric over- or under-feeding.
Media	Addition of growth-influencing components may change historical feeding profiles completely and make a direct comparison between experiments difficult as these changes have further implications on the process.
Process parameters	Changes in temperature, stirrer speed, pO_2, pH or pCO_2 levels may affect gas solubility, buffer capacity, offgas profiles, cellular stress level, and growth and may change the metabolic requirements for both adaptively or historically calculated feed rates.

We do not claim to offer solutions to all mentioned points but want to make the reader aware of their existence and propose solutions for some of them in our own work: the herewith presented adaptive feeding strategy was based on estimation of biomass in the current process, making it independent of historical data. The method we used in our labs to estimate biomass, and with it the adaptive feeding rate, has been shown in a previous contribution to be transferrable between clones and scales [52].

We coped with changing metabolic yields by limiting glucose concentrations and the glucose uptake rate to a desired range. The yield was still subject to variation, but could not exceed a particular range—the rate at which substrate was supplied corresponded to the desired consumption rate of the cells in real time, which made it controllable. The method is based on a first principle investigation using correlations between the specific rates such as glucose or lactic acid to determine the feed rate set-point. The metabolic response may differ from clone to clone in the slope of the yield Y L/G, but once established, this correlation will hold for one clone which makes the general

methodology applicable to different clones. Uncertainty consisting of both metabolic state and biomass was considered by applying a range for the feed set-point; even though the yields and biomass error are subject to change, certain signals (*i.e.*, pH, among others) could be used to switch the feeding rate automatically to ensure sufficient supply of substrate for the whole process time.

1.4. Goal

The goal of this study was two-fold: to reduce lactic acid levels and to minimize the pH control actions by additional acid-base feeding using a dynamic process environment as encountered in a process of industrial relevance. This contribution therefore proposes a novel adaptive feeding strategy, which is based on estimating the viable cell concentration with a capacitance probe in glucose-limited growth conditions, targeting a low lactic acid/glucose yield (Y L/G) and taking pH variation into account in the control strategy. The workflow followed in this contribution was (i) to analyze historical process development data to get an understanding in which range specific glucose consumption shows a favorable Y L/G profile; (ii) MVDA to determine the contributions of important parameters influencing lactic acid production; and (iii) development of the feeding strategy based on a high and low uncertainty of the metabolic state and biomass estimation error. The results show that lactic acid build-up can be decreased by confining glucose flux to a particularly low range, regardless of the pH set-point. However, pH may have helped to keep glucose consumption higher than reference processes late in the process, which in turn may have had a positive effect on productivity. The challenge was finding a way to address this in a dynamic fed-batch process; therefore, our methodology aims at being directly transferrable from process development back to manufacturing conditions.

2. Experimental Section

2.1. Cell Lines

All data in this contribution was collected from one single engineered CHO clone (derived from CHO DG44), subsequently referred to as "Clone B". This clone was kindly provided by Boehringer Ingelheim (Ingelheim, Germany) for the necessary experiments at the VUT (Vienna University of Technology). Only Clone B data was used for data analysis purposes as the metabolic behavior is different from Clone A data. However, for completeness, the data to construct and validate the multivariate model for biomass estimation (see chapter MVDA) were recorded with "Clone A" clones. The biomass model was shown to be transferrable and also estimate Clone B data. Clone A data is not discussed further in this contribution, but more information could be found in our recent publication [52].

2.2. Available Dataset

A historical Clone B dataset consisting of 29 fed-batch fermentations from process development (2L and 80L) was kindly provided by Boehringer Ingelheim (Ingelheim, Germany) to generate process understanding using MVDA techniques. These data sets are also referred to as "Historical data". On this basis, fed-batch fermentations were mirrored in a scale-down model (2L), and two experiments, "R-30" and "R-31", were selected and described in detail in this contribution.

2.3. Media

Media for the seed train and fed-batches are proprietary in composition and subject to variations in starting levels of metabolites, growth factors, selection pressure, *etc.* All components were serum-free, animal-component-free, and chemically defined.

2.4. Process Setup

Clone B cells were cultivated in shake-flasks (Corning Inc., Corning, NY, USA) in incubators (Minitron, Infors, Bottmingen, Switzerland) with 5% partial CO_2 pressure at physiological temperature

(35–37 °C) on orbital shakers at 120 rpm (orbit 50 mm). Passaging was performed every third day in proprietary chemically defined media and the bioreactors (3.6L, Infors HT, Bottmingen, Switzerland) were inoculated at an initial seed density of 3×10^4 to more than 1×10^6 cells/mL. Process information was logged using the process management system Lucullus (PIMS, Lucullus, Biospectra, Switzerland). A capacitance probe (Biomass Sensor, Hamilton Bonaduz AG, Bonaduz, GR, Switzerland) in scanning mode was used with an excitation frequency ranging from 0.3 MHz to 10 MHz, the signal was collected in real time and recorded between one and more than 60 min steps. Offline concentration of total cells, viable cells and viability was measured using an image-based white/dark classification algorithm after automatic trypan blue staining integrated in the cell counter (Cedex HiRes, Roche, Basel, Switzerland). Main metabolite concentrations (Glucose, Lactic acid and IgG) were measured on-line and off-line using a photometric robot (CubianXC, Optocell, Bielefeld, Germany).

2.5. Online Enzymatic Analyzer

An enzymatic analyzer (CubianXC, Optocell, Germany) was coupled to the process by withdrawing sample via a ceramic membrane made of Al_2O_3 (pore size 0.2 µm, membrane area 17.8 cm^2, membrane thickness 1.6 mm; IBA, Heiligenstadt, Germany). The supernatant was removed from the reactor at certain intervals at a flow rate between 0.2–1.0 mL/min. The sampling intervals of the analyzer were tested between 3–12 h in experiment R-30 and a designated sampling interval of 6 h was set for the experiment R-31 to measure, among others, glucose and lactic acid concentrations. Purging the lines and sampling took a total of 30 min where a maximum of 12 mL cell-free supernatant was withdrawn per sample. Faster purging and more frequent sampling (up to 30 min intervals) was technically possible, but not required for our purpose. A more detailed description of the enzymatic robot's set-up and substrate and metabolite measurements is described elsewhere [53].

2.6. MVDA

2.6.1. Biomass Model

In brief, multivariate models perform better at capturing the declining phase of a culture than linear models. One of these was available to estimate biomass in historical Clone A runs. The resulting model was transferrable to another clone (here: Clone B) by adaptation of the slope of the multivariate model for this new clone. For more information please refer to our publication [52]. The model was constructed prior to the experiment with Datalab 3.5 software (kindly provided by Prof. Lohninger, Vienna University of Technology, Vienna, Austria) [54] and the transferred Clone A model was used to estimate Clone B biomass in real time.

2.6.2. Data Mining

Historical Clone B data from process development was harmonized by using the same calculation routine in Matlab (Mathworks, Natick, MA, USA) for all runs, including selected experiments R-30 and R-31 at the institute, to exclude possible variations after data treatment. In order to gather process understanding of how and by which importance the process parameters relate to each other, we used PLS-R (partial least squares regression) in Simca 13.0.3 (Umetrics, Sweden) to analyze the historical dataset. These parameters were chosen with considerations explained a little later in the results section and encompassed the specific rates and concentrations of glucose and lactic acid as well as pH, which was used in its non-logarithmic representation of H^+ concentration.

2.7. Process Control

2.7.1. Process Control Scheme

The recorded spectra were recorded every minute and directly translated into a real-time feed rate. The following script (Figure 1) controlled the process in the experiments. First, biomass was

estimated using the coefficients calculated from a multivariate model. We used the assumption that glucose is limited and solely the feed rate is determining the volumetric glucose uptake rate. Both entities are combined to a specific glucose uptake rate, which is in turn used as a set-point to calculate a real-time feed rate. The calculation of specific substrate consumption qs was done by dividing the volumetric rate by the viable cell concentration. As derivative rates tend to be noisy, all specific rates were treated using Matlab's *rlowess* function [55]. The set-point of the feeding rate was corrected in real time by either of two factors: by an online enzymatic analyzer result (such as threshold glucose concentration), or by pH (for instance a pH upper or lower band), or a combination of both.

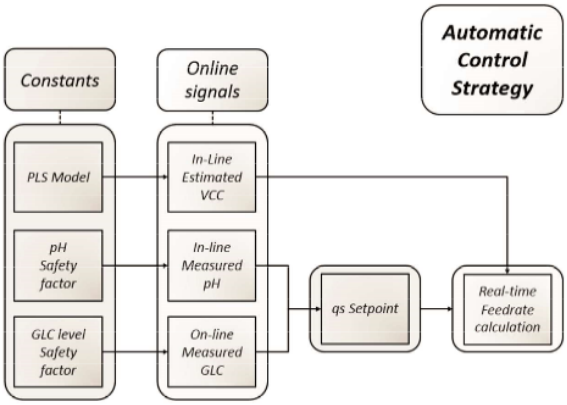


Figure 1. Automatic control strategy using exclusively real-time available signals. Constants (safety factors and the biomass model) were defined *a priori* which led to either a rather loose or very strict adherence to a desired qs set-point.

2.7.2. Feed Rate Calculation

We estimate biomass and feed stoichiometrically for the specific consumption rate q of a substrate s which the cells should ideally hold. A material balance for substrate glucose was used [56] to calculate qs and historical data analysis was performed to find a useful qs set-point which should be held over time. As qs is not a static value and subject to change, the question of which qs should be set was a very important one. We address it and the implications a bit later in the results section of our work. The feed set-point was controlled gravimetrically by a pump (Lambda Preciflow, Czech Republic) which received the adaptive set-point from a process management system Lucullus (PIMS, Lucullus, Biospectra, Switzerland) in real time. The adaptive feed rate (see Equation (1)) increased and decreased with biomass, which was estimated using a PLS model. The rate at which cells were fed stoichiometrically depended on the experiment and could be influenced by certain process events (Equation (2) and Table 2) to prevent starvation of the cells without having to rely on operator interventions.

The real-time automatic feed rate is described by Equation (1), where qs (which may be glucose or any other substrate) is kept constant.

$$Feed\left[\frac{ml}{h}\right] = \frac{VCC\left[\frac{c}{ml}\right] \cdot V_{Reactor}\,[mL] \cdot q_s\left[\frac{mg}{c \cdot h}\right]}{Feed\ concentration_{Substrate}\left[\frac{mg}{ml}\right]} \quad (1)$$

We used a qGlc *range* instead of a single, fixed set-point, which depends on a previously defined safety factor (ranging here from −25% to +100%). As a switch between qs set-points we used inline pH and online available signals of substrates or metabolites measured by the online analyzer (here glucose); therefore, the qs we specified to be desirable to reduce Y L/G is multiplied by a factor which corresponds to the expected variation in the process. For both experiments, the factor (Equation (2)) by which qs changed was between 0.75 and 2.0 (1.0 minus safety factor of 25% and 1.0 plus safety factor of 100%) and if 1.0 is replaced by the actual feeding set-point, qs reads 8 (rounded up) to 30, depending on the experiment. The safety factor specification by either pH or online analyzer was distinct in both experiments and is explained in detail in Table 2.

$$qs_{\,adapted} = qs - \ safety\ factor\ f\,(pH,\ online\ analyzer)\ \cdot qs \tag{2}$$

The final equation (Equation (3)) as it was used in the control strategy now reads:

$$Feed\left[\frac{ml}{h}\right] = \frac{VCC\left[\frac{c}{ml}\right] \cdot V_{Reactor}\,[mL]\ \cdot q_{s\ adapted}\ \left[\frac{mg}{c \cdot h}\right]}{Feed\ concentration_{S,in}\ \left[\frac{mg}{ml}\right]} \tag{3}$$

Table 2. Control specifications: Initial qs set-point (SP) in [pg/ch] (picogram per cell per hour), correction by a defined safety factor in [%] under certain conditions (pH and Online Analyzer), and Target SP.

Control Specifications				
	Switch Conditions		*Target SP [pg/ch]*	
Experiment	***R-30***	***R-31***	***R-30***	***R-31***
Initialization	Start feed after 2 h	Start feed after 2 h	−15	−10
		pH control †		
pH high	If pH ⩾ 7.1. increase qs by 100%	If pH ⩾ 7.4. increase qs by 25%	−30	−13
pH OK	If pH between 7.1 and 6.9 use the desired qs	If pH between 7.1 and 6.9 use the desired qs	−15	−10
pH low	If pH ⩽ 6.9 reduce qs by 25%	If pH ⩽ 6.9 reduce qs by 25%	−11	−8
		Online Analyzer control ‡		
GLC low	If Gluc ⩽ 0.4 increase qs by 100%	Monitoring	−30	−10
GLC high ʃ	If Gluc > 0.4 use the desired qs	Monitoring	−15	−10

† The pH range was chosen to lie in the physiological range. However, it could be extended to conditions close to the maximum tolerance which may lie somewhere between pH 6.5 and up to 8.0 for mammalian cells [27]; ‡ The correction order of the set-point is as follows: first pH, then online analyzer. This is important because if the feed is already reduced by pH, it was not done so a second time by the online analyzer. The online analyzer in experiment R-31 had no purpose other than monitoring the metabolite concentrations; ʃ A high glucose concentration is the current status quo in most industrial mammalian cell culture processes.

2.7.3. Operation Window for qs

The set-point for feeding was not fixed but assumed several distinct values in two experiments, which are summarized in Table 2. The rationale for the range calculation of the set-point is based on a possible biomass estimation error and will be explained a little later in both Sections 3 and 4 including an exemplary calculation. In brief, the initial qs set-point was selected to lie in an area relevant for control purposes (reduced or negative Y L/G), in our case −10 and −15 p/cell*h.

3. Results and Discussion

3.1. MVDA for Assessing Lactate Metabolism

To capture the high dimensionality of a heterogeneous dataset, we employed PLS-R, a chemometrics tool [57,58] relatively straight-forward in its application [59]. The single most important

feature we were interested in describing was the lactic acid metabolism. Therefore, only a handful of predictor variables were selected with the following considerations in mind:

1. We wanted to be able to set or influence the selected parameters easily, which is the case for the ones we selected, *i.e.*, pH.
2. We did not wish to become dependent on other parameters by including them in the analysis, which might then not be frequently analyzed, such as amino acids.
3. We did not want to include parameters or define new parameter ranges, which are fixed in a platform/manufacturing process, *i.e.*, pO_2 or the pH (acid/base) regulation strategy.

From the available variables and the considerations above, the following were always available in all processes, for the whole process time, and therefore the most straightforward to describe lactic acid production or consumption (qLac): (i) glucose concentration (GLC); (ii) specific glucose consumption (qGlc); (iii) pH in the form of hydrogen ion concentration (H^+); and (iv) lactic acid concentration (LAC).

A score scatter plot was used to identify all data inside a 95% Hotelling's confidence ellipsoid and used in the subsequent analysis (Figure 2A). A new PLS-R model was built without the outliers, and variable variation (R^2) and variable prediction (Q^2) were calculated. We were interested in the ranking of parameters according to their relevance to reduce qLac to get a better understanding on what to focus on in our experiments. The resulting fit did not play a role in accepting or rejecting the predictive quality of the results as we did not use the model in a quantitative way (Figure 2B). High LAC as well as H^+ prevented further increase of qLac [28], which could be mechanistically explained by chemical gradient action between the inside and outside of the cell [60,61]. The specific glucose consumption, meanwhile, was quite naturally indirectly proportional to lactic acid production and a much better predictor of metabolic behavior than, *i.e.*, GLC levels.

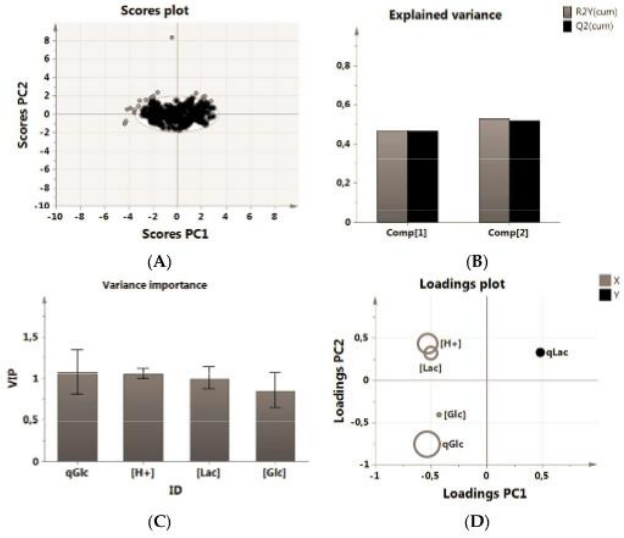


Figure 2. MVDA of the historical data set. (**A**) removal of outliers prior to analysis; (**B**) PLS-R model fit using two Principal Components, here Comp 1 and 2. R2Y(cum) indicates the explained variance, while Q2(cum) explains the predictive quality using one resp. two principal components; (**C**) VIP plot of the most important variables in the analysis sorted by relevance; (**D**) Loading scatter plot of data relationship between predictors (X) and predicted (Y) variable discriminated by VIP size.

A VIP (variance importance of the projection) analysis was used to simplify the multivariate data analysis by calculating a relative score for those predictor variables, which were used to describe the predicted variable qLac. A high VIP score (Figure 2C) resulted in a larger size representation of the given parameter in Figure 2D. The positions of predicted variable qLac together with the predictor variables explained the overall relationship of all parameters; H^+ and LAC were clustered closely to each other and indicate a closer relationship, while all parameters, including GLC and qGlc, were positioned opposite of qLac, indicating an inverse relationship (see the Appendix for more information on the importance of GLC in this analysis, in particular Figures A1 and A2). As GLC had the lowest VIP score, its importance may be understood to be low in this data set.

High GLC has an effect early in the culture while cells are "fit" to take up large amounts of glucose, so to speak. High GLC or overfeeding has almost no effect on cells later on—their maximum qs simply decreases over time, even if GLC was held perfectly constant over time. We can only speak for our experiments, but maybe this was already observed by other labs in which CHO cells were metabolically engineered to yield a reduced lactic acid profile. One would expect qLac to increase along with rising GLC levels, but from historical data this is not exactly the case, which implies that not GLC but the specific glucose consumption rate needs to be tightly controlled.

Summarizing our findings, we found that qGlc was by far the most important parameter to describe qLac behavior in process development runs, while H^+ as well as LAC apparently played a role as well, although much less pronounced. GLC was statistically correlated with qLac due to the nature of the clone to consume lactic acid also during high glucose concentrations, but this particular finding was of little practical use. Therefore, the following experiments featured a modification of the most important parameters, which are qGlc and H^+, in a glucose-limited state.

3.2. Lactic Acid Metabolism

The clone we investigated already featured a reduced LAC profile and was capable of lactic acid consumption as can be seen in Figure 3A,B. MVDA revealed that qLac can be expressed as a function of the parameters qGlc, LAC, H^+, but not necessarily GLC; hence, the feeding had to be based on qGLC and not on GLC itself to improve lactic acid profiles, the latter being the normal control entity in industrial processes.

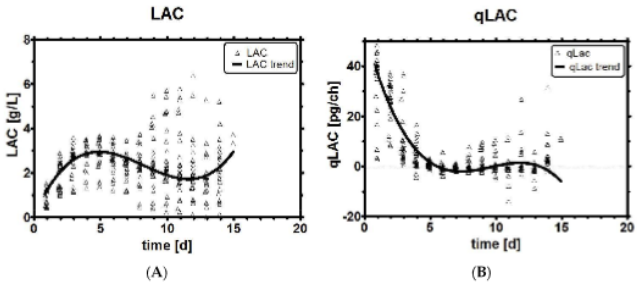


Figure 3. Historical LAC profiles (**A**); Historical qLac profiles (**B**).

A strong early consumption of glucose is the main reason for high LAC levels later on, as will be seen a bit later in chapter 3.6. Historical process development data, where glucose flux was limited and low lactic acid levels were observed during the process, supports this statement. Le [4] shows that early intervention in the process affects the process outcome, which may be, for instance, the lactic acid profile. This intervention was demonstrated in this contribution with an early limitation of glucose availability, especially while the cells are still growing well. The benefit was to prevent LAC buildup and avoid many challenges with LAC and pH in cell culture. As any limitation of substrate

may lead to reduced cell counts, we have tried both a weak and a strong limitation of qGlc to improve the process.

3.3. Impact of pH on qGlc

The historical processes all featured a pH set-point with an upper and lower range, inside of which no control agents are yet used—as result, the pH can assume several values, depending on the current control set-point. This raises the question if pH could be kept somehow stable by design and de-correlated from the typically concomitantly occurring undesired rise in LAC levels at alkaline pH. We differentiate between two different pH states: "pH low" (6.6–6.9) and "pH high" (6.9–7.2).

Figure 4A,B show qGlc and on-line pH profiles of all 29 process development runs. The pH of all observations was split into two groups, higher or lower than pH 6.9, and plotted against qGlc. Under high pH conditions, a high qGLC range was observed, while at acidic pH, qGlc was more restricted (Figure 4C,D). As a result, high pH was often associated with high qGlc and *vice versa*. During the mid- and end-phase, pH did not impact or change qGlc much, while the effect that pH might exert on qGlc was strongest while the cells were growing well. Therefore, pH may be used indirectly to change qLac, as both qGlc and qLac are highly correlated (Figure 2D).

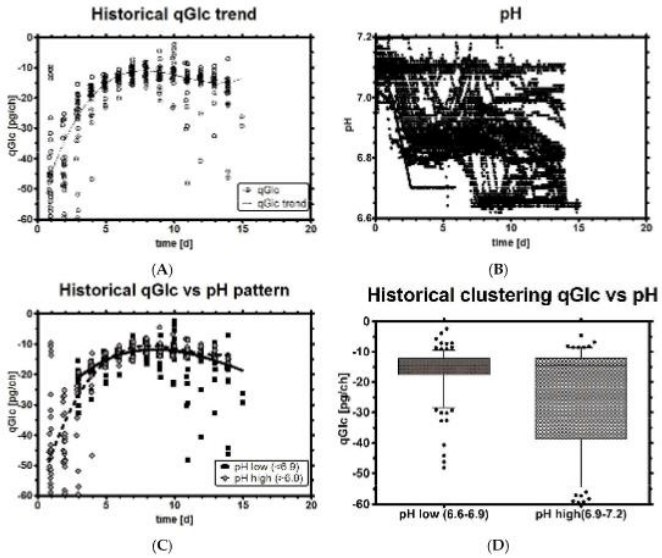


Figure 4. Historical qGlc profiles (**A**); On-line pH profile (**B**); Clustering of qGlc with a pH classifier (dark: low pH, bright: high pH) over time (**C**) and as box plot (**D**). A robust third-order polynomial function was applied in Graphpad Prism software to capture the general trend of a typical qGlc profile in (**A**) and qGlc in (**C**). No runs started with pH < 6.9 as pH initially falls and is then controlled.

3.4. Set-Point Selection for qGlc

The most important parameter which influences qLac and therefore LAC levels in the fermentation is qGlc. Other important parameters which also influence qLac, such as oxygen availability [62] or methods to improve pyruvate uptake into the tricarboxylic acid circle (TCA) [6], are acknowledged but out of scope of this publication. Historical data showed that the cells apparently never consumed less glucose than approximately 8 pg/ch (pg per cell per hour) and the gross of all measured qGlc lied in

the range of 10–20 pg/ch. Plotting qGlc *versus* the yield qLac/qGlc (in short Y L/G) implies that, with the control of qGlc, a particular target yield might be achieved (Figure 5). This relationship may also be used to design a stoichiometric feed by trying to feed at a defined qGlc set-point in further experiments.

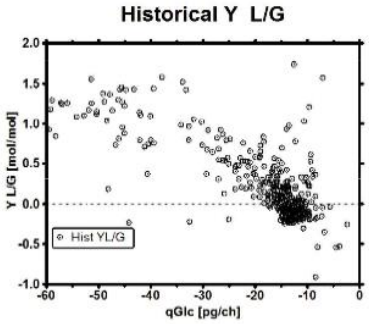
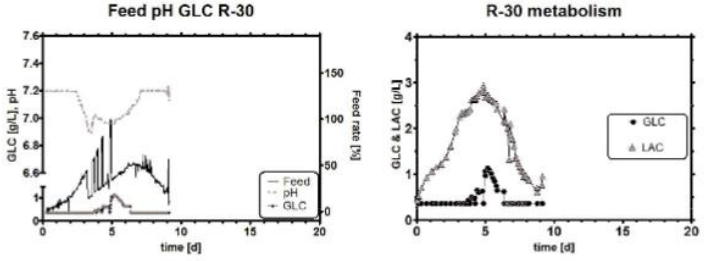


Figure 5. Historical yield qLac/qGlc (Y L/G) *versus* qGlc showing the consequence of particular qGlc on the yield.

3.5. *Impact of a Broad Range as qGlc Set-Point on the Lactic Acid Profile*

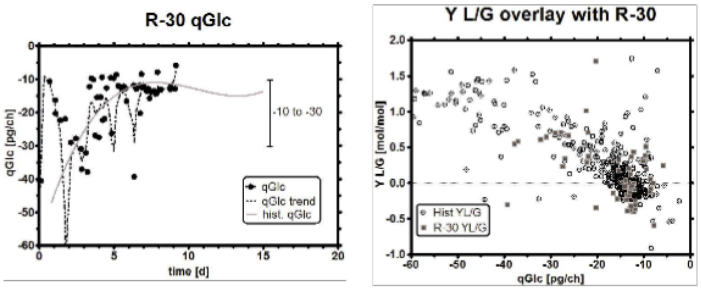
First, we attempted to hold the cell's qGlc set-point in a broad range between 10 and 30 pg/ch over the whole cultivation time by feeding at a target set-point rate which changes stoichiometrically with biomass. The feeding strategy is described in the experimental section (Table 2). For most of the time, the culture was controlled at a glucose level <0.36 g/L with a short excursion to up to 1 g/L when a high feeding set-point was automatically selected and the cells could not consume the substrate quickly enough before the next measurement.

In Figure 6A, it can be seen that the target specific rate of qGlc switched automatically between high and low set-point when certain conditions were fulfilled. Simply put, on-line pH was used as the main switch to affect the feed rate by changing the qGlc set-points between −10 and −30 pg/ch in real time. A second switch was the GLC level measured by an online analyzer, but pH had a higher hierarchy level as it was a faster available signal than the metabolite levels which were measured every couple of hours and so the associated events were rarely ever triggered. When the viable cell concentration reached a maximum, the high qGlc set-point was exchanged with the default set-point (10 resp. 15), because we know from experience that this clone does not require so much substrate so late in the cultivation.



(A) (B)

Figure 6. *Cont.*



(**C**) (**D**)

Figure 6. Experiment R-30, feeding profile with pH and GLC level (**A**); GLC and LAC concentrations (**B**); qGlc with trend and historical qGlc (**C**); overlay of resulting Y L/G profile with historical data (**D**).

Overall, this feeding strategy did not reduce LAC levels by much (Figure 6B) compared to historic reference runs (shown later). This was to be expected as we selected a purposely high qGlc set-point resulting in high glucose uptake rates. The qGlc trend of this run very much resembles the one from the historical runs which are overlaid in Figure 6C,D. As can be seen, qGlc falls well within the previously defined limits and shows the characteristics of a non-glucose-limited run (but *is* glucose- limited), which is relevant because it means two things: first, to achieve a lower qLac, a lower qGlc set-point must be selected (which could be relatively high so that the total carbon influx into TCA is not greatly reduced), and second, and perhaps more interestingly, a high GLC concentration is *not* necessarily required to keep the culture well. This is a very important statement because it implies that mammalian cell cultures may be operated in limitations if the biomass and qGlc can be estimated well. Both come with a particular error and, for this reason, a range rather than a fixed stoichiometric feeding point is necessary to keep the culture well supplied with enough nutrients over the whole process time. In the next chapter we briefly describe how a cell culture may be stoichiometrically fed if errors are rather low with experiment R-31.

3.6. Impact of a Tight Range as qGlc Set-Point on the Lactic Acid Profile

We select a second experiment, in which we used a lower qGlc set-point criterion and selected a very stringent band between −8 and −13 pg/ch as the operational range. The rationale for such a strong glucose inhibition was that lactic acid production should, in theory, be very low because Y L/G approaches zero for this qGlc, as seen in Figure 5. The experiment R-31 shows a few noteworthy differences when compared to any historical process development runs, and, in fact, any runs we have found so far in the literature: first, pH was extremely stable (Figure 7A) during the whole process time and no control action by either acid or base was required. If, as a consequence of this strategy, pH should begin to rise too much in a run, acids are easier to add than to remove from the system. During the whole fermentation time, the tight qGlc set-point was never breached by more than 25% (which corresponds to the error of our biomass estimation method). Even when overestimating biomass, the consequence resulted in a total increase of qGlc by roughly 3 pg/ch, equaling roughly 16 pg/ch at the time of the largest error. This implies that even though biomass estimation is not perfectly accurate and deviates within a certain error range (as observed for the capacitance signal [63], but also any other signal), the qGlc set-point range simply needs to be selected high enough. A possible biomass underestimation cannot threaten the culture and if the set-point is still low enough, the benefits from a reduced LAC buildup can be still fully reaped (see Section 4 for an example).

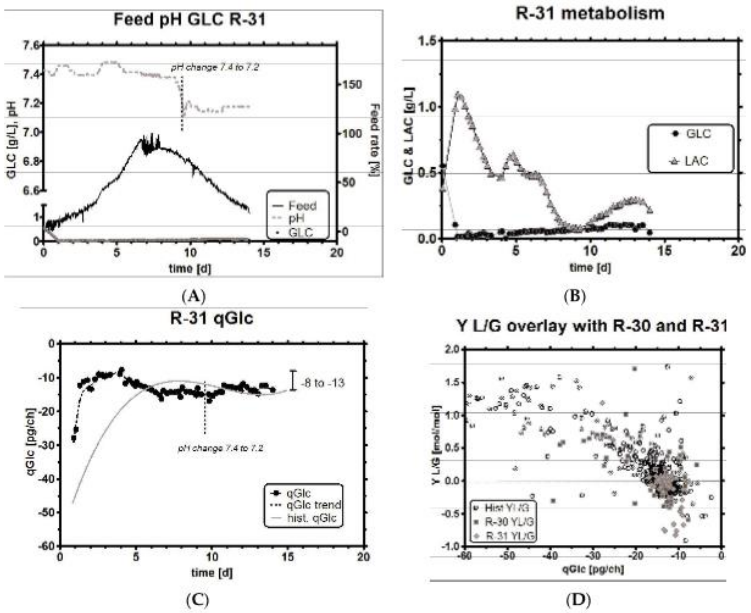


Figure 7. Experiment R-31, feeding profile with pH and GLC level (**A**); GLC and LAC concentrations (note that leftover GLC is consumed and results in a small LAC until it is consumed, but qs changes immediately to the target value once GLC is limited by the stoichiometric feeding strategy) (**B**); qGlc (including dotted trend line) and historical qGlc (full line) (**C**); overlay of resulting Y L/G profile with historical data (**D**). Both targets (qGlc and resulting lower Y L/G) can be better achieved due to the tighter control specification, when compared with run R-30. A comparison follows in Table 3.

Table 3. Statistical evaluation: in experiment R-30, a slight improvement of Y L/G and P/G compared to the historical runs can be seen. In experiment R-31, the improvements for Y L/G and P/G are most pronounced. The variation of qGlc could be massively reduced as can be seen in the low value for the standard deviation. N describes the number of available sampling events in the historical data and the experiments R-30 and R-31.

	qGlc [pg/ch]			Y L/G [-]			Y P/G [-]		
ID	**Historical**	**R-30**	**R-31**	**Historical**	**R-30**	**R-31**	**Historical**	**R-30**	**R-31**
MEAN	−22.39	−19.15	−13.92	0.41	0.18	0.03	0.29	0.31	0.36
SD	16.53	14.62	3.64	0.63	0.49	0.31	0.12	0.05	0.1
$N_{All} = 667, N_{R-30} = 60, N_{R-31} = 56$									

Another observation was the potential effect that pH may have had on qGlc. In all historical experiments, qLac was strongly correlated with pH, but here, for the first time, we could successfully decouple both parameters and investigate what happens if a high pH meets low LAC. During biomass overestimation, we could see that qGlc is higher than the historical reference but within the previously selected qGlc range of −8 to −13 pg/ch (Figure 7C). After the viable cell concentration maximum was reached, a pH change of −0.2 units brought the pH down from 7.4 to 7.2 (Figure 7A). After a delay of one to two days, qGlc decreased slightly by *ca.* 15%, which cannot be seriously discussed in terms of statistical significance. However, the direction of qs change after pH change would fit well into the

overall picture of pH affecting qs (pH acidic: decreased GLC influx, pH basic: more GLC influx into the cells). It may be too early to conclude that pH could be used to modulate the maximum possible qGlc so extremely easily (which depends, among others, also on amino acid availability). However, the data suggest that exactly this may be the case, so this claim might be worth being explored by other research groups in the future.

We claim to have reached the goal of this study to reduce lactic acid profiles successfully, as can be seen in Figure 7B, with a minimum amount of experiments. Even though we exposed cell cultures to an unusually high pH of 7.4, which is usually correlated with high qLac, our feeding strategy showed exceptionally low LAC levels by controlling qGlc tightly and accurately. The culture could be kept in a metabolically highly interesting state for a duration of 14 days in a dynamic fed-batch process. The resulting Y L/G was, therefore, around zero most of the time and also assumed negative values, which are found to correlate with high titers in cell culture (Figure 7D) [5,64].

3.7. *Adaptive Feeding Using Real-Time Switches*

3.7.1. Adaptive Feeding Using pH Correction

It is not in the nature of an error to announce an over- or under-estimation, so when exactly can the target set-point be switched on or off with an upper or lower range set-point, and how high or low can those possible set-points be defined? The answer to these questions lies in the enormous wealth of on-line signals which can be used exactly for this purpose [35,65]. In the first experiment, pH and the GLC level were used, and in another only pH was enough to switch the qGlc set-point to particular levels automatically. As the scale-down bioreactor model mirrors a manufacturing process with a floating pH dead band, we found pH suitable to automatically correct the feeding set-point. As a consequence, acid/base control for pH was not required at all. This may have had a positive impact on cellular physiology and, as a consequence, on the final titer. The strategy to include pH in set-point adaptation was adapted from Gagnon [42], but our control is more tightened due to the real-time and closed-loop approach instead of using predetermined rates. In the first experiment, we selected the switching criteria a little too broadly, and the set-point jumped at the smallest occasion, as can be seen in the feeding profile of run R-30 (Figure 6B). The reason for this was partially an unlucky first selection of the Boolean which often activated and deactivated the switch, but also the combined application of another switch from the online analyzer, as we had both signals at our disposal.

3.7.2. Adaptive Feeding Using an Online Metabolic Analyzer

Ozturk described how online analyzers were used to hold a particular substrate or metabolite level [38]; however, for our problem statement this was not necessary, as the low glucose level posed no threat to the culture due to robust stoichiometric feeding. The online analyzer was working well but was actually not required to keep lactic acid levels low in a dynamic fed-batch. Therefore, we restricted the use of the set-point switch in experiment R-31 to pH only (Figure 7B). We conclude that one well-chosen switching criterion is enough for the task to keep the control strategy simple and intuitive. Although signals are available every minute, we advise to reduce control action to every half hour or more—otherwise, tiny excursions of the control range lead to feed corrections in the same time which makes the feeding profile impossible to interpret without strong smoothing. However, process development may very much benefit from highly frequently measured concentrations (*i.e.*, in the hour range) and unlock precise event detection, *i.e.*, after pulsing or shifting [66,67], to develop process understanding which may then translate into better process development and control.

3.8. *Comparison of Experiments with Historical Performance*

Direct comparison of key differences between historical data and the experiments reveals the following findings: Selecting the qGlc band for a stoichiometric feeding strategy in a very high range from −10 to −30 pg/ch led to comparable concentrations (Figure 8A) and metabolic profiles (Figure 8C)

to the reference runs, even though glucose was limited at levels around 0.36 g/L (with a temporary shoot-up up to 1.2 g/L GLC). The operation in limited GLC conditions did not lead to a reduction of growth at all, as can be seen in the exceptionally high cell concentration (Figure 8E), but as cells started dying very fast in this process, a high final product titer was finally not obtained in this fermentation (Figure 8F).

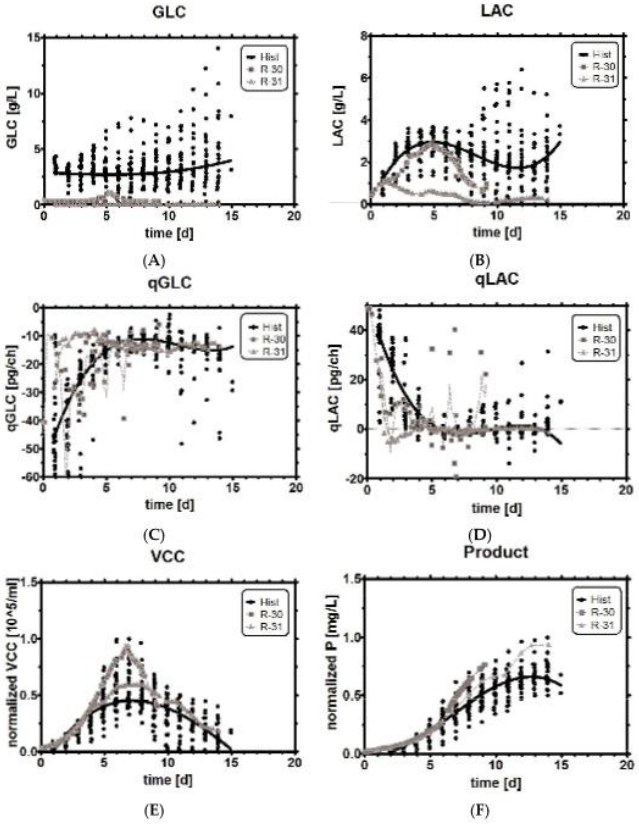


Figure 8. Benchmarking limited glucose control with historical data (black dots: historical data, squares: experiment R-30 with a broad qGlc set-point, triangles: experiment R-31 with a tight qGlc set-point. (**A**) and (**B**): metabolic profiles of glucose and lactic acid; (**C**) and (**D**): Specific glucose and lactic acid rates; (**E**) and (**F**): Viable cell concentration and product titer compared to historical data.

In contrast to a very broad selection of qGlc, a very tightly controlled qGlc band between −8 to −13 pg/ch resulted in almost immediate LAC uptake from the environment (Figure 8B,D), while at the same time limiting the availability of GLC to a stoichiometric feed (Figure 8A,C). As can be seen, such a strong substrate limitation showed an effect on the growth profile, and on the maximum cell concentration (Figure 8E), which was somewhat low but still higher than the average mammalian cell cultivation. However, the declining phase was much softer and flatter, which may have been the reason for the high final titer at the end of the process (Figure 8E,F) as cells were still productive

instead of being in the process of quickly dying. To be more precise, we observed that the majority of the cells in the tight qGlc limitation did not turn into dead cells, but lysed, which could be seen in the high viability (80%) at the end of the process run time (Figure A3 in the Appendix). This could indicate that cells took a different way to die than under regular process conditions. In a speculative afterthought, development of drugs which improve resistance to either cell lysis or apoptosis [68–70] might help to prolong cultivation (and production) time. A lower maximum viable cell count (VCC), but at the same time higher titer, suggests that the culture may have been simply more productive or productive for longer than the references.

From the experiments shown, we cannot truly distinguish which of the two parameters (qGlc or pH) finally led to more product. We assume that qGlc had the higher impact, simply because there is no truly direct mechanistic link between productivity and pH. It may have been a combination of both effects, but this statement would require much more rigorous testing, *i.e.*, by running a DoE to determine the exact contributions and interactions (of H+, qGlc set-point, qGlc upper and lower range, Y L/G, other low GLC levels, *etc.*), which was far out of scope of this contribution. Table 3 summarizes the experimental findings with regard to qGlc, Y L/G and Y P/G. The strongest limitation of average qGlc over the whole process time (−62% compared to the reference) resulted in a decrease of the Y L/G yield to almost zero, while at the same time the product yield was increased by 25%. For this we only needed to estimate biomass using a capacitance probe, which is already an accepted standard in many development labs. Usually, an increase in product yield has to be critically regarded, especially if the final titer turns out to be lower overall than the reference, but this is very clearly not the case here.

4. Conclusions

4.1. Eliminating the Root Cause for High Lactic Acid Concentrations

Historical data was analyzed by MVDA techniques where the specific glucose uptake rate was identified as the most relevant parameter to reduce the specific lactic acid production rate and, as a consequence, the lactic acid profiles of the culture. A feeding strategy was developed on this basis to set up a target specific glucose consumption rate for the whole process runtime of a fed-batch. This target rate was set up by estimating cell count via a capacitance probe, which can be used either for process monitoring [71–74] or control purposes [47,75]. Feeding was realized by supplying the cells stoichiometrically in a range between 8 to 30 pg/ch in real time with feed solution instead of using a previously defined off-line feeding profile which cannot react to deviations in the present process. Glucose is shown to be efficiently held constant to levels around 2–3 times the Km of mammalian cell lines during the whole process time without having a negative effect on viable cell concentration, as long as the specific glucose consumption feeding rate is selected to be high. Although the main effects leading to a reduction of lactic acid happen in the first days, it is well worth mentioning the remainder of the experiment. The proposed strategy also keeps the culture in the decline phase in a stoichiometrically fed state. This is a very important finding because it eliminates otherwise-required operator interventions and makes it easy to integrate it in an automatized manufacturing environment. We demonstrate that, if a strong limitation of the specific glucose consumption rate is selected, a very positive effect on the stability of the pH signal was observed together with a complete prevention of lactic acid build-up.

4.2. Directions from MVDA

Multivariate data analysis (MVDA) techniques can be tricky to use and sometimes misleading, as we have seen with the parameter GLC. The historical data set was comprised of glucose levels which are far above the cell's physiological limit, where any effect on the specific glucose consumption rate could be expected. "Low" glucose levels (around 2–3 g/L) correlated to high lactic acid production rates, which makes no sense biochemically, but the largest part of the given data did not feature a glucose range in which qGlc was mechanistically affected. Because glucose concentration had the

overall lowest weight in the analysis, and because the model was used in an explorative rather than quantitative way, the fit could be, from a statistical point of view, accepted. PLS-R was still very useful to correctly identify the exceptional importance of qGlc and the close relationship between H^+ and LAC. Ranking their importance with regard to the impact on qLac could, furthermore, save valuable time by running the most interesting experiments first, instead of screening all eventualities. Conclusively, MVDA results were helpful to save time, but must be always reviewed carefully by considering whether the used data is really adequate and truly representative for the given problem.

4.3. Hypothesis for the Positive Effect of pH on Productivity

In our opinion, one facet of pH-dependent flux modulation remains overlooked until today and might be used to maximize the production of high-quality therapeutics of tomorrow: reverting the same method which researchers were using for years to reduce metabolic flux [30,76] into lactic acid and redirecting it to the TCA [50,77], we *increase* pH to maximize metabolism in substrate limitations. This is especially important in the light that lactic acid toxicity is actually more pronounced at lower pH than at higher pH [61,78–80] and may be the cause of more cell lysis [81], resulting in the lower productivity of a process. Flux increase is hypothesized to result from the cell's regulation of its internal pH by a plethora of transporters [82–85] in the following manner: many enzymes involved in glycolysis are strongly affected by pH, among them those considered as metabolic bottlenecks, such as HK (hexokinase), PFK (phosphofructokinase) and others [86–89]. Even though the cytosol of the cell is strongly buffered, a certain gradient difference of around 1.5 units cannot be crossed if the cell should stay alive, not even if the cell in question is one of the strongest lactic acid producers [90] we are aware of today. Therefore, it is assumed that the internal pH, however buffered, must be affected by the external pH; thus, metabolism can be directed to different fates. The increase in flux is possibly distributed to all nodes and not just one node of the metabolism, among them the one responsible for product formation [77,91], as symbolized in Figure 9. We believe that the herein presented novel concept may lead to more biomass, which, in turn, may lead to higher titers [19], since more substrate leads to more cells which can produce more product due to an improved utilization of energy and nutrients.

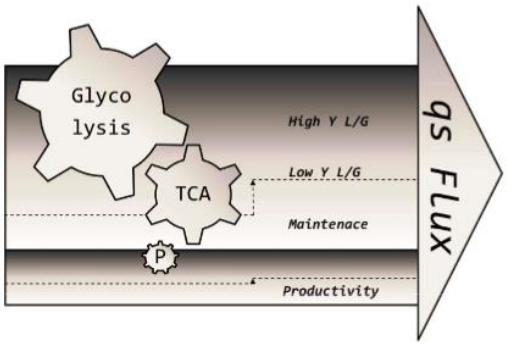


Figure 9. Suggested flux dependency on productivity. A higher overall flux may lead to a higher flux in all metabolic nodes, including lactic acid secretion and the one responsible for product formation. Usually, the basic consumption of cells does not change much because pH is controlled tightly. However, by setting and holding physico-chemical conditions which act on this flux by a high pH, while still operating in a relatively low Y L/G range, the overall higher flux may be cumulatively translated into higher productivity.

4.4. Living with Uncertainty

The biggest risk in the experiments was biomass underestimation, as this might mean starvation and an early stop of the process. In case that biomass estimation comes with a high error, two possibilities can be considered to cope with the uncertainty: changing the target qGlc set-point to a safer operation set-point or giving the target set-point a safe *range*, in which it can move and still satisfy the goals. The maximum error of biomass estimation may depend on the model by which it is assessed and have an asymmetrical distribution, *i.e.*, in an imaginary worst case scenario (Figure 10) of an overestimation of up to +50% and underestimation up to −33% (our biomass estimation in these experiments was in the range of maximum 25% over the whole process time, including the death phase of the culture). We want to point out that such a large uncertainty might not be as problematic as it sounds, as variation is a part of the daily business when dealing with biological systems [92]. In this contribution, we demonstrated how a robust qGlc set-point could be developed for any process using viable cell count estimation to calculate a safe feeding rate: in the historical data set we see that this clone almost always consumed at least 10 pg/ch to maintain metabolic integrity (Figure 7D). We exemplarily calculated a target set-point on the basis of the maximum error from biomass underestimation (Table 4) so, even with a high error on biomass, the culture will be supplied with sufficient nutrients in a glucose-limited fed-batch.

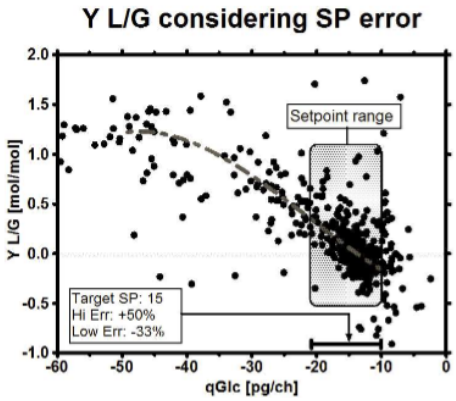


Figure 10. Metabolic control window with a suggested robust qGlc set-point. Data points represent both historical and experimental data. Depending on the task, the set-point range may be selected either conservatively (better overfeeding than starving, higher qGlc) or more courageously (improved metabolic state at the cost of reduced growth, lower qGlc). In general, the error of the online biomass estimation method may serve as a good first estimate for the range. It must be noted that the cells might not truly consume at a high qGlc just because it is desired by the operator and fed in this way, especially in the mid-phase of the culture—a fact which we suggest to solve by applying a high pH.

Table 4. Consequences of a high error for a robust set-point (SP) selection. To prevent starvation of the culture, the lower set-point range must be high enough to compensate for a possible underestimation of biomass.

	SP	*Lower Range SP*	*Upper Range SP*
Error	0%	−33%	+50%
Desired set-point considering error	10	6.7	15
Robust target set-point	**14.9**	**10**	**22.4**

4.5. Limitations and Suggested Improvements

The range for qGlc enabled the introduction of a safety margin, which may be set up according to the error of the biomass estimation if certain conditions are fulfilled (*i.e.*, on-line available metabolic results *or* on-line pH). Especially in the case of a signal as simple as pH, of course not all is well. Especially in a buffered system, pH hardly changes due to cellular activity early in the process, but rather, because of process events, pH probes are known to drift. Therefore, other signals which may replace pH as the set-point switch are, for instance, capacitance, conductivity, pO_2, pCO_2, offgas, mass spectroscopy (MS), turbidity, base, inline microscopy, heat exchange, fluorescence, infrared, metabolic ratios, mass flow controller, online analyzer and many more [35,93]. If expensive high-tech equipment is not available, the controller output from proportional-integrative-derivative (PID) control [94–96] of suitable devices might pose an interesting alternative. A change in metabolic activity might be even better detected by using derivatives [2] or otherwise modified raw signals in combination with a robust signal-processing algorithm which may solve the possible problem with outliers very elegantly [55,97–99].

4.6. Summary and Outlook

The goal of the study was to reduce lactic acid accumulation in fed-batch fermentations in process development to increase overall process performance. We developed an adaptive feeding strategy which was based on real-time signals, capacitance and pH, and found that an on-line analyzer was not necessary for our intended control approach in glucose limitation. We have shown for the first time that the deliberate decrease of pH [26,29,32], which impairs lactic acid production but also cellular proliferation, was not necessary to stop lactic acid accumulation in the bioreactor, if the metabolic state could be controlled with a stoichiometric feeding strategy in which uncertainty is taken into account by design. As high lactic acid levels no longer posed a threat to the culture, a higher glucose consumption rate, favored, *i.e.*, by a high pH set-point, may boost productivity [81]. Experimental demonstration furthermore led to a more consistent pH and glucose profile, which might increase product consistency, as was reported by other authors [32,100,101], or simply process robustness. The uncertainty of biomass estimation was solved by designing a range of uncertainty-based feeding set-points, which may switch the feed rate in real time when combined with the previously suggested on-line-accessible parameters. The basis of our methodology, using specific uptake rates, is independent of scale, location and initial conditions. Therefore, the simplicity of the suggested method allowed a transfer of the desired *metabolic state* between piloting and manufacturing, which fits perfectly in the context of Process Analytical Technology (PAT) and Quality by Design (QbD) with respect to an improvement of robustness [1,102–104].

A process run under such conditions from start to end holds a few yet unmentioned benefits in terms of cost of goods and productivity [105]: Healthy cultures with high viability are likely to create less cell debris because of the lower dead cell count, and the here-proposed feeding strategy is likely to result in a low residual substrate concentration at the end of the fermentation, which may potentially improve harvest efficiency. This naturally has an impact on subsequent unit operations [106], which, for example, include time-intensive separation, washing and purification steps [92,107] (Figure 11). We therefore hypothesize that the proposed strategy may be suitable to increase plant productivity and costs in the unit operations after the actual fermentation [108], which leads to leaner processes with a shorter turnover time and result in an overall improvement in productivity.

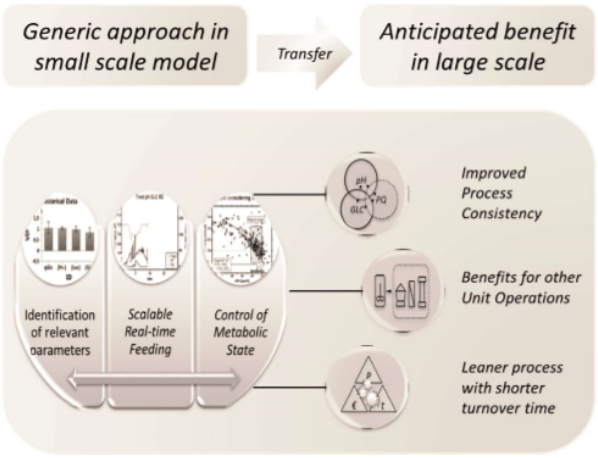


Figure 11. Generic approach for the holistic improvement during mammalian cell culture development.

Acknowledgments: We sincerely acknowledge Civan Yagtu for assistance in the lab, Patrick Wechselberger for providing technical support (Lucullus PIMS) also outside working hours, Christian Dietzsch for configuration of the online analyzer (CubianXC) and Hans Lohninger from the Vienna University of Technology for kindly providing us with a full version of Datalab [54] which was used to generate the model to estimate biomass for our feeding strategy.

Author Contributions: Viktor Konakovsky designed, planned and performed the experiments at the Vienna University of Technology, analyzed the data, performed statistical analysis, prepared tables, Figures, additional files, drafted and wrote the manuscript. Christoph Clemens gave valuable inputs while the manuscript was drafted. Martina Berger, Markus Michael Müller, Christoph Clemens and Jan Bechmann contributed the largest part of the data, helped with critical discussions and gave valuable support also for advanced feeding control. Christoph Herwig passed on personal and professional experience, especially in automation, logic, and structure, helped drafting the manuscript and supported training activities during the project. Christoph Herwig and Stefan Schlatter supervised research and conceived the study. All authors read and approved the final manuscript. Partial financial support from the CATMAT Doctorate College of the Vienna University of Technology is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CHO: Chinese Hamster Ovary
CHO-DG44: Clone B is derived from this cell line
Clone A: Clone in use for a biomass model based on capacitance
Clone B: Clone in this contribution (historical and experimental data)
CVRMSE: Coefficient of variation of RMSE (Root Mean Square Error)
DoE: Design of Experiments
Feed concentration, $_{Substrate}$: Here: Glucose substrate concentration in the feed [mg/mL]
GLC: Glucose concentration [g/L]
GS-CHO: Glutamine-Synthetase CHO
H^+: Hydrogen ion concentration [mol/L]
Hist: Historical runs from process development
HK: Hexokinase
IVCC: Integrated viable cell concentration [c/mL]
Km: Clone-dependent Monod constant for glucose affinity (0.1–1 g/L) [g/L]

LAC: Lactic acid concentration [g/L]
mAbs: Monoclonal Antibodies
MVDA: Multivariate Data Analysis
N: Number of observations
OD: Optical density [-]
PAT: Process Analytical Technology
pg/ch: Picogram per Cell per hour
PFK: 6-Phosphofructo-1-kinase
PLS: Partial Least Squares
PLS-R: Partial Least Squares Regression
Q^2: Variable prediction coefficient
QbD: Quality by Design
qGlc: Specific glucose uptake rate [pg/ch]
qLac: Specific lactic acid uptake/production rate [pg/ch]
qOUR: Specific oxygen uptake rate [mmol/ch]
qs adapted: New qs set-point after online control action (*i.e.*, pH or online analyzer)
qs: General specific uptake rate notation, identical with qGlc as s stands for glucose [pg/ch]
R^2: Regression Coefficient [-]
R-30: Experiment R-30, broad range for target qGlc set-point
R-31: Experiment R-31, tight range for target qGlc set-point
rlowess: Robust local regression using weighted linear least squares
SP: Set-point
SD: Standard Deviation
$V_{Reactorx}$: Volume of the reactor [mL]
VCC: Viable Cell Concentration [cells/mL]
VIP: Variance importance of the projection, a measure of relevance of the parameter
Y GLC/OX: Yield glucose per oxygen [mol/mol]
Y L/G: Yield lactic acid per glucose [mol/mol]
Y P/G: Yield product per glucose [mg/g]
Y X/GLC: Yield biomass per glucose [cells/g]
Y X/GLN: Yield biomass per glutamine [cells/g]

Appendix A. The Importance of Glucose

A high glucose concentration was not necessarily required to keep the culture well, as was shown in this contribution. Often, cell cultures in manufacturing and process development are operated "between 1 and 6 g/L" without further ado. One reason for this lies in the natural fear to lose a culture if the cells should be exposed to substrate limitations and starve, for instance because a stoichiometric feeding regime did not work as expected. Another reason for running at a high target glucose concentration range may have to do with glycosylation patterns of the final product which may be influenced by GLC levels [109]. However, the intracellular flux and concentrations might surpass the extracellular concentration of substrates in importance for the final structure of a complex protein [110]. Finally, some clones do not grow well below a certain glucose threshold and a limitation strategy is not the right approach for success; however, in this study, it worked very well.

To understand why GLC was often not a crucial parameter for lactic acid metabolism, it is necessary to take a look at the raw data. All historical runs started off with a remarkably high qGlc, which, over the course of a fed-batch fermentation, wore off and approached a certain plateau. In other words, qGlc (Figure 4A) was not solely governed by GLC as often assumed in microbial cultivations and appeared to be somewhat insensitive to a *high* GLC presence (see Figure A1) late in the culture, which means that even concentrated bolus shots would not change qGlc at this time point. GLC levels

in cell culture were traditionally almost always operated way above the Monod constant Km, which depends on the genetic traits of the clone and may range between 0.1–1 g/L [111].

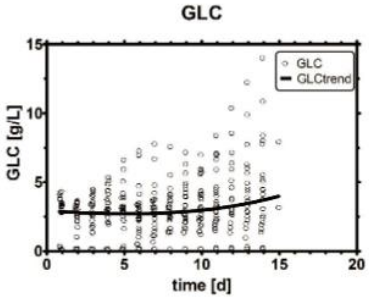


Figure A1. Historical GLC profiles. The GLC level itself was often inadequate to describe metabolic effects, such as the relation to qGlc in fed-batches. A robust third-order polynomial indicated an average trend of all runs, regardless of the individual trend in each cultivation.

Appendix B. Coefficients in Multivariate Data Analysis (MVDA)

The importance of parameters can be obtained by looking at the coefficients of the principal components, which also shows their effect on the target variable. (Here: qLac; positive coefficients indicate a direct correlation, while negative coefficients indicate inverse correlation. All coefficients were negative in this analysis). For their variable importance ranking, the coefficients are sorted regardless of their effect on the target parameter with decreasing importance, and this can be visualized by the VIP plot. In this dataset, the coefficients for glucose concentration (GLC) made apparently no sense. A negative coefficient would imply that a low glucose concentration was related to a high lactic acid production rate. GLC was almost always very high (>>>Km of glucose of the clone) and this clone was capable of consuming lactic acid even when there was plenty of glucose available. As a consequence, this behavior was captured in the statistical analysis and finally led to negative coefficients for glucose in this data set.

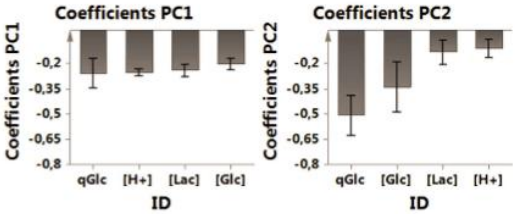


Figure A2. MVDA coefficients of PC1 and PC2 capturing direction of parameter effect on the target variable (qLac). PC1 captures most variance in the following order: qGlc, H+, Lac, GLC, while PC2 captures some remaining variance in qGlc and Glc but not Lac and H+ as they are already well represented by PC1.

Appendix C. Viability

R-31 shows very high viability of the remaining cell population at the end of 14 days, while at the same time the total viable cell count was decreasing. A large portion of cells were dying in a much more different way than in the other process development bioprocesses. Whether this should be

traced back to the difference in pH or the metabolic state could not be investigated in more detail in this contribution.

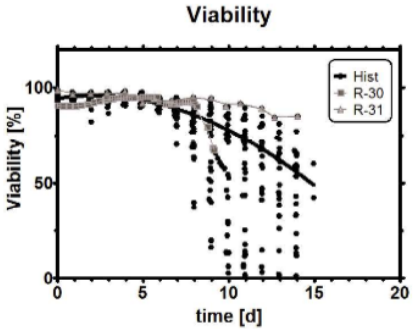


Figure A3. Viability trend in historical data, R-30 and R-31. The latter run had the highest viability at the process end. R-31 was tightly controlled throughout the process time in terms of metabolic state and featured a pH change from 7.4 to 7.2 between days 9 and 10.

Notes: All figures were optimized for printing; however, they are also made available in color in the supplementary materials of this journal. We encourage the reader to take a look at them as tiny differences inside the paper become much clearer when looking at the figures in full scale at a high resolution.

References

1. Rathore, A.S.; Mhatre, R. *Quality by Design for Biopharmaceuticals: Principles and Case Studies*; Auflage: 1. Wiley-Interscience: Hoboken, NJ, USA, 2011.
2. Charaniya, S.P. Systems Analysis of Complex Biological Data for Bioprocess Enhancement. Ph.D. Dissertation, University of Minnesota, Minneapolis, MN, USA, 2008.
3. Croughan, M.S.; Freund, N.W. Strategy to Reduce Lactic Acid Production and Control PH in Animal Cell Culture. U.S. Patent US8470552 B2, 25 June 2013. Available online: http://www.google.com/patents/US8470552 (accessed on 6 January 2016).
4. Le, H.; Kabbur, S.; Pollastrini, L.; Sun, Z.; Mills, K.; Johnson, K.; Karypis, G.; Hu, W.-S. Multivariate analysis of cell culture bioprocess data—Lactate consumption as process indicator. *J. Biotechnol.* **2012**, *162*, 210–223. [CrossRef] [PubMed]
5. Le, H.; Castro-Melchor, M.; Hakemeyer, C.; Jung, C.; Szperalski, B.; Karypis, G.; Hu, W.-S. Discerning key parameters influencing high productivity and quality through recognition of patterns in process data. *BMC Proc.* **2011**, *5*. [CrossRef] [PubMed]
6. Zhou, M.; Crawford, Y.; Ng, D.; Tung, J.; Pynn, A.F.J.; Meier, A.; Yuk, I.H.; Vijayasankaran, N.; Leach, K.; Joly, J.; *et al.* Decreasing lactate level and increasing antibody production in Chinese Hamster Ovary cells (CHO) by reducing the expression of lactate dehydrogenase and pyruvate dehydrogenase kinases. *J. Biotechnol.* **2011**, *153*, 27–34. [CrossRef] [PubMed]
7. Hu, W.-S.; Kantardjieff, A.; Mulukutla, B.C. Cell Lines that Overexpress Lacatate Dehydrogenase c. Patent WO2012075124 A2, 7 June 2012.
8. Brown, N.J.; Higham, S.E.; Perunovic, B.; Arafa, M.; Balasubramanian, S.; Rehman, I. Lactate Dehydrogenase-B Is Silenced by Promoter Methylation in a High Frequency of Human Breast Cancers. *PLoS ONE* **2013**, *8*. [CrossRef] [PubMed]
9. Legmann, R.; Melito, J.; Belzer, I.; Ferrick, D. Analysis of glycolytic flux as a rapid screen to identify low lactate producing CHO cell lines with desirable monoclonal antibody yield and glycan profile. *BMC Proc.* **2011**. [CrossRef] [PubMed]

10. Ma, N.; Ellet, J.; Okediadi, C.; Hermes, P.; McCormick, E.; Casnocha, S. A single nutrient feed supports both chemically defined NS0 and CHO fed-batch processes: Improved productivity and lactate metabolism. *Biotechnol. Prog.* **2009**, *25*, 1353–1363. [CrossRef] [PubMed]
11. Altamirano, C.; Paredes, C.; Illanes, A.; Cairó, J.J.; Gòdia, F. Strategies for fed-batch cultivation of t-PA producing CHO cells: Substitution of glucose and glutamine and rational design of culture medium. *J. Biotechnol.* **2004**, *110*, 171–179. [CrossRef] [PubMed]
12. Wahrheit, J.; Nicolae, A.; Heinzle, E. Dynamics of growth and metabolism controlled by glutamine availability in Chinese hamster ovary cells. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 1771–1783. [CrossRef] [PubMed]
13. Kim, S.H.; Lee, G.M. Down-regulation of lactate dehydrogenase-A by siRNAs for reduced lactic acid formation of Chinese hamster ovary cells producing thrombopoietin. *Appl. Microbiol. Biotechnol.* **2007**, *74*, 152–159. [CrossRef] [PubMed]
14. Müller, D.; Katinger, H.; Grillari, J. MicroRNAs as targets for engineering of CHO cell factories. *Trends Biotechnol.* **2008**, *26*, 359–365. [CrossRef] [PubMed]
15. Luan, Y.-T.; Stanek, T.C.; Drapeau, D. Controlling Lactic Acid Production in Fed-Batch Cell Cultures via Variation in Glucose Concentration; Bioreactors and Heterologous Gene Expression. U.S. Patent US7429491 B2, 30 September 2008. Available online: http://www.google.com/patents/US7429491 (accessed on 6 January 2016).
16. Drapeau, D.; Luan, Y.-T.; Stanek, T.C. Restricted Glucose Feed for Animal Cell Culture. Patent WO2004104186 A1, 2 December 2004. Available online: http://www.google.com/patents/WO2004104186A1 (accessed on 6 January 2016).
17. Basch, J.O.; Gangloff, S.; Joosten, C.E.; Kothari, D.; Lee, S.S.; Leister, K.; Matlock, L.; Sakhamuri, S.; Schilling, B.M.; Zegarelli, S.G. Product Quality Enhancement in Mammalian Cell Culture Processes for Protein Production. Patent WO2004058944 A2, 15 July 2004. Available online: http://www.google.com/patents/WO2004058944A2 (accessed on 6 January 2016).
18. Sauer, P.W.; Burky, J.E.; Wesson, M.C.; Sternard, H.D.; Qu, L. A high-yielding, generic fed-batch cell culture process for production of recombinant antibodies. *Biotechnol. Bioeng.* **2000**, *67*, 585–597. [CrossRef]
19. Templeton, N.; Dean, J.; Reddy, P.; Young, J.D. Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture. *Biotechnol. Bioeng.* **2013**, *110*, 2013–2024. [CrossRef] [PubMed]
20. Mulukutla, B.C.; Khan, S.; Lange, A.; Hu, W.-S. Glucose metabolism in mammalian cell culture: New insights for tweaking vintage pathways. *Trends Biotechnol.* **2010**, *28*, 476–484. [CrossRef] [PubMed]
21. Yongky, A.; Lee, J.; Le, T.; Mulukutla, B.C.; Daoutidis, P.; Hu, W.-S. Mechanism for multiplicity of steady states with distinct cell concentration in continuous culture of mammalian cells. *Biotechnol. Bioeng.* **2015**, *112*, 1437–1445. [CrossRef] [PubMed]
22. Mulukutla, B.C.; Yongky, A.; Daoutidis, P.; Hu, W.-S. Bistability in Glycolysis Pathway as a Physiological Switch in Energy Metabolism. *PLoS ONE* **2014**, *9*. [CrossRef] [PubMed]
23. Wahrheit, J.; Niklas, J.; Heinzle, E. Metabolic control at the cytosol-mitochondria interface in different growth phases of CHO cells. *Metab. Eng.* **2014**, *23*, 9–21. [CrossRef] [PubMed]
24. Miller, W.M.; Blanch, H.W.; Wilke, C.R. A kinetic analysis of hybridoma growth and metabolism in batch and continuous suspension culture: Effect of nutrient concentration, dilution rate, and pH. *Biotechnol. Bioeng.* **1988**, *32*, 947–965. [CrossRef] [PubMed]
25. Ivarsson, M.; Noh, H.; Morbidelli, M.; Soos, M. Insights into pH-induced metabolic switch by flux balance analysis. *Biotechnol. Prog.* **2015**, *31*, 347–357. [CrossRef] [PubMed]
26. Trummer, E.; Fauland, K.; Seidinger, S.; Schriebl, K.; Lattenmayer, C.; Kunert, R.; Vorauer-Uhl, K.; Weik, R.; Borth, N.; Katinger, H.; *et al.* Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors. *Biotechnol. Bioeng.* **2006**, *94*, 1033–1044. [CrossRef] [PubMed]
27. Osman, J.J.; Birch, J.; Varley, J. The response of GS-NS0 myeloma cells to pH shifts and pH perturbations. *Biotechnol. Bioeng.* **2001**, *75*, 63–73. [CrossRef] [PubMed]
28. Mulukutla, B.C.; Gramer, M.; Hu, W.-S. On metabolic shift to lactate consumption in fed-batch culture of mammalian cells. *Metab. Eng.* **2012**, *14*, 138–149. [CrossRef] [PubMed]
29. Joosten, C.E.; Leist, C.; Schmidt, J. Cell Cultivation Process. U.S. Patent US8765413 B2, 1 July 2014. Available online: http://www.google.com/patents/US8765413 (accessed on 6 January 2016).

30. Ozturk, S.S.; Palsson, B.O. Growth, metabolic, and antibody production kinetics of hybridoma cell culture: 2. Effects of serum concentration, dissolved oxygen concentration, and medium pH in a batch reactor. *Biotechnol. Prog.* **1991**, *7*, 481–494. [CrossRef] [PubMed]
31. Patel, S.D.; Papoutsakis, E.T.; Winter, J.N.; Miller, W.M. The lactate issue revisited: Novel feeding protocols to examine inhibition of cell proliferation and glucose metabolism in hematopoietic cell cultures. *Biotechnol. Prog.* **2000**, *16*, 885–892. [CrossRef] [PubMed]
32. 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]
33. L'Allemain, G.; Paris, S.; Pouysségur, J. Growth factor action and intracellular pH regulation in fibroblasts. Evidence for a major role of the Na^+/H^+ antiport. *J. Biol. Chem.* **1984**, *259*, 5809–5815. [PubMed]
34. Rathore, A.S.; Bhambure, R.; Ghare, V. Process analytical technology (PAT) for biopharmaceutical products. *Anal. Bioanal. Chem.* **2010**, *398*, 137–154. [CrossRef] [PubMed]
35. Luttmann, R.; Bracewell, D.G.; Cornelissen, G.; Gernaey, K.V.; Glassey, J.; Hass, V.C.; Kaiser, C.; Preusse, C.; Striedner, G.; Mandenius, C.-F. Soft sensors in bioprocessing: A status report and recommendations. *Biotechnol. J.* **2012**, *7*, 1040–1048. [CrossRef] [PubMed]
36. Wechselberger, P.; Sagmeister, P.; Herwig, C. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess Biosyst. Eng.* **2013**, *36*, 1205–1218. [CrossRef] [PubMed]
37. Lu, F.; Toh, P.C.; Burnett, I.; Li, F.; Hudson, T.; Amanullah, A.; Li, J. Automated dynamic fed-batch process and media optimization for high productivity cell culture process development. *Biotechnol. Bioeng.* **2013**, *110*, 191–205. [CrossRef] [PubMed]
38. Ozturk, S.S.; Thrift, J.C.; Blackie, J.D.; Naveh, D. Real-time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor. *Biotechnol. Bioeng.* **1997**, *53*, 372–378. [CrossRef]
39. Franze, R.; Link, T.; Takuma, S.; Takagi, Y.; Hirashima, C.; Tsuda, Y. Method for the Production of a Glycosylated Immunoglobulin. U.S. Patent US20110117087 A1, 19 May 2011. Available online: https://www.google.com/patents/US20110117087 (accessed on 6 January 2016).
40. Lenas, P.; Kitade, T.; Watanabe, H.; Honda, H.; Kobayashi, T. Adaptive fuzzy control of nutrients concentration in fed-batch culture of mammalian cells. *Cytotechnology* **1997**, *25*, 9–15. [CrossRef] [PubMed]
41. Luan, Y.; Stanek, T.C.; Drapeau, D. Restricted Glucose Feed for Animal Cell Culture. U.S. Patent US7429491. Available online: http://www.sumobrain.com/patents/us/Restricted-glucose-feed-animal-cell/US7429491.html (accessed on 6 January 2016).
42. Gagnon, M.; Hiller, G.; Luan, Y.-T.; Kittredge, A.; DeFelice, J.; Drapeau, D. High-End pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO Fed-batch cultures. *Biotechnol. Bioeng.* **2011**, *108*, 1328–1337. [CrossRef] [PubMed]
43. KWlaschin, F.; Hu, W.-S. Fedbatch culture and dynamic nutrient feeding. *Adv. Biochem. Eng. Biotechnol.* **2006**, *101*, 43–74.
44. Aehle, M.; Schaepe, S.; Kuprijanov, A.; Simutis, R.; Lübbert, A. Simple and efficient control of CHO cell cultures. *J. Biotechnol.* **2011**, *153*, 56–61. [CrossRef] [PubMed]
45. Lin, H.; Bezaire, J. Pre-Programmed Non-Feedback Controlled Continuous Feeding of Cell Cultures. Patent WO2013040444 A1, 21 March 2013.
46. Zhou, W.; Hu, W.-S. On-line characterization of a hybridoma cell culture process. *Biotechnol. Bioeng.* **1994**, *44*, 170–177. [CrossRef] [PubMed]
47. Noll, T.; Biselli, M. Dielectric spectroscopy in the cultivation of suspended and immobilized hybridoma cells. *J. Biotechnol.* **1998**, *63*, 187–198. [CrossRef]
48. JDowd, E.; Jubb, A.; Kwok, K.E.; Piret, J.M. Optimization and control of perfusion cultures using a viable cell probe and cell specific perfusion rates. *Cytotechnology* **2003**, *42*, 35–45.
49. Zhou, W.; Rehm, J.; Hu, W.-S. High viable cell concentration fed-batch cultures of hybridoma cells through on-line nutrient feeding. *Biotechnol. Bioeng.* **1995**, *46*, 579–587. [CrossRef] [PubMed]
50. Europa, A.F.; Gambhir, A.; Fu, P.-C.; Hu, W.-S. Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells. *Biotechnol. Bioeng.* **2000**, *67*, 25–34. [CrossRef]
51. Mi, L.; Feng, Q.; Li, L.; Wang, X.H. Method for Parameter Control of the Process for Culturing Serum-Suspension Free Animal Cell. Patent CN1557948 A, 29 December 2004. Available online: http://www.google.com/patents/CN1557948A (accessed on 6 January 2016).

52. Konakovsky, V.; Yagtu, A.C.; Clemens, C.; Müller, M.M.; Berger, M.; Schlatter, S.; Herwig, C. Universal Capacitance Model for Real-Time Biomass in Cell Culture. *Sensors* **2015**, *15*, 22128–22150. [CrossRef] [PubMed]
53. Dietzsch, C.; Spadiut, O.; Herwig, C. On-line multiple component analysis for efficient quantitative bioprocess development. *J. Biotechnol.* **2013**, *163*, 362–370. [CrossRef] [PubMed]
54. Lohninger, H. Datalab 3.5, A Programme for Statistical Analysis. 2000. Available online: http://datalab.epina.at.
55. MATLAB, Inc. Robust Local Regression Using Weighted Linear Least Squares in Matlab (RLOWESS). Available online: http://de.mathworks.com/help/curvefit/smoothing-data.html#bq_6ys3–3 (accessed on 6 January 2016).
56. Jobé, A.M.; Herwig, C.; Surzyn, M.; Walker, B.; Marison, I.; von Stockar, U. Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. *Biotechnol. Bioeng.* **2003**, *82*, 627–639. [CrossRef] [PubMed]
57. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
58. Glassey, J.; Gernaey, K.V.; Clemens, C.; Schulz, T.W.; Oliveira, R.; Striedner, G.; Mandenius, C.-F. Process analytical technology (PAT) for biopharmaceuticals. *Biotechnol. J.* **2011**, *6*, 369–377. [CrossRef] [PubMed]
59. Haenlein, M.; Kaplan, A.M. A Beginner's Guide to Partial Least Squares Analysis. *Underst. Stat.* **2004**, *3*, 283–297. [CrossRef]
60. Halestrap, A.P.; Price, N.T. The proton-linked monocarboxylate transporter (MCT) family: structure, function and regulation. *Biochem. J.* **1999**, *343*, 281–299. [CrossRef] [PubMed]
61. Smerilli, M.; Neureiter, M.; Wurz, S.; Haas, C.; Frühauf, S.; Fuchs, W. Direct fermentation of potato starch and potato residues to lactic acid by Geobacillus stearothermophilus under non-sterile conditions. *J. Chem. Technol. Biotechnol.* **2015**, *90*, 648–657. [CrossRef] [PubMed]
62. Kurano, N.; Leist, C.; Messi, F.; Kurano, S.; Fiechter, A. Growth behavior of Chinese hamster ovary cells in a compact loop bioreactor: 1. Effects of physical and chemical environments. *J. Biotechnol.* **1990**, *15*, 101–111. [CrossRef]
63. Ansorge, S.; Esteban, G.; Ghommidh, C.; Schmid, G. Monitoring Nutrient Limitations by Online Capacitance Measurements in Batch & Fed-batch CHO Fermentations. In *Cell Technology for Cell Products*; Smith, R., Ed.; Springer: Houten, Netherlands, 2007; pp. 723–726.
64. Yongky, A. Analysis of Central Metabolic Pathways in Cultured Mammalian Cells. Ph.D. Dissertation, University of Minnesota, Minneapolis, MN, USA, 2014.
65. Ündey, C.; Ertunç, S.; Mistretta, T.; Looze, B. Applied advanced process analytics in biopharmaceutical manufacturing: Challenges and prospects in real-time monitoring and control. *J. Process Control* **2010**, *20*, 1009–1018. [CrossRef]
66. Spadiut, O.; Rittmann, S.; Dietzsch, C.; Herwig, C. Dynamic process conditions in bioprocess development. *Eng. Life Sci.* **2013**, *13*, 88–101. [CrossRef]
67. Kantardjieff, A. Transcriptome Analysis in Mammalian Cell Culture: Applications in Process Development and Characterization. Ph.D. Dissertation, University of Minnesota, Minneapolis, MN, USA, 2009.
68. Fadok, V.A.; Bratton, D.L.; Guthrie, L.; Henson, P.M. Differential effects of apoptotic *versus* lysed cells on macrophage production of cytokines: role of proteases. *J. Immunol. (Baltim. Md.: 1950)* **2001**, *166*, 6847–6854. [CrossRef]
69. Rathmell, J.C.; Fox, C.J.; Plas, D.R.; Hammerman, P.S.; Cinalli, R.M.; Thompson, C.B. Akt-directed glucose metabolism can prevent Bax conformation change and promote growth factor-independent survival. *Mol. Cell. Biol.* **2003**, *23*, 7315–7328. [CrossRef] [PubMed]
70. Zheng, L.; Dengler, T.J.; Kluger, M.S.; Madge, L.A.; Schechner, J.S.; Maher, S.E.; Pober, J.S.; Bothwell, A.L. Cytoprotection of human umbilical vein endothelial cells against apoptosis and CTL-mediated lysis provided by caspase-resistant Bcl-2 without alterations in growth or activation responses. *J. Immunol. (Baltim. Md.: 1950)* **2000**, *164*, 4665–4671. [CrossRef]
71. Ansorge, S.; Esteban, G.; Schmid, G. On-line monitoring of responses to nutrient feed additions by multi-frequency permittivity measurements in fed-batch cultivations of CHO cells. *Cytotechnology* **2010**, *62*, 121–132. [CrossRef] [PubMed]

72. Ansorge, S.; Esteban, G.; Schmid, G. On-line monitoring of infected Sf-9 insect cell cultures by scanning permittivity measurements and comparison with off-line biovolume measurements. *Cytotechnology* **2007**, *55*, 115–124. [CrossRef] [PubMed]
73. Dabros, M.; Dennewald, D.; Currie, D.J.; Lee, M.H.; Todd, R.W.; Marison, I.W.; Stockar, U. Cole–Cole, linear and multivariate modeling of capacitance data for on-line monitoring of biomass. *Bioprocess Biosyst. Eng.* **2008**, *32*, 161–173. [CrossRef] [PubMed]
74. Cannizzaro, C.; Gügerli, R.; Marison, I.; von Stockar, U. On-line biomass monitoring of CHO perfusion culture with scanning dielectric spectroscopy. *Biotechnol. Bioeng.* **2003**, *84*, 597–610. [CrossRef] [PubMed]
75. Carvell, J.P.; Dowd, J.E. On-line Measurements and Control of Viable Cell Density in Cell Culture Manufacturing Processes using Radio-frequency Impedance. *Cytotechnology* **2006**, *50*, 35–48. [CrossRef] [PubMed]
76. De Jesus, M.J.; Bourgeois, M.; Baumgartner, G.; Tromba, P.; Jordan, D.M.; Amstutz, M.H.; Wurm, P.F.M. The Influence of pH on Cell Growth and Specific Productivity of Two CHO Cell Lines Producing Human Anti Rh D IgG. In *Animal Cell Technology: From Target to Market*; Lindner-Olsson, D.E., Chatzissavidou, M.N., Lüllau, D.E., Eds.; Springer: Houten, Netherlands, 2001; pp. 197–199.
77. Gambhir, A.; Korke, R.; Lee, J.; Fu, P.-C.; Europa, A.; Hu, W.-S. Analysis of cellular metabolism of hybridoma cells at distinct physiological states. *J. Biosci. Bioeng.* **2003**, *95*, 317–327. [CrossRef]
78. Eyal, A.M.; Starr, J.N.; Fisher, R.; Hazan, B.; Canari, R.; Witzke, D.R.; Gruber, P.R.; Kolstad, J.J. Lactic Acid Processing; Methods; Arrangements; and, Product. U.S. Patent US6320077 B1, 20 November 2001. Available online: http://www.google.com/patents/US6320077 (accessed on 6 January 2016).
79. Abdel-Rahman, M.A.; Tashiro, Y.; Sonomoto, K. Recent advances in lactic acid production by microbial fermentation processes. *Biotechnol. Adv.* **2013**, *31*, 877–902. [CrossRef] [PubMed]
80. Wu, C.; Huang, J.; Zhou, R. Progress in engineering acid stress resistance of lactic acid bacteria. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 1055–1063. [CrossRef] [PubMed]
81. Klein, T.; Heinzel, N.; Kroll, P.; Brunner, M.; Herwig, C.; Neutsch, L. Quantification of cell lysis during CHO bioprocesses: Impact on cell count, growth kinetics and productivity. *J. Biotechnol.* **2015**, *207*, 67–76. [CrossRef] [PubMed]
82. Boron, W.F. Regulation of intracellular pH. *Adv. Physiol. Educ.* **2004**, *28*, 160–179. [CrossRef] [PubMed]
83. Dechant, R.; Binda, M.; Lee, S.S.; Pelet, S.; Winderickx, J.; Peter, M. Cytosolic pH is a second messenger for glucose and regulates the PKA pathway through V-ATPase. *EMBO J.* **2010**, *29*, 2515–2526. [CrossRef] [PubMed]
84. Boyer, M.J.; Tannock, I.F. Regulation of intracellular pH in tumor cell lines: Influence of microenvironmental conditions. *Cancer Res.* **1992**, *52*, 4441–4447. [PubMed]
85. Olsnes, S.; Tønnessen, T.I.; Sandvig, K. pH-regulated anion antiport in nucleated mammalian cells. *J. Cell Biol.* **1986**, *102*, 967–971. [CrossRef] [PubMed]
86. Pilkis, S.J.; Claus, T.H. Hepatic Gluconeogenesis/Glycolysis: Regulation and Structure/Function Relationships of Substrate Cycle Enzymes. *Annu. Rev. Nutr.* **1991**, *11*, 465–515. [CrossRef] [PubMed]
87. Okar, D.A.; Wu, C.; Lange, A.J. Regulation of the regulatory enzyme, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. *Adv. Enzym. Regul.* **2004**, *44*, 123–154. [CrossRef] [PubMed]
88. Somberg, E.W.; Mehlman, M.A. Regulation of gluconeogenesis and lipogenesis. The regulation of mitochondrial pyruvate metabolism in guinea-pig liver synthesizing precursors for gluconeogenesis. *Biochem. J.* **1969**, *112*, 435–447. [CrossRef] [PubMed]
89. Mulquiney, P.J.; Kuchel, P.W. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: Equations and parameter refinement. *Biochem. J.* **1999**, *342*, 581–596. [CrossRef] [PubMed]
90. Hutkins, R.W.; Nannen, N.L. pH Homeostasis in Lactic Acid Bacteria. *J. Dairy Sci.* **1993**, *76*, 2354–2365. [CrossRef]
91. Nolan, R.P.; Lee, K. Dynamic model of CHO cell metabolism. *Metab. Eng.* **2011**, *13*, 108–124. [CrossRef] [PubMed]
92. Shivhare, M.; McCreath, G. Practical Considerations for DoE Implementation in Quality by Design. *BioProcess Int.* **2010**, 22–30. Available online: http://www.bioprocessintl.com/wp-content/uploads/bpi-content/BPI_A_100806AR03_O_98037a.pdf (accessed on 6 January 2016).

93. Carrondo, M.J.T.; Alves, P.M.; Carinhas, N.; Glassey, J.; Hesse, F.; Merten, O.-W.; Micheletti, M.; Noll, T.; Oliveira, R.; Reichl, U.; *et al.* How can measurement, monitoring, modeling and control advance cell culture in industrial biotechnology? *Biotechnol. J.* **2012**, *7*, 1522–1529. [CrossRef] [PubMed]
94. Oeggerli, A.; Eyer, K.; Heinzle, E. On-line gas analysis in animal cell cultivation: I. Control of dissolved oxygen and pH. *Biotechnol. Bioeng.* **1995**, *45*, 42–53. [CrossRef] [PubMed]
95. Åström, K.J.; Murray, R.M. Feedback Systems Web Site. 5 April 2008. Available online: http://www.cds.caltech.edu/~murray/amwiki. (accessed on 29 May 2015).
96. Dumont, G. EECE Courses Prof. Guy Dumont: Lecture Notes. Available online: http://www.phoneoximeter.org/ece-courses/eece-460/lecture-notes/ (accessed on 29 May 2015).
97. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2007.
98. Motulsky, H. *Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking*; Oxford University Press: New York, NY, USA, 2013.
99. Motulsky, H. *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*; 1. Aufl.; Oxford University Press: New York, NY, USA, 2004.
100. Aghamohseni, H.; Ohadi, K.; Spearman, M.; Krahn, N.; Moo-Young, M.; Scharer, J.M.; Butler, M.; Budman, H.M. Effects of nutrient levels and average culture pH on the glycosylation pattern of camelid-humanized monoclonal antibody. *J. Biotechnol.* **2014**, *186*, 98–109. [CrossRef] [PubMed]
101. Fan, Y.; Del Val Jimenez, I.; Müller, C.; Wagtberg Sen, J.; Rasmussen, S.K.; Kontoravdi, C.; Weilguny, D.; Andersen, M.R. Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnol. Bioeng.* **2015**, *112*, 521–535. [CrossRef] [PubMed]
102. Wechselberger, P.; Herwig, C. Model-based analysis on the relationship of signal quality to real-time extraction of information in bioprocesses. *Biotechnol. Prog.* **2012**, *28*, 265–275. [CrossRef] [PubMed]
103. Zalai, D.; Dietzsch, C.; Herwig, C. Risk-based Process Development of Biosimilars as Part of the Quality by Design Paradigm. *PDA J. Pharm. Sci. Technol. PDA* **2013**, *67*, 569–580. [CrossRef] [PubMed]
104. Gnoth, S.; Jenzsch, M.; Simutis, R.; Lübbert, A. Process Analytical Technology (PAT): Batch-to-batch reproducibility of fermentation processes by robust process operational design and control. *J. Biotechnol.* **2007**, *132*, 180–186. [CrossRef] [PubMed]
105. Li, F.; Vijayasankaran, N.; Shen, A.; Kiss, R.; Amanullah, A. Cell culture processes for monoclonal antibody production. *mAbs* **2010**, *2*, 466–477. [CrossRef] [PubMed]
106. Abu-Absi, S.F.; Yang, L.; Thompson, P.; Jiang, C.; Kandula, S.; Schilling, B.; Shukla, A.A. Defining process design space for monoclonal antibody cell culture. *Biotechnol. Bioeng.* **2010**, *106*, 894–905. [CrossRef] [PubMed]
107. Gronemeyer, P.; Ditz, R.; Strube, J. Trends in Upstream and Downstream Process Development for Antibody Manufacturing. *Bioengineering* **2014**, *1*, 188–212. [CrossRef]
108. Gao, Y.; Kipling, K.; Glassey, J.; Willis, M.; Montague, G.; Zhou, Y.; Titchener-Hooker, N.J. Application of agent-based system for bioprocess description and process improvement. *Biotechnol. Prog.* **2010**, *26*, 706–716. [CrossRef] [PubMed]
109. Green, A.; Glassey, J. Multivariate analysis of the effect of operating conditions on hybridoma cell metabolism and glycosylation of produced antibody: Effect of operating conditions on mAb glycosylation. *J. Chem. Technol. Biotechnol.* **2015**, *90*, 303–313. [CrossRef]
110. Hossler, P.; Mulukutla, B.C.; Hu, W.-S. Systems Analysis of N-Glycan Processing in Mammalian Cells. *PLoS ONE* **2007**, *2*, e713. [CrossRef] [PubMed]
111. Le, H.T.N. Mining High-Dimensional Bioprocess and Gene Expression Data for Enhanced Process Performance. Ph.D. Dissertation, University of Minnesota, Minneapolis, MN, USA, 2012.

2.3 Chapter 3

Marshall Plan Scholarship Paper

Viktor Konakovsky

Vienna University of Technology
(Vienna, Austria)

University of Minnesota
(Minneapolis, Minnesota, USA)

Advisors:

Wei Shou Hu, Ph.D.

Univ.Prof. Dipl.-Ing. Dr.techn.Christoph Herwig

2015

Improved metabolic process control by analysis of genetic clone background in mammalian cell culture

Viktor Konakovsky[1‡], Tung Le[2‡], Andrew Yongky[2], Ravali Raju[2], Nandita Vishvanathan[2], Christoph Herwig[1], Wei Shou Hu[2]

[1] Institute of Chemical Engineering, Division of Biochemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A 166-4, 1060, Vienna, Austria,
E-mails: konak004@umn.edu, christoph.herwig@tuwien.ac.at
[2] Department of Chemical Engineering and Materials Science, University of Minnesota, 421 Washington Avenue SE, Minneapolis, MN 55455-0132 USA
[‡] These Authors contributed equally to this work.
E-mails: acre@cems.umn.edu

Abstract

Currently, the analysis of mammalian cell culture bioprocesses is discriminated into a multitude of highly complex disciplines which share several common goals: to make bioprocesses safer, cheaper and easier to control. In mammalian fed-batch process control, one topic has been dominating and dictating research in this area for many years: The problem in build-up of high lactic acid concentrations [1].

Interdisciplinary discoveries in several fields are able to combine and expand knowledge in the individual technologies and address today's challenges in industry [2][3]. We present a systems approach comprising process control, modeling and cell line engineering, where the genetic background of mammalian cells and tissues, more precisely their measured transcript level and isoform composition, is linked to a highly customizable mathematical model. This model can be used to test the cell's *in-silico* behavior in various process modes and simulate the metabolic response under physiological conditions. One of the targets of modeling was to show different pathways how to reduce lactic acid build-up and to explore the model to gather new ideas to accomplish this in different ways. By incorporation of the clone's genetic identity we i) show differences in the dynamics of metabolic shifts ii) reveal potential targets to improve cell engineering and iii) provide the basis for screening suitable process and control modes for potential production cell lines *ab initio* [4].

Keywords: mammalian cell culture, systems analysis, mathematical modeling, CHO cell metabolism, process control

1. Introduction

1.1. Problem statement

In process development, clones need to be selected. The main selection criteria for this are high titers and high growth kinetics. However high growth kinetics reveal overflow metabolism, such as lactate production [5][6], which is a disadvantageous in respect to metabolic efficiency [7][8] which often corresponds to a lower product titer [9]. This metabolic efficiency can be modulated by operating in a chemical environment of high productivity and metabolic efficiency, and the tool to do this is using intelligent process control. The challenge we face is that the genetic background of every producing cell line is unique, and their response to one process scenario and its control may be very different. This leads to a different performance under the same chemical conditions and therefore unpredictability and insecurity in regard to the question which conditions may lead to optimal performance *in situ*.

As a consequence, we wish to contribute to an area of hardly any understanding, namely i) how the genetic heritage of clones, more precisely isoform composition, can affect cell metabolism, and ii) how process conditions, such as timing and intensity of shifts or pulses, can lead to a modification of metabolism. Apart from the enzymatic composition, which determines the genetic architecture, the up- or down-regulation in transcriptomic levels [10] can be considered as well to reveal how a stronger transcription may contribute to metabolic regulations. Different tissues, cell lines or high and low producers of the same product can then be analyzed under the same standards and by exposure to identical *in silico* process conditions to find i.e. that a carbon limitation strategy may be beneficial for one clone but not for another due to a general lower glucose uptake rate due to genetic background.

Usually, enough data is available way ahead of the first bioreactor run in form of transcriptomics which analyzed how well the genetic modifications on the clone were integrated. Exactly this data is highly interesting for us and was used to describe particular clone-dependent behavior in simulations. Transcriptomic data from different clones and tissues is publicly available and the model we use for these simulations is based on work published under [8][11][12].

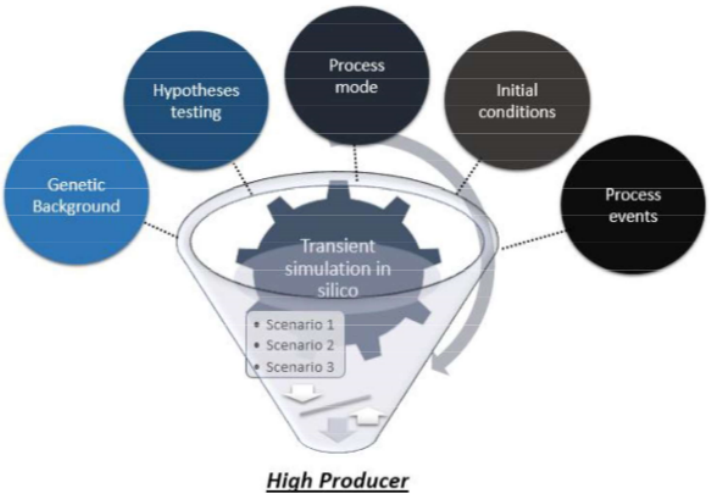


Figure 1: What makes a high producer a high producer?

Each clone reacts differently to process changes. Anticipating these changes by simulating how the real process might possibly run under the given conditions *in silico,* we believe that this tools may help us in finding clone-dependent optimal operating conditions.

1.2. State of the art

Modeling the intricate metabolic behavior of living microorganisms is a challenge which is faced best with as few assumptions a possible. Those which have to be made should be as general and reasonable as possible, optimally without empirical equations and supported by relevant literature. But the most important consideration should be to generate a model which will be able to describe the current problem appropriately. That includes the premise that the given system will fulfill the assumptions most of the time, and this might be the hardest task.

There are many published mathematical models which use mass balances to describe input and output but often a simple model means hefty assumptions which run the danger of having to be recalculated in a parameter estimation fit every single time the model is tested with new data. Such models often include very specific parameters concerning cell growth and death and changing metabolic yields and may lack transferability to other processes because recalibration may not be successful.

The general structure of current mathematical models can be described with Monod-Wyman-Changeaux [13] kinetics, where a rate decreases with a lower concentration of the limiting substrate,

$$r = r_{max} \cdot \frac{C_A}{K + C_A} \tag{1}$$

where r is the forward or backward rate, rmax the maximum possible rate, C metabolite concentration of compound A and K the concentration at which the speed of the reaction assumes exactly half maximum speed. Most rate changes, i.e. for metabolite consumption, but even growth rate are described with variants of this relationship. It does not come as a surprise that mostly batch fermentations, but also continuous cultivations can be sufficiently described with variants of this basic relationship and enjoy a great popularity because they work for their intended purpose [14][15]. While in a batch culture the substrate will eventually reach zero, and uptake must stop, in continuous cultivations a particular concentration can be held, and a certain flux can be therefore set up and ensured.

All of these rates involve a forward (fw) and a backward (bw) reaction, therefore each of the reactions r can be written as rfw and rbw [16]. Both rates go into opposite directions, and the resulting final rate determines in which direction the reaction is headed and if a substance is produced or consumed. However, there is more to it, especially if multiple intermediates are involved and the reactions are close to equilibrium. The denominator N explains allosteric regulations which happen further downstream in the equations and can contain several activators and inhibitors which affect the final outcome significantly.

$$r = \frac{r_{fw} - r_{bw}}{N} \tag{2}$$

The full term may look in a simple form as the following equation 3, however the appendix will contain a list of much more complex reactions:

$$r = \frac{r_{fw} \cdot \frac{C_A}{K} - r_{bw} \cdot \frac{C_B}{K}}{1 + \frac{C_A}{K} + \frac{C_B}{K}} \tag{3}$$

Mathematical models often use these kind of descriptions in their equations. They are our best mechanistic attempt to simplify complex effects such as cell growth in batch and fed-batch [17][18][19], but also chemostat cultivations [20][7][21]. There are plenty of models available which extend the simple models by inclusion of parts of the TCA cycle [22][23][24] to gain even more insights into the mechanisms involved, but as more parameters are introduced, complexity level increases while general validity decreases because often the parameters become bound to one clone or process strategy.

In the end, all models may have to prove themselves by showing their practical use in the intended application. Most mammalian processes today are "simple" fed-batches in which cell death is prevented as long as possible by shifts in the external conditions [25] and genetic architecture of the cells [26]. Differences between clones make a generalization of a simple model difficult and may lead to the situation where even after validation in one process the benefit of application is restricted to only this one clone, one type of medium, one type of process condition or all those points at the same time.

Therefore, oversimplified models trying to describe fed-batch cultivation processes might or might not capture enough of the most critical parameters in a dynamic fermentation. To make things even more difficult, an inevitable decrease in cell viability over time massively affects the cell's metabolism [27][28][29]. The inability to provide a valid and general metabolic model for mammalian processes is symbolized by the huge numbers of different and at the same time similar attempts to reduce a model down to its most essential parts, as summarized by Pörtner et al. [30]. This problem is even more accentuated by the striking lack of mathematical descriptors of critical factors, which influence cell metabolism in a dynamic fermentation format such as batches and fed-batches, where the lack of parameters only shows in the longer fed-batch processes where metabolism is very hard to represent accurately.

Such missing parameters may involve the dynamic changes in pH and pCO2, which are not always perfectly controlled and osmolality, which is usually not controlled at all in fed-batches [31][32][32]. Homeostasis between external and internal pH, like in lactic acid bacteria [33], is rarely mechanistically considered, although practitioners and researchers know that a shift in external pH shift leads often to a shift in internal pH within less than half an hour [34][35][36] and makes it possible to act directly on the cell's metabolism [37][38][39][40] by affecting the reaction rate of glycolytic and other enzymes [41][42][43][44].

And yet there is still an insatiable hunger for more simple, universal and accurate models. Because they can be only as good as their assumptions, the premises and expectations of the model's result should be stated in advance and compared to what is scientifically reasonable to model accurately and what is not. This concerns the following most important factors: i) predictable cell growth and ii) no delayed time effects affecting metabolism. These two conditions hold mostly true for two process formats: short-term batch fermentations and continuous cultivations, in which the chemical environment is refreshed continuously and guarantees stable growth conditions without limitations or inhibitions for a long time. The currently presented model is capable of mimicking all process modes but was intended to mainly generate knowledge from the cell's genetic architecture and isoform specific mechanism under different chemical conditions.

Pulses in a low glucose environment as well as shifts in pH can lead to different responses between clones and tissues which can be analyzed qualitatively and assessed quantitatively in experiments. Enzyme parameters, sensitivities towards inhibitory and activating substrates as well as any initial metabolic environment can be fully edited and tested under any thinkable range of simulation conditions, and provide interesting insights into the cell's response to stimuli.

A computer simulation also has a significant advantage over wet lab experiments: the environment is not influenced by any unintended process events which may modify the cells' behavior and can be tested as many times as necessary for the intended purpose by applying minute changes to the process conditions or enzymatic regulation.

Our purpose is to provide a useful tool to improve the understanding in a systems approach in several engineering disciplines [45]. We combine rational protein design, cell clone engineering and accelerated process development by exposing the cells *in silico* to a relevant challenge in several scenarios as described by Kitano [46]. We want to make clear that we do not actually design or modify real proteins in this contribution. We are merely interested in the results of the simulation after protein modification. Such a modification may involve, for instance, substitution of one or several amino acids, which can affect the protein's binding specificity to i.e. a substrate, activator or inhibitor [47][48]. We also advise to remain cautious about the model's qualitative dry lab output which may have to be adapted to the quantitative wet lab reality.

1.3. Approach

Our approach in improving current process control strategies encompasses a mechanistic explanation for differences in clone metabolism based on the clone's genetic background. In order to investigate the differences, the already existing mechanistic model developed by [7][49] was fed clone-specific transcript data [12]. The previously described model is extended by isozyme-specific parameters which are retrieved from literature databases [50][51]. Enzyme levels and ratios can be exploratory varied to find those enzymes which have the largest impact and help to explain most of the in-silico metabolic behavior of the culture. Additionally, a simulation can be conducted in several possible cultivation scenarios, in which metabolic response and clone performance can be assessed *in silico*.

The underlying model was modified and extended in the following respects:

- The most important allosteric regulations happen in glycolysis, therefore the TCA Cycle and PPP were kept constant to prevent over-complication of the given problem [49].
- Isoform fractions of not only a few selected, but all isoforms in the reaction network catalyzing the same reaction are introduced in the rate equations to account for the genetic diversity of clones. The list is incomplete and can be further expanded when feedback regulations in the form of equations become better understood and their kinetic parameters become known.
- Differences in transcript level are used to increase or decrease the calculated rate rmax, which opens up simulations of the metabolic response related to variable enzyme levels.
- External process parameters can be varied (i.e. timing and power of shifts in pH, glucose and lactic acid concentration) during simulation to elucidate their implication on cell metabolism and process control [4].

- Several optional settings, i.e. pH sensitivity of glycolytic flux and effect of low glycolytic flux on growth are implemented but not validated, and therefore open for exploration only. We hope they may help in calibrating the model to real data sets in the future.
- The code, originally written in Matlab, is now partially editable in Microsoft Excel, which makes it easier to implement changes on genetic composition, protein mechanism or process control level. Matlab is still required to run the script, but it accesses all previously modified parameters from Microsoft Excel, runs the simulations under the new conditions and saves a report file together with figures of the results.

1.4. Novelty

To the author's knowledge, a clone's full genetic information was never attempted to be included in simulations of its metabolism. We want to focus on the impact of enzyme composition involved in the glycolytic pathway. The simulations are expected to unveil clone-specific responses to dynamic perturbations in the chemical environment as seen in real processes. The individual responses after exposure to different stimuli (glucose pulse, pH shift) are believed to lead to more knowledge in the fields of directed protein design, clone engineering and radical improvements of cell-specific process parameter set-points instead of current arbitrary initial guesses as practiced when characterizing a new clone in industry.

The novelty of this approach is the combination of transcriptomic data, modeling and process control strategies in mammalian cell culture, thus enabling simulation and possibly prediction of the dynamic behavior of different industrial mammalian cell lines and tissues based solely on their genetic background with a mechanistic model. As tangible output qualitative information about an unknown clone's metabolism may be obtained in comparison with another clone such as: strong or weak producer of lactic acid under standard conditions, or good or bad growth under glucose-limited conditions. However, these statements are indicated in the simulations, not in wet lab experiments yet. Therefore, a small part of very differently behaving clones *in silico* should be used for verification of these hypotheses under standardized conditions. This may mean that an industrial strain will leave its proprietary feed and medium environment and may not have the same quantitative behavior as during simulation, but a qualitative analysis should be generally possible.

1.5. Goal

This contribution is dedicated to developing a comprehensive understanding how different isozyme compositions can affect the metabolic fate of a cell line in development and its resulting underlying process control strategy. Such a fate can encompass i.e. metabolic overflow as described by [24][52] or more efficient metabolism [53][54][55]. As simplified representation, the simulations are understood to be qualitative rather than quantitative in their output and serve as a guide to develop lean clones by identifying crucial metabolic regulations and bottlenecks as well as show the path to improvements in

clone engineering and process control in a very early screening of clones. As very practical outcome, this work can be immediately used to screen large data sets of yet untested next-generation high producers [10][12][56] to compare their performance in several synthetic benchmarking scenarios [57].

1.6. Workflow

These process parameters include, but are not limited to pH, substrate and metabolite concentrations, online set-point controls as well as the type of feeding profile [58][59]. Our simplified mechanistic model for glycolysis under different levels and types of isoforms is expected to illuminate the intricate allosteric regulations and dynamic effects of genetically very differently tuned industrial strains and empower process engineers to select the best-suited process control strategy by design.

This contribution is expected to weld the invisible band between cell characteristics on the genomic level, description thereof by mathematical modeling, and final implementation in the bioreactor with state-of-the-art process control. We believe that a holistic perspective is not only capable but also necessary to solve otherwise isolated and therefore challenging problems with the most simple and intuitive techniques from either technology. Our vision is the description of a systematic methodology which may serve as a prototype to describe more complex problems than lactic acid metabolism in the future by bridging several disciplines in bioprocess development [3][60][61].

2. Materials and Methods

2.1. Data set

The available dataset includes different tissues and cell lines and was collected and kindly provided by Nandita Vishvanathan and Ravali Raju [12]. Transcriptomic information was obtained by RNA-seq and microarray measurements of different specimen after 4 and 7 days in culture [10].

2.2. Isoform influence

The levels of each class were statistically analyzed for the min, mean and max values of all clones. The weight of each isoform was calculated for ratios. If only one isoform exists, the ratio is 1, or 100%. An example is given below. The rate multiplies by this ratio and all involved isoforms are added up to determine the final rate. In a simplified example, where rHK1 would weight 0,77 and rHK2 0,21, the final rate calculation would follow the form r = rHK1 + rHK2. The differences between both enzymes is their specificity towards glucose, which is closer described in the excel sheet containing all isoform-specific parameters, can be fully modified and serves as direct input for the simulations.

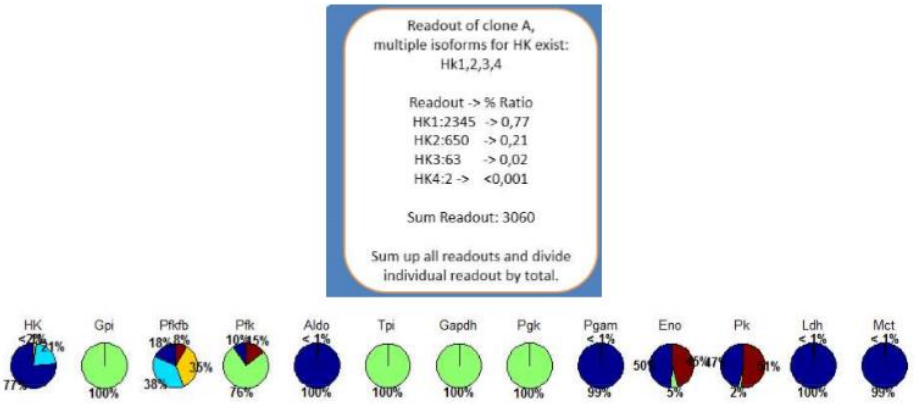


Figure 2: Isoform ratios. Every pie represents a clone's genetic isoform composition in percent.

2.3. Transcript level normalization

The data was normalized by dividing the genes' absolute intensity of the readout by the level of a house-keeping gene, whose abundance is assumed to be evenly distributed among the cell lines under investigation and in this case Gapdh [62][63]. The factor (how much more of one particular isoform exists than average) can be then plotted for each isoform and each clone.

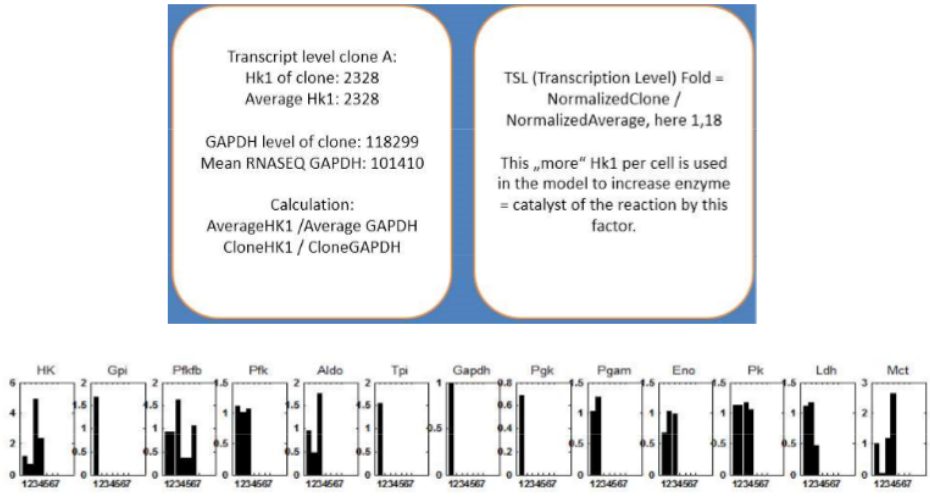


Figure 3: Transcript levels. Every bar represents a clone's fold-change compared to the mean transcript level of all clones of this isoform.

Finally, the data is automatically read out by Matlab and used for simulation conditions which can be closer specified in the excel sheet containing the starting parameters.

3. Results

3.1. Scenarios

The following scenarios were compiled to test the given dataset under all kinds of different process conditions and to show the difference in metabolic response.

- Impact of clone architecture to model behavior (considering isoform and transcript level)
- Continuous feeding with pH shifts and glucose pulses
- Varying power, direction and timing of pH shifts
- Exploring interesting targets for protein design (*in-silico* modification of substrate/inhibitor binding specificity of i.e. LDH, HK…)

3.2. Simulation of continuous behavior

The following experiments were performed to show the effect of isozyme composition and transcript level on metabolic behavior between different clones.

First, the scenario entails a glucose shift from a strongly limiting concentration of 0,1 mM glucose to 5,4 mM (roughly 1 g/L) glucose. On the left, glucose concentration in mM is shown, on the right the specific glucose consumption rate of the cells in mM/h. As can be seen in this very simple example, a change in glucose concentration leads to an instant increase in the consumption rate.

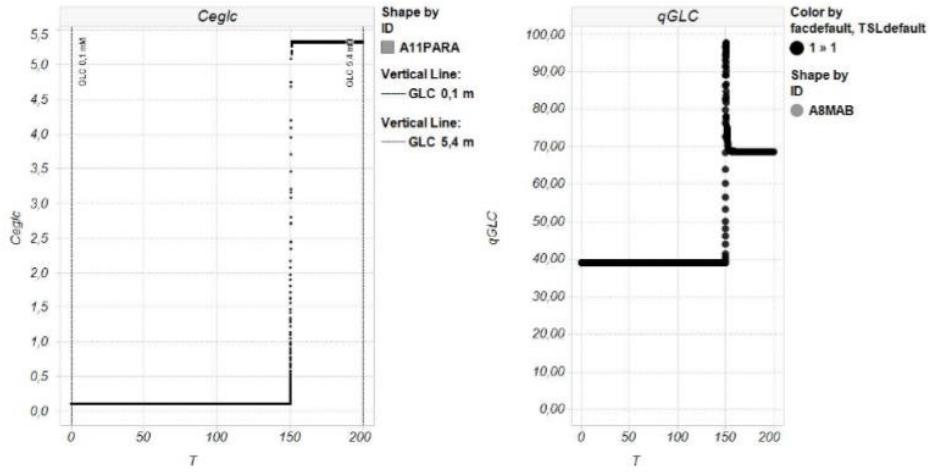


Figure 4: Simple glucose pulse experiment, Glucose concentration in mM and glucose flux in mM/h

The scenario is repeated but this time isozyme ratios and transcript levels of the particular clone are considered. The following graph is used as visual reference about the cell's isoform structure and genetic background.

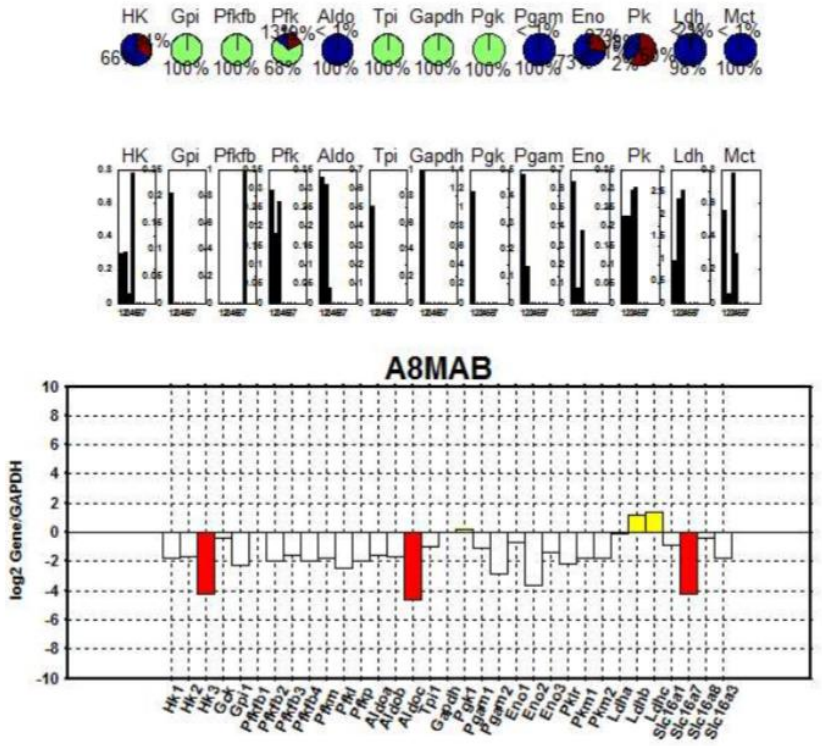


Figure 5: Simple glucose pulse experiment taking the clone's isoform composition into account - Genetic background of clone A8MAB. The bottom bars indicate the biggest three upregulated and smallest three downregulated enzymes, in respect to GAPDH, a housekeeping gene to which transcript levels were scaled.

Indicated as pie are the cell's major isoforms in Matlab's standard colors for categorical data. If only one isoform is available, only one color fills the pie, otherwise more colors are used. The bar charts signify the transcript level (or fold change) of the reaction rate for each isozyme normalized to a housekeeping gene. The y-axis is the numeric fold change and the x-axis the isoform in question in regard to the mean level of this isoform which was determined from all clones. Finally another bar chart represents by which standard the cell line's transcript levels were normalized on log2-scale for a full view of present isozymes compared to the mean and the highest and lowest three genes highlighted in yellow resp. red. A very detailed description of the calculation and a concrete example is given in the appendix and in the formula below. Our contribution does not only include a few selected, but all possible isoforms, their ratio in any given clone, and the amount of transcript readout relative to the mean readout.

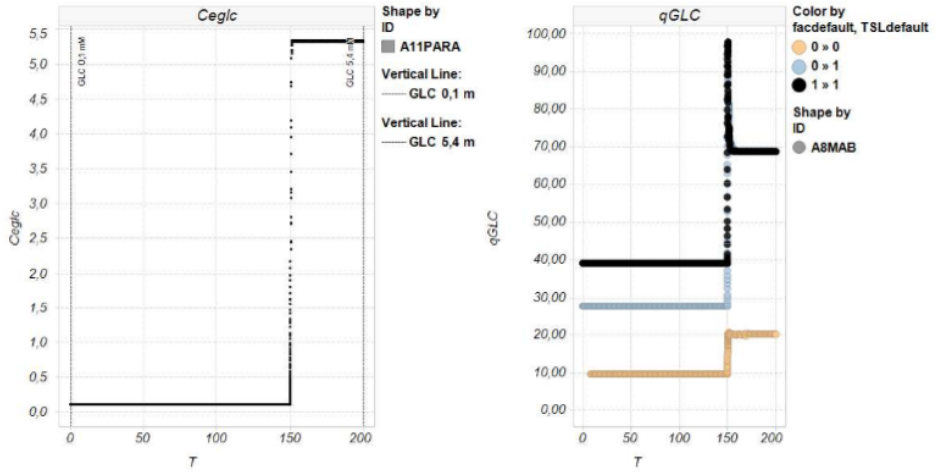


Figure 6: The decision to include isoform composition (blue) or additionally also transcript level (orange) changes the results of the simulations for A8MAB. The default model is black. Glucose in mM, glycolytic flux in mM/h.

The ratio of isozymes is incorporated in the rate equations by modifying the forward reaction with TSL (transcript level), which represents more of the enzyme in question, fractions of the individual rate equations for each isozyme or both. Facdefault 1 represents a default ratio where no isozymes are considered but one dominant form, where the dominant form is the same as in [49]. TSLdefault 1 neglects the effect of increased or decreased enzyme levels and the rate equations are multiplied by a factor of 1. An example of the modified rate equations is given below. Prior experiments were conducted by modifying the isozyme ratio of HK, PFK and PK and transcript levels of PFKFB in a clone, for which the mathematical model was calibrated for [7][11][49].

This contribution considers not only all available isoforms in their exact composition as calculated from the provided transcriptomic data [10], but also the transcriptomic level of any clone in this data sheet. As a consequence, the model may have to be adapted to yield reasonable intracellular concentrations for all clones under investigation. The general form of how transcript levels are incorporated into the existing model is described in the following formula:

$$r_{fw1} = \frac{r_{fw1} \cdot TSL_1 - r_{bw1}}{N_1} \tag{4}$$

$$r_{fw2} = \frac{r_{fw2} \cdot TSL_2 - r_{bw2}}{N_2} \tag{5}$$

$$r = r_{fw1} \cdot Fac_1 + r_{fw2} \cdot Fac_2 \tag{6}$$

The investigated clone A8MAB has a decreased amount of nearly all isozymes and GAPDH, resulting in an absolute decrease of all rate equation modifiers except those for LDHB and LDHC which are increased. Reasons for such low transcript levels may be genetic engineering because of an interest to drive the clone's metabolism towards a metabolism of reduced lactic acid formation.

Plotting A8MAB versus another clone, A11PARA, under the same conditions yielded very large differences in the genetic architecture of some particular genes responsible for increased glycolytic activity, namely PFKFB3 [64] , PFKL [65], and both HK 1 and 2 [66][67] which have a high affinity to glucose even at low glucose levels.

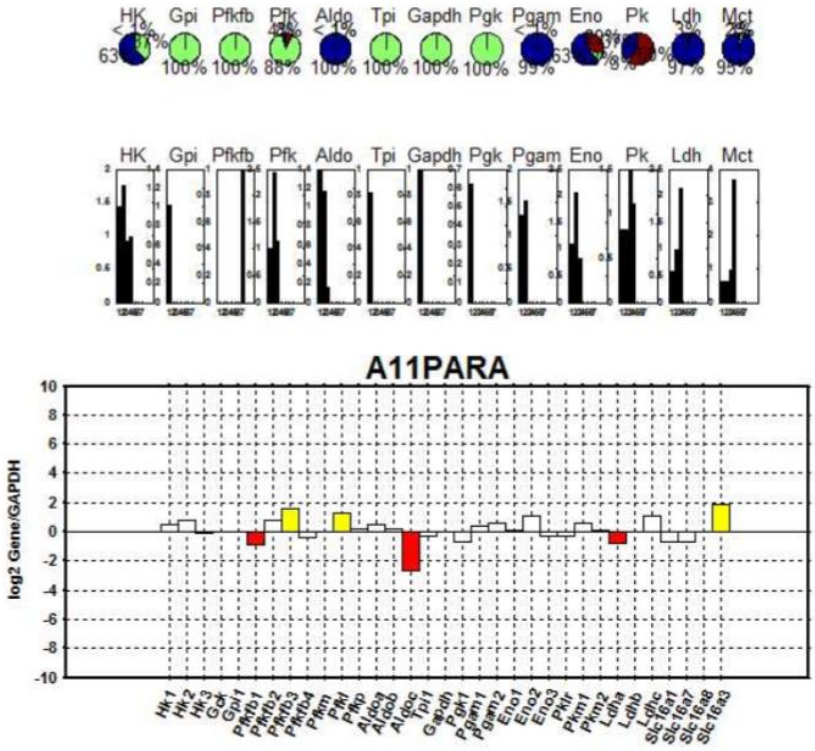


Figure 7: Genetic background of clone A11PARA.

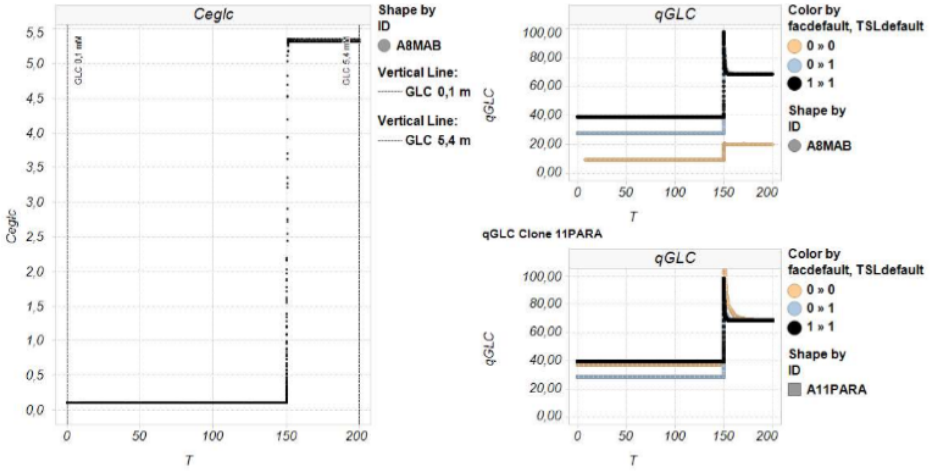


Figure 8: Comparison of metabolic response between two different clones (A8MAB: top, A11PARA: bottom) when isoform composition is taken into account (blue) and when transcript levels are taken into account additionally (orange). The black line represents default behavior. Glucose in mM, glycolytic flux in mM/h.

Simulation results and comparison with the first cell line reveals that glycolytic flux is higher in this clone. They also show that it matters whether only isoform composition or additionally the amount of enzyme according to the measured transcript readouts is considered or not. For further investigations, simulations considering both isoform composition as well as transcript levels were taken into account with the exception of PFKFB3 which stays exactly the same for all clones. The reason for not including changes of this important regulatory enzyme is that an important assumption was made in the early versions of this model and requires more attention.

KbP, an early transcript level-like parameter describing the ratio between forward and backward reaction for the F6P –> F2,6bP node was set to 10, and acting on this value with two more factors could lead to a situation where more F2,6bP is consumed than is available because the backward reaction rate becomes larger than the forward rate even when F2,6bP is already zero. This problem arises because KbP must be known, at least approximately. It can assume values between 0,4-710 [49] between tissues and clones. An excel sheet is provided to better explain the problem. However, if the exact KbP is known or well estimated, it can be set individually to the desired value and a change of PFKFB3 can be simulated for any individual cell line instead of making this assumption for all clones.

Conclusively, as a first step, transcriptomic information was integrated into the mathematical model under the following considerations:

- The original model without implementing isoform composition or process events can be used to test metabolic response to a different chemical environment
- By introducing isoform ratios, metabolic changes between all clones can be observed, i.e. under glucose limited conditions. However, isoform composition can potentially be very similar and result in almost no difference between clones in the simulation.
- Transcript levels of enzymes may yield more interesting results since some clones have more of the same isoform than others. This quantity is difficult to represent for two reasons: first, a higher transcript readout does not automatically mean a higher enzyme concentration and vice versa. On the other hand, not including such an important effector of rate speed may lead to oversimplification of otherwise very different transcript levels between clones.
- Calibration – some assumptions may need revision, such as the transcript level and basal reaction rate of PFKFB3 or other enzymes, which are modified by both a transcript level and isoform composition factor. If clones should be tested under the same standardized conditions, concentrations of any intermediates may not turn negative or reach unphysiological concentrations in the simulations.

We have shown that the way to analyze transcriptomic integration into a mathematical model may affect the results of the simulations. We decided to use both isoform composition and transcript levels in further simulations to reflect both effects in distinct scenarios as accurate as possible.

3.3. Influence of pH shift

The basic model involves a transporter equation for lactic acid which explains how chemical equilibrium between cytosolic pH and cytosolic lactic acid and extracellular pH and extracellular lactic acid can drive lactic acid production or consumption. To show the mechanism in action, a pH shift was simulated for both clones and its impact on the lactic acid profile and viable cell concentration shown. To gain the most insights from the experiment, the experiment shows cell metabolism i) at a low glucose concentration and high pH, ii) at a high glucose concentration but low pH, and iii) at a high glucose concentration and high pH.

While in the simulations both cell lines start alike, their behavior during a pH shift diverges due to a difference in genetic composition. After the first pH shift by -0,2 units, lactic acid production is reduced for both clones. But while clone A11PARA continues producing lactic acid, A8MAB begins consumption. After glucose is restored to a higher level, lactic acid production sets in again due to an increase in glycolytic flux, albeit with different specific production rates. A final pH shift back in the same order of magnitude leads to increased lactic acid production once more.

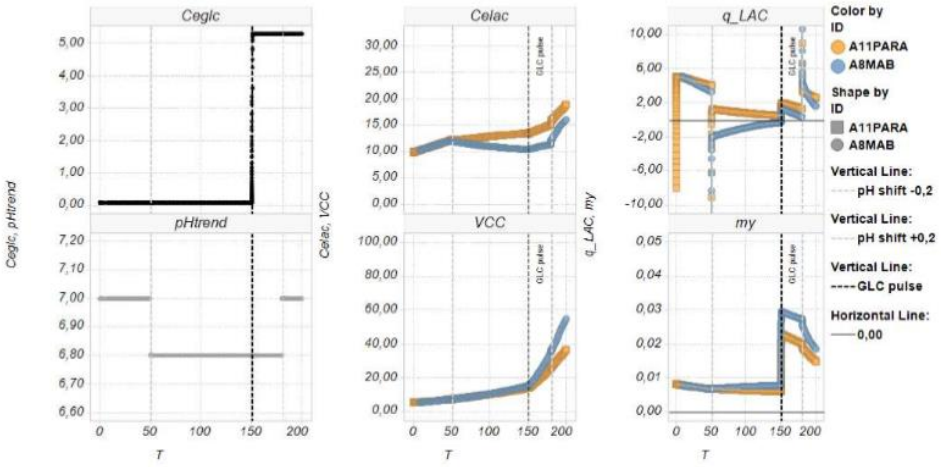


Figure 9: Influence of pH shift. Two clones are compared in their metabolic response of pH shift: A8MAB (blue) and A11PARA (orange). External glucose and lactic acid concentration in mM, VCC in E5 cells/mL, lactic acid flux in mM/h, growth rate in h-1.

The simulation results indicate that A11PARA may be a stronger consumer of substrate and stronger producer of metabolites in the form of lactic acid, whereas A8MAB might consume less substrate but finally be less inhibited by lactic acid as seen in the total number of viable cells at the end of the simulation.

Growth is explained by the standard model using the following relationship between glucose availability and lactic acid inhibition, which was used in a continuous culture experiment [7] but may have to be modified to better fit the scenario's reality.

$$\mu_{standard} = 0.055 \frac{C_{GLC}}{C_{GLC} + 0{,}3} \cdot \frac{144}{{C_{LAC}}^2 + 144} \tag{7}$$

A pH shift may lead to a reduction in lactic acid production, but it may have also other less desirable effects, such as reduced cell growth. An external pH may shift the internal pH by a few units depending slightly on cell clone [36], which then may act directly on the cell's rate limiting enzymes in glycolysis, such as HK [42][68] and PFK [69][70].

3.4. Influence of pHe on pHi

How much the external pH correlates with internal pH can be assessed by one-time calibration. However, this is somewhat problematic for a screening tool with different cell lines, because finally, the calibration curve may vary slightly between clones. Still, we believe that assuming a constant pH gradient during pH shifts is farther from reality than using an experimental reference conducted with mammalian cells and therefore enable this reversible option for our model simulations. Allemain et al. have measured exactly this relationship in fibroblasts and found that different clones have a relatively constant slope.

An increase of roughly 0,55-0,6 units pHi per 1 unit increase in pHe [36] can give an estimate how much other cell lines might differ from each other. The slope is relateively constant, but the intercept may be different. Since we are interested mainly in the slope, which is the relationship between pHi and pHe, we find this less problematic for our purpose. Therefore, even though clone to clone variance is possible, this relationship may hold fairly true for most simulations and therefore be used to investigate how metabolism can change over time. The relevance of this finding may help modeling cells even more accurately in a pH-dependent manner and is explained in more detail below.

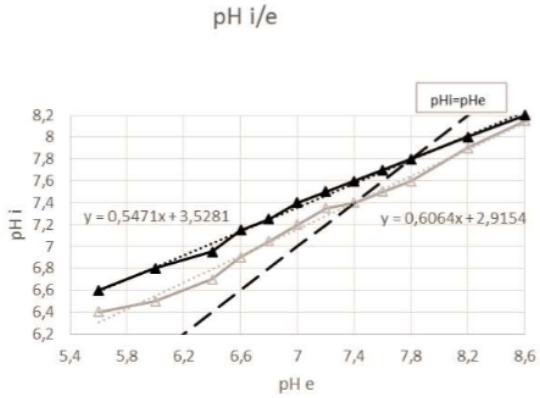


Figure 10: Relationship between pHi and pHe as published by Allemain [36]**. The data was used to calculate slope and intercept. Two different fibroblasts were compared in culture, one paternal line (black) and a mutant lacking Na+/H+ exchange (white). For our purposes, the slope from the wild type (black) was considered.**

The question is how pH change could be used to alter the rate in our glycolysis model. Mulquiney [44] modeled the impact on the forward reaction of HK by a bell function,

$$pH\ factor\ = \frac{1}{1 + \frac{10^{-pH}}{10^{-pK1}} + \frac{10^{-pK2}}{10^{-pH}}} \quad (8)$$

where pK1 is 7,02 and pK2 9,0. At pH 7,3, the fold increase factor is 0,65. Because the maximum flux in the bell function is around pH 8, but the model flux used pHi 7,3, normalization was required. Changing pH between 6,6 and 8,0 now corresponds to a change in glycolytic flux by a factor between 0,58 – 1,3.

Comparison with literature, where the impact of external pH on glycolytic flux was investigated [39] reveals that a flux change is reasonable as the glycolytic rate in experiments decreased to little more than a third of its typical rate in case of a pHe shift from 7,2 to 6,8 and increased after pHe shift from 7,2 to 7,8 by more than two-fold. However, these experiments were conducted in the first two days of a batch culture, where specific glucose uptake is subject to very large changes. Our calculated "squeeze" of glycolytic flux rate is rather conservative in comparison by assuming that a pHe change of 0,2 units would not increase the glycolytic flux by more than 0,1 pHi units. However, often batch or even fed-batch cultures start at a much higher pH because of the anticipated lactic acid formation, thus a total pH change from 7,4 to 6,8 is possible which is in the experimentally observed flux change range.

Conclusively, this adds an aspect to the simulation without exaggerating its effects in the range pHi 6,6 – 8 which corresponds to a pHe of 5,6-8,2 as depicted in the pHi/pHe relationship above with pHi of 7,3 as norm. As pHi does not change massively, but slightly, we hope to point out how cultivations at different pH or long-time pH shifts may lead to a different metabolic profile of the culture, especially if the explored difference in pH is very large and must be taken into account.

Table 1. Relationship between pHi and glycolytic flux. At pHi 7,3 there is no effect on glycolytic flux

pHi	*final pH factor influencing flux*
6,6	*0,42*
6,8	0,58
6,9	*0,66*
7	0,75
7,1	*0,84*
7,2	0,92
7,3	*0,99*
7,4	1,07
7,5	*1,13*
7,6	1,18
7,8	*1,26*
8	1,28

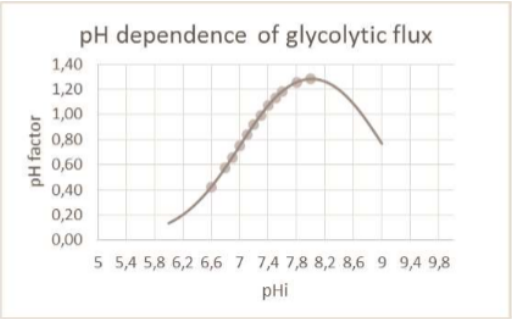


Figure 11: Relationship between pHi and glycolytic flux in modeled erythrocytes. The pH factor is a dimensionless multiplier of the forward reaction of pH sensitive enzymes, such as HK or PFK [39].

The pH factor as presented by Mulquiney's bell-curved relationship [39] between metabolic flux and pHi was used to multiply the forward rates of HK and PFK, the rate-limiting steps in glycolysis, by assuming that pHi has the same effect on both two enzymes. Finally, a decrease in glycolytic flux is considered to influence growth rate negatively and is represented in form of the following modified growth equation of the original model:

$$\mu_{custom} = 0.055 \cdot pHfactor \cdot \frac{C_{GLC}}{C_{GLC} + 0{,}3} \cdot \frac{144}{{C_{LAC}}^2 + 144} + death\ constant \tag{9}$$

At least one term in the growth equation should consider a decrease in cell growth to better reflect fed-batch behavior in which cell viability decreases over time. It may depend on many other factors than the linear functions with which it is represented and is set to zero in these simulations but may be used in fed-batch simulations where cell concentrations decrease after some days of cultivation.

The additional coefficients pH factor can be used if pH or excessive lactic acid is believed to have an effect on the specific growth rate which could result in slower growth as experienced in practice. Osman et al. have probed the growth behavior of cell culture after pH shifts in a wide range between 6,5 and 9, and came to the conclusion that cells grow best between pH 7,3-7,5 and antibody productivity was highest around pH 7,1 [71]. Several reasons for an increase in glycolytic flux at higher pH were proposed, including an alteration in membrane potential which facilitates glucose import, more lactic acid production to counter alkaline pHe to maintain pHi [72]. If for the current problem statement pH is controlled perfectly, or for other reasons have no effect on growth, the standard growth equation (7) can still be used instead. For our following next simulations, pH was considered.

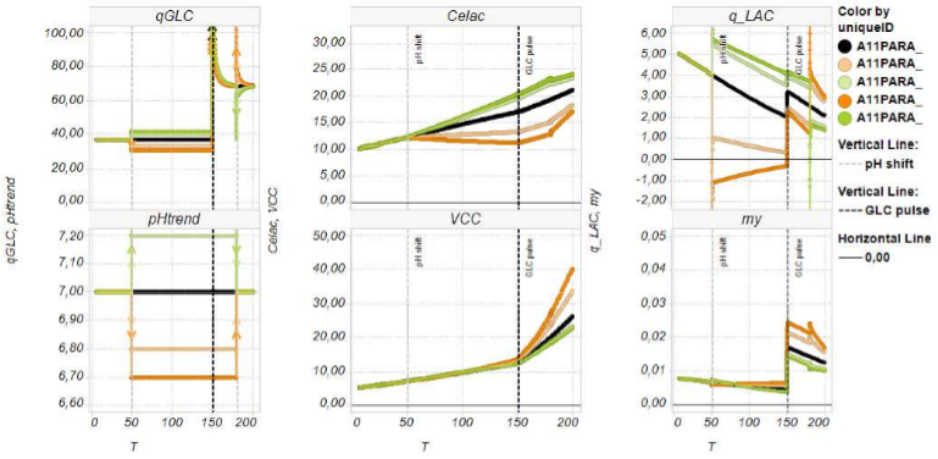


Figure 12: Effect of pH shifts of different intensity (shade) and direction (green: + 0,2 or + 0,3 units, orange: -0,2 or -0,3 units). Lactic acid in mM, VCC in E5 cells/mL, glycolytic and lactic acid flux in mM/h, growth rate in h-1. Shift in pH and glucose concentration (from 0,1mM to 5,3 mM) is added as grey and black dashed lines on all plots.

The simulations show that an acidic pH shift reduces the specific glucose and growth rate slightly (from pH 7 to 6,8), which does not have a big impact in limitation, except that lactic acid levels are kept low due to the reduced lactic acid production rate, which, depending on the intensity of the pH shift, can turn to lactic acid consumption. Comparing these simulations results to literature, Kurano et al. observed a decrease of growth rate, specific glucose uptake rate and specific lactic acid production rate as an effect of pH [73], where growth peaked at a pH of 7,6. Osman et al. [71] produced similar results in fed batch culture, where he observed peak cell growth at pH 7,3-7,5 but peak volumetric antibody production at pH 7, indicating a strong link to metabolic efficiency.

Our results indicate that the major process technological advantage of this cultivation strategy may lie in the ability to keep viability high in the form of an initially lower, but constant growth rate [74]. Upon glucose pulse, all specific glucose uptake rates approach their physiological maximum because the enzymes work at their full capacity. A pH shift back to norm conditions perturbs the equilibrium between inner and outer lactic acid and H+ concentration once more and leads to adaptation of the specific lactic acid production rate.

We wish to remind the reader that this is a scenario in which the cell line in question is tested for its individual metabolic response. By providing a simulation including both a pH shift down and shift up at two distinct glucose concentrations, we are able to make a statement about metabolism in both limited

and non-limited process conditions. The results could then be used as guidance whether the clone would be suitable under limited conditions or not and how much a shift in pH might affect cells in a limited or non-limited environment in order to find the most efficient metabolic behavior.

3.5. Timing and intensity of pH shift

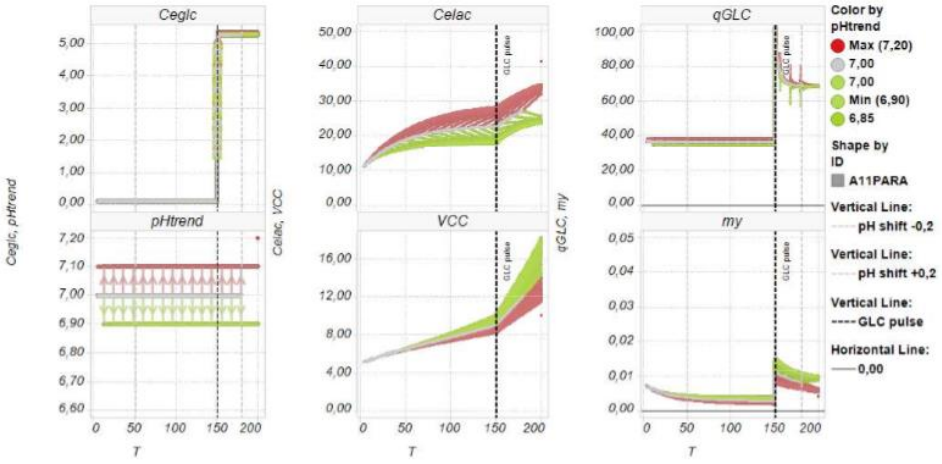


Figure 13: Effect of different timing of pH shifts of different direction (green: pH shift - 0,1 units, red: pH shift + 0,1 units). Only one pH shift was performed per simulation, the time window for pH shift exploration was 12h and all results are plotted in one window. Glucose and lactic acid in mM, VCC in E5 cells/mL, glycolytic flux in mM/h, growth rate in h-1. Shift of glucose concentration is added as dashed line on all plots.

Simulations show an effect in the timing and intensity of pH shifts, which can be as important as the decision how to hold a particular glucose concentration. The scenario was set to simulate a unique pH shift of 0,1 after 12h, and the scenario was repeated by delaying the pH shift another 12h without returning to the initial pH, but a glucose pulse at the end to determine the effect of non-limited behavior after some time of incubation in limited glucose conditions. As a result, every scenario has one shift in its runtime, the direction of the shift is distinguished by color (green: shift down 0,1 pH units, red: shift up 0,1 pH units), and all of them are plotted together above.

A small shift of only 0,1 pH units resulted in a transient lactic acid difference which caused the growth rate to differ by as much as 50% upon glucose shift just because of reduced lactic acid concentration. Such a scenario, with a less severe limitation as in this example, could become relevant during seed train

expansion where cells are transferred into fresh medium and experience a sudden growth impulse after each stage and where the optimal timing between transfers would be of interest to modification.

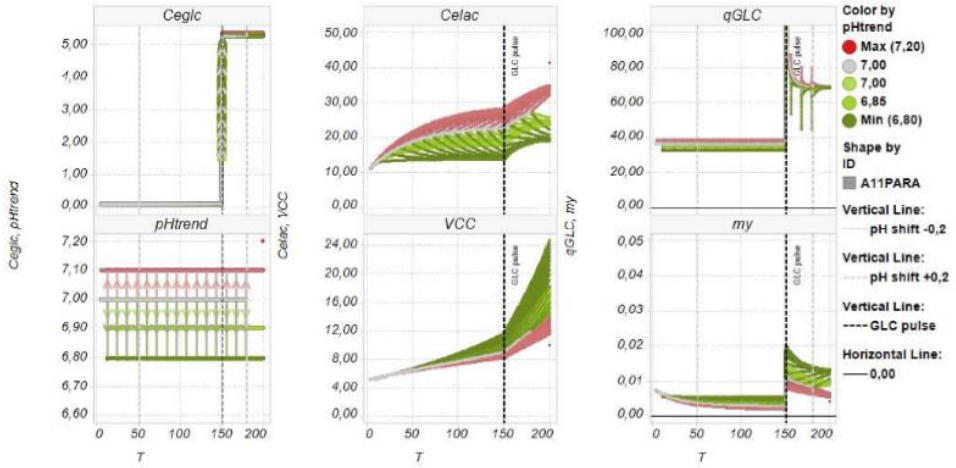


Figure 14: Effect of different timing of pH shifts of different direction and intensity (green: pH shift - 0,1 and - 0,2 units, red: +0,1 units). Only one pH shift was performed per simulation, the time window for pH shift exploration was 12h. Glucose and lactic acid in mM, VCC in E5 cells/mL, glycolytic flux in mM/h, growth rate in h-1. Shift of glucose concentration is added as dashed line on all plots.

The shift experiment was repeated for -0,2 units in pH and differences in the speed at which lactic acid was consumed could be observed. The later the pH shift happened, the more lactic acid was already present in the extracellular environment and the more cells could use the larger gradient to increase their lactic acid consumption rate as described in literature [75]. Comparison between 4 pH ranges indicates that although both glycolytic rate and growth rate is slightly reduced at lower pH, the growth rate benefits more from the reduction in lactic acid than it suffers under a pH shift and assumes up to two fold larger values after the glucose pulse than cells which were cultivated at a higher pH and have a larger lactic acid inhibition by then.

A pulse-wise dosage of glucose results in glucose concentrations in a range way above the physiological level does not only harbor the risk of glucose starvation as a result of manual feeding, i.e. when operators make a mistake in administering glucose or add it too late or at a too low dosage [76], but the metabolic profile is expected to be more growth-inhibiting due to lactic acid accumulation compared to a continuous, 'feed what you need' strategy.

Such a continuous feeding strategy is possible only with appropriate online-capable equipment such as off-gas analysis [77][78][79], capacitance probe [80][81][82] or combinations of those and other methods [60], [83]–[86] and often glucose concentration comes close to the Km range of glucose transporters [87]. Such a mode of operation may be not only beneficial in terms of keeping viability higher [88] than in a standard process, but also improve the stability of the chemical environment (external glucose and lactic acid concentrations), which may stabilize pH control [89] and reduce process events, and finally lead to a higher titer [90], better product quality [25][91] or a synergetic mixture of all.

The point in not using i.e. a black-box model is, that different cells react differently to process control strategies, but especially so under glucose limiting conditions. Most mammalian production cell lines are difficult to describe mechanistically when the culture operates always way above the physiological range in a non-limited environment because often the resulting parameters may change when the model is tested on other clones. Our approach is to describe growth mechanistically under conditions where real enzyme parameters are used in rate equations and a different composition of isoforms has an impact on metabolic flux. For instance, glucose influx can be described as a function of glucose concentration, and depending on which isoform composition is available in the clone and how much is present, the resulting rate will differ greatly among clones and some examples will be provided in the next section to explain this in detail.

3.6. Explorative Protein engineering

This model has fully explorative kinetic parameters which can be changed any time to better understand the system. To do this, the provided excel sheet can be used to change i.e. affinities to substrate. By getting to know the regulations first-handed in simple and short simulations, parameters may be also changed to see their impact on the process scenario result.

The following parameters were changed to simulate how a modulation of isoform could have a different outcome in the simulation. Differences in isoform behavior (terms in rate equations) and kinetic properties (numeric parameters such as substrate affinity) are not fully explored for all isoforms yet. For instance, there are 4 isoforms for LDH, but often important information is missing, such as the exact composition of the tetramer fraction of the isozyme. These tetramers can then i.e. influence the parameter for product inhibition, and affect lactic acid behavior of the clone [92]. In lieu of concrete isoform data, the following approximations may be as informative instead: a ten-fold increase of an available “norm”, which is determined from literature usually not considering isoforms, is increased or decreased to simulate how LDH isoforms having those fold changes in important parameters would change metabolism. The parameters which were modified were the Km for pyruvate and LDH enzyme level.

Table 2. Possible targets for enzyme modifications

Enzyme	*Target*	*Color code*		
		Black	Blue	Green
LDH	Km Pyruvate	Current norm (0,2mM)	0,02mM	2mM
LDH	Enzyme level	Current norm (0,00343)	0,000343 (10-fold decrease)	0,0343 (10-fold increase)
HK	Km Glucose	Mixture depends on selected clone	0,001mM	0,1mM
HK	Enzyme level	Current norm (0,047)	(0,0047) (10-fold decrease)	(0,47) (10-fold increase)
PK	Km F1,6bP	Mixture depends on selected clone	0,01 mM	1000 mM (virtually no binding to activator)
PFK	Enzyme level	Current norm (0,0002)	0,00002 (10-fold decrease)	0,002 (10-fold increase)

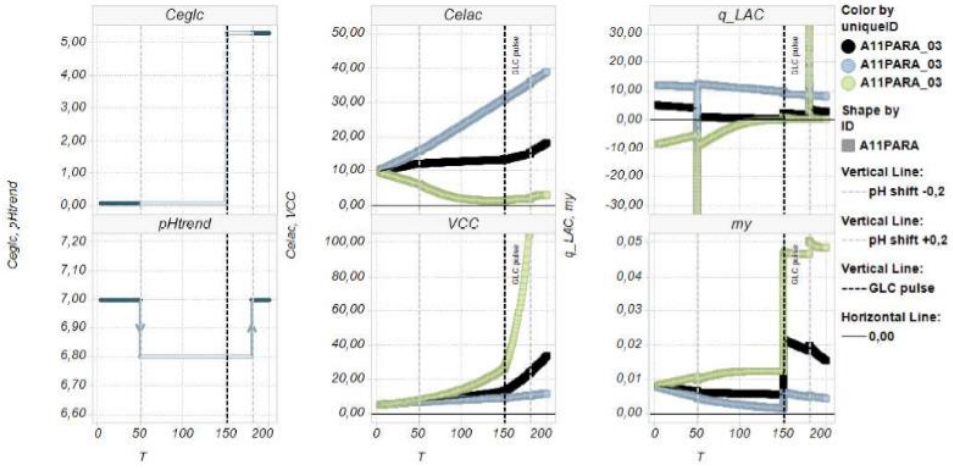


Figure 15: Effect of changed kinetic parameters in LDH: substrate specificity. Glucose and lactic acid in mM, VCC in E5 cells/mL, lactic acid flux in mM/h, growth rate in h-1. Shift of pH and glucose concentration is added as grey and black dashed lines on all plots.

Upon a change in Km of LDH to pyruvate by a factor of 10, lactic acid consumption behavior was markedly affected. A tenfold increase in Km of pyruvate resulted in immediate lactic acid consumption which not even a bolus shift or pH shifts could noticeably affect. As a result, our modeled growth rate

was at first limited, while glucose was limited, but after the pulse shifted to its maximum value as no inhibitory external lactic acid was present at all. On the contrary, decreasing Km of pyruvate tenfold resulted in a constant lactic acid production.

Conclusively, protein engineering of LDH may result in low lactic acid levels in newly engineered clones which may have a more efficient metabolism and be easier to cultivate in culture because they would by themselves not produce as much lactic acid as the paternal line. However, in practice this could also mean that their ability to grow might be impaired. This is why it is important to understand which parameters influence metabolite formation and which influence growth. By setting up for instance a much higher pH, a low growth rate might be increased and the clone would potentially still have a low lactic acid production rate.

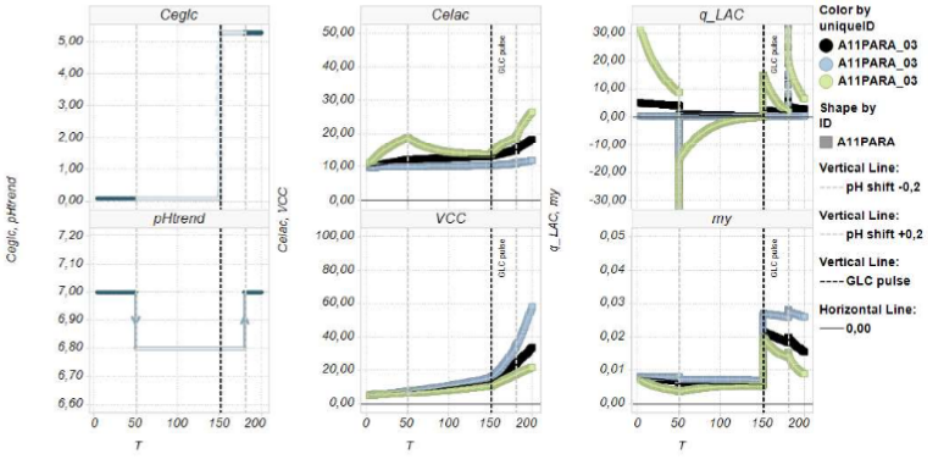


Figure 16: Effect of changed kinetic parameters in LDH: enzyme transcript level. Glucose and lactic acid in mM, VCC in E5 cells/mL, lactic acid flux in mM/h, growth rate in h-1. Shift of pH and glucose concentration is added as grey and black dashed lines on all plots.

Modifying the transcript level of LDH has also a direct metabolic effect. With more LDH, lactic acid production and consumption are both stronger pronounced in the given process conditions, while with less LDH it takes longer to reach equilibrium.

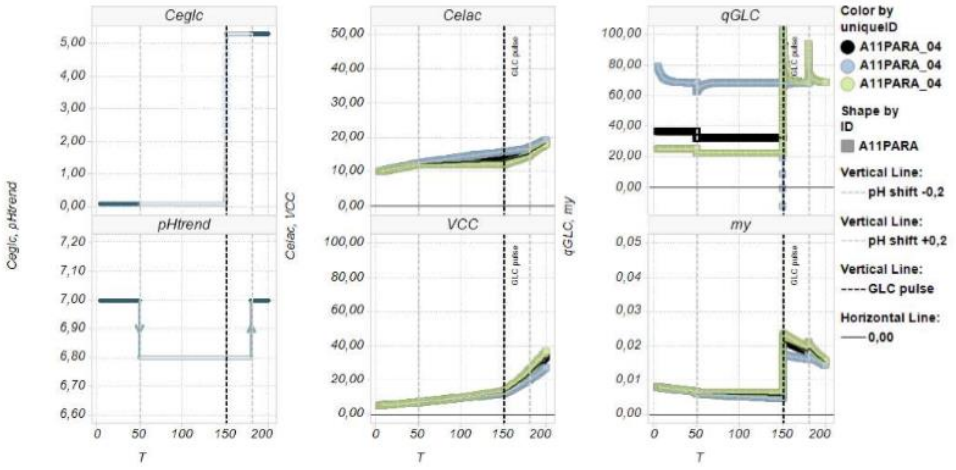


Figure 17: Simulation of pure isoforms in HK. Black: default clone, blue: Km glucose artificially set to 0,001mM (closest to HK III isoform), green: Km glucose artificially set to 0,1 mM (closest to HK I isoform).

Applying the same artificial changes to HK results in an interesting artifact which is based on the simulation scenario: increasing specificity towards substrate (low Km, blue) results in maximum glycolytic flux, but low lactic acid production rate because of flux channeling to the TCA. The growth rate meanwhile is rather low because the model's growth is described by glucose concentration and lactic acid inhibition, but not glycolytic flux itself. The flux, but not growth rate of the lowest substrate specificity to glucose (high Km, green) is lowest in this experiment. All fluxes rise when sufficient levels of glucose are available again, because the concentration is far above Km in this scenario.

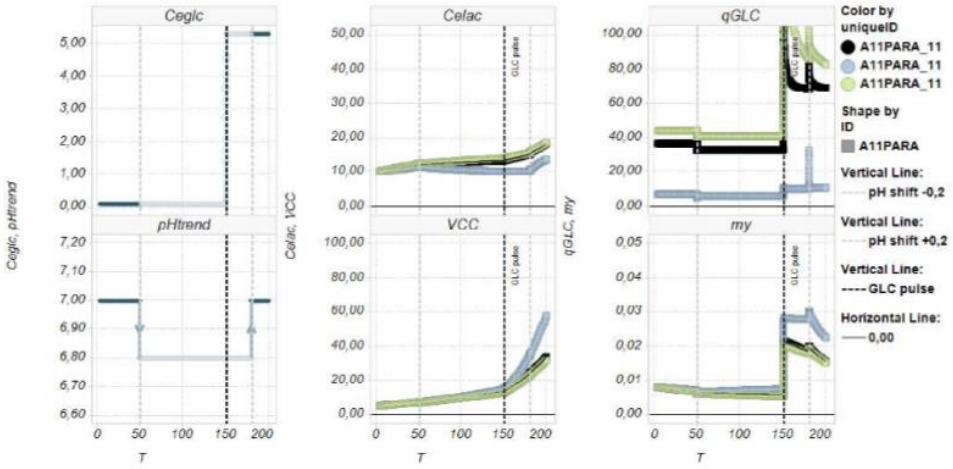


Figure 18: Simulation of different enzyme levels in HK. Black: default clone, blue: enzyme level of standard model artificially decreased ten-fold, green: enzyme level of standard model artificially increased ten-fold.

Changing (all) HK enzyme levels artificially by ten-fold reveals that the isoforms have an impact on glycolytic behavior, because more enzyme will not change flux dramatically, if substrate specificity is as low as in the default model. But if enzyme levels are decreased ten-fold, glycolytic flux is decreased. This makes sense because few available catalysts are not enough to sustain a high flux, even though the chemical environment would allow it with the given specificity of the catalyst. This becomes even more apparent when glucose is shifted up, and the flux is barely increased. Conversely, viable cell count would be highest in this simulation, simply because a high concentration of glucose and low concentration of lactic acid would lead to a very high calculated growth rate with the given formula.

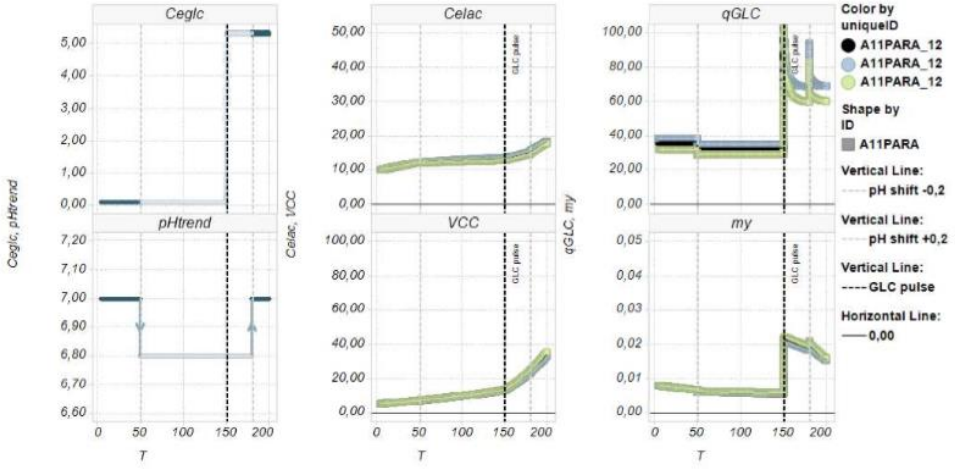


Figure 19: Simulation of binding to an activator in PK. Black: default clone, blue: high specificity to activator F1,6bP (PK-L type, Km 0,01), green: low specificity to activator (PK-M1 type, Km 1000).

Changing Km F1,6bP of PK to high substrate affinity (low Km, blue) results in a slight increase in flux with respect to the default behavior. Upon changing glucose concentration, the already strong affinity does not influence the flux any further, but it can be seen that a lower substrate affinity of the glycolysis activator F1,6bP leads to a reduced glycolytic flux by ca. 10% (69 versus 60 mM/h) which indicates that the reduction of substrate affinity can be used to direct cells metabolism to "get stuck", or rate-limited, at different stations of glycolysis, depending on allosteric regulation of the other intracellular metabolites. This involves the concentration of the activating or repressing substrate.

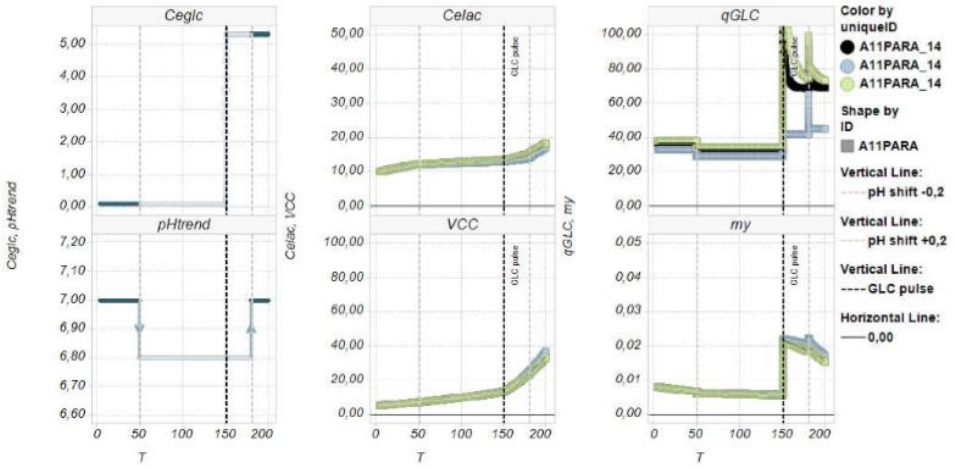


Figure 20: Simulation increased enzyme level in PFK. Black: default clone, blue: low (10-fold decreased) concentration of PFK leading to low levels of F1,6bP, green: high (10-fold more than default) concentration of PFK leading to low levels of F1,6bP. In the extracellular rates and concentrations, differences are noticed only in reduced glycolytic flux when PFK is reduced.

In our simulations, many assumptions from a calibrated model were passed on into this model, which is in its current form designed to simulate metabolic behavior in different cell lines. By doing so, several intermediate concentrations are out of their original bounds because parameters which were fitted to one clone may not fit to all the other clones any longer [44][93][94][95][96][97]. One very important intermediate which may require adaptation to physiological concentrations is F1,6bP.

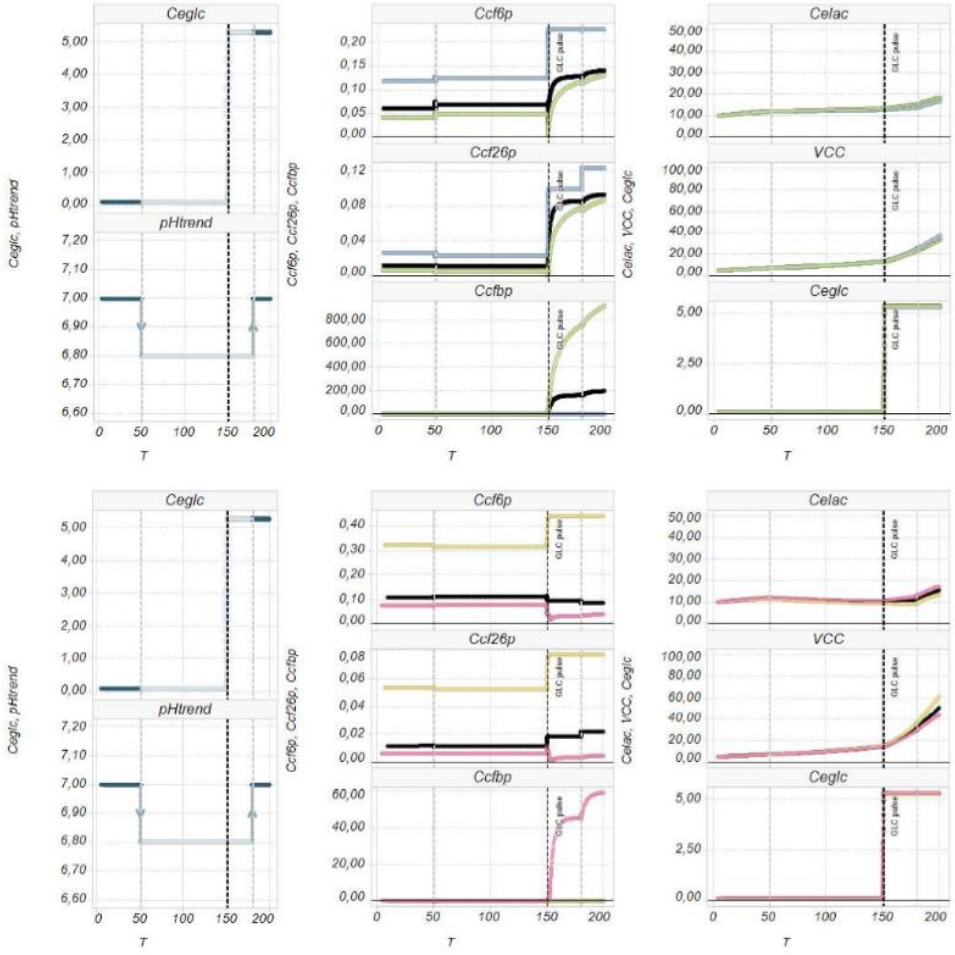


Figure 21: Sensitivity of the PFK node to parameter modification. Black: default clone behavior 11PARA/8MAB, blue/yellow: low (10-fold decreased) concentration of PFK enzyme level, Clone 11PARA/8MAB, green/pink: high (10-fold more than default) concentration of PFK enzyme level, Clone 11PARA/8MAB. Concentrations from top to bottom: F1,6bP, F2,6bP, F6P, lactic acid and glucose in mM, VCC in E5 cells/mL

The problem of a necessary parameter fit can be rarely seen in extracellular rates or concentrations alone, but rather in intracellular accumulations of metabolites in concentrations beyond natural limits. These occur when very sensitive reactions are modified with multipliers like isoform composition or transcript levels as was the case here. For allosteric regulation it does not matter much whether the concentration

is 10 or 100 mM, when the activation occurs at values over 1mM to 100% already. But it does matter when the chemical environment changes, and rates changes sign, simply because 1000mM are harder to remove per time interval than just 1mM.

As a consequence, we wish to point out that the model was applied for the first time on a big dataset. It is unlikely that the parameter estimation from one clone was so universal that it is valid for future, novel data under all kinds of different chemical conditions. But as soon as new clones are available, a new fit can and should be made by specifying upper and lower limits to all concentrations so that future simulations will lie within physiologically reasonable boundaries of the new dataset. To demonstrate this, we show exemplarily how a factor of 10 in enzyme level of PFK can lead to unphysiological F1,6bP concentrations for two clones above 60 and 900 mM. This could potentially happen when transcript level and isoform composition push one or several rates multiplicatively to such values.

4. Discussion

4.1. Insights from simulations

Our simulations showed that two parameters played a pivotal role in mammalian metabolism: i) the way glucose flux and concentration is controlled and the timing, and ii) power and direction of pH shifts. Both parameters are thus considered critical process parameters in QbD jargon [98][59][99] and can change the course of a fermentation significantly. The effect of pHi on pHe can help to explain some effects, such as reduced growth or flux under more acidic conditions, and comparison between extremely low with extremely high conditions may appear to have a high impact on flux, however, in our simulation we have shown that using a very reasonable pHi adaptation to pHe results in approximately 7-9% in flux increase or decrease per each 0,2 units of pHe which matters when the pH is prone to strong fluctuations such as at the beginning of a batch cultivation until the first 48h where the pH can easily encompass a shift of 0,6 units, (i.e. from 7,4 to 6,8). As a consequence, pH can influence both the equilibrium between lactic acid and H+ ions in and outside the cell [100]. Its effect on internal enzymes is pivotal in describing a modulation of glycolytic flux and why pH shifts are simple means to reduce or prevent overflow metabolism [39].

4.2. Link to predictive process control

Ensuring the cell's metabolic demands has been a playground for engineers for decades [84][101] and continues to be both interesting and relevant to industry even today. But the desire for developing a process with the highest final titer is challenging [25][102][103]. To stay competitive, companies have to focus on reducing the unexplainable sources of variability, and one of them is the pH profile. Lactic acid, aeration strategy, feed and medium composition all share part of the blame. It is well known that being able to hold pH without adding control reagents like CO2 or base is a challenge not many are willing or able to accept in fed-batch cultivations.

However, controlling flux and lactic acid build-up is one very potent way to ensure a consistent pH profile, which will in turn affects important parameters such as product quality, productivity, growth rate

and viability of the culture and potentially prolong cultivation time [104]. The intricate relationship between pHi and pHe may become easier to handle which in turn allows a better description of intracellular pH dependencies such as key glycolytic enzymes [42][69].

Conclusively, our simulations were conducted under two industrially relevant process modes: limitation and non-limited conditions. Both conditions were probed with pH shifts and showed to have an effect on lactic acid concentration at the time of glucose pulsing. Growth rate was shown to be dependent on timing and intensity of the process events due to a transient reduction of lactic acid build-up. However, care must be taken of the assumptions about growth. We use a very simple relationship of glucose limitation and lactic acid inhibition to describe the growth rate, which is not as simple as that in real life. In principle our assumption holds true for some process modes, but it is prone to artifacts. One such artifact was demonstrated by imposing a very low glycolytic flux, and as a consequence having no active metabolism but large growth. We therefore ask the operators to consider that only an active metabolism can make growth possible.

4.3. Link to holistic cell line engineering

Variation in isoform can be explored in depth as was demonstrated with LDH and lead to cues and ideas for improved cell line engineering, both by trying to overexpress, knock-out or modify any participating enzyme. Our results show that apart from the usual suspects HK, PFK and PK also differences in LDH isoforms [105] are worth closer inspection as they affect more directly lactic acid equilibrium inside and outside the cells. With bigger and more accurate datasets, less known enzymes may turn out to play an important part in the overall metabolic response. We therefore hypothesize that these genes might be promising targets for epigenetic engineering in future efforts to prevent lactic acid accumulation and growth inhibition.

Most genes in glycolysis are well-known in the field for years, but their differential expression leads to new cues and leverage to fine-tune metabolism *in silico* [62]. The knowledge of detailed metabolic mechanisms has already been partially integrated in constructing improved cell lines for bioprocess development which produce less lactic acid, even without a reduced glucose environment. Still, most genetic backgrounds can be further refined, for instance by making them more robust towards otherwise negative process events such as pH or even too large substrate shifts by integrating isoforms usually found in other tissues [106][107][108] which are more tolerant towards high glucose already by nature.

4.4. Link to directed protein design

Examples for successful protein engineering which has led to a favorable metabolic behavior is seldom loudly broadcasted because it represents years of research and knowledge that had to accumulate, be maintained by generations of scientist and brought to success in private and industrial research labs. However, we hope to shed some light on the process technological consequences of engineering efforts and show how metabolism as a whole is affected by a change in parameters of LDH, HK, PK or PFK.

Researchers are willing to go the long way of knock-out mutants and cycles of clone development to test the effect of different overexpression of certain particular isoforms, which end up in the wet lab for testing. Our modeling tool may be of use to visualize the mechanistic behavior prior to stepping into a laboratory and run experiments under a variety of conditions which may be used to gather important knowledge. We are aware that a satisfactory model takes a long time to develop, but we believe that even our current reduced model representing only the crudest glycolysis pathways in 13 reactions is enough to shed some light on allosteric loops which affect a culture's metabolism under norm conditions like in a bioreactor greatly. Later wet-lab experiments should then be easier to set up to draw the most out of the consequences of an alteration in the genetic code of mammalian cells [109]. We hope that future engineering efforts can be modularly extended to consider the results of simulations under a variety of cultivation conditions and process modes. As logical consequence, targeted protein design will become more important in the future and this tool might assist in speeding up bioprocess development as a whole.

5. Conclusions

This contribution shows that linking traditional technologies with each other may have several benefits over their isolated observation. First, we recreate mammalian cell culture metabolism in an artificial mathematical environment. Then, we simulated metabolic behavior under glucose limited and non-limited conditions with pH shifts of different intensity and timing. The *in silico* results show that both have a distinct effect on lactic acid profiles and also on growth. We wish to point out that metabolism can be highly clone-dependent and that our model parameters must first be recalibrated to represent internal metabolites more accurately in their physiological range. After this, two clones which behave very differently should be picked for a validation of the modified model. While we do not have these clones ready to run in our hands, we believe that this contribution is valuable to gather process knowledge outside the lab.

This model can be readily calibrated to one particular clone by changing the parameters of interest and used to explore different process control or protein engineering consequences on the metabolism. Several process modes can be set up, from batch, fed-batch or pulse-feeding scenarios to controlled limitations or pH shifts. Many unwanted and unexpected side-effects and additional complexity can be simply ruled out because the variance in the outcome of the simulations has one defined source: the variation of a parameter in the model.

In this synthetic environment, we were able to observe clone-specific differences when different genetic background, but also different process conditions are set up in a series of scenarios to test the clone's metabolic performance. It was shown that not just one, but several strategies can lead to a strong reduction in a clone's lactic acid formation in culture. This plethora of individual strategies allows for tremendous synergies, therefore they can and should be linked in order to discover the most favorable cultivation conditions for future clones.

6. Outlook

This contribution was one of the first steps to show how several disciplines can be bridged in a systems approach to explain glycolytic behavior in mammalian cell culture. We are still at the beginning to grasp the consequences of cell engineering on metabolism. Few attempts were made at all to combine related, yet so distant fields as process control and clone design. Our contribution has focused on how to integrate protein-specific substrate specificity, genetic distribution of isoforms and process control by pH shifts or glucose pulsing into one platform to make a qualitative statement about the clone's future metabolic behavior. Real wet lab verification might therefore encompass the same shifts in pH or pulses of glucose in and outside glucose-limited conditions for two opposite clones.

In our simulations, two clones had a very distinct lactic acid metabolism *in silico*. These clones were A8MAB and A11PARA, and they showed different behavior in a series of different scenarios. However, the clones are not available in our lab and we would have to expose two industrial strains to exactly the same conditions, involving proprietary medium and feed compositions. We may or may not be able to obtain the same epigenetic response if we deviate from the clone's natural cultivation conditions. Further efforts should be therefore directed to in-house clone analysis, simulation and prediction to make a really independent qualitative statement about the capabilities of this model.

List of Abbreviations

A11PARA	Clone 1
A8MAB	Clone 2
ALDO	Aldolase
CA	Concentration of compound A
Ccf26p	Cytosolic concentration of fructose-2,6-bisphosphate
Ccf6p	Cytosolic concentration of fructose-6-phosphate
Ccfbp	Cytosolic concentration of fructose-1,6-bisphosphate
Ceglc	Extracellular concentration of glucose
Celac	Extracellular concentration of lactic acid
ENO	Enolase
F6P	Fructose-6-phosphate
F1,6bP	Fructose-1,6-bisphosphate
F2,6bP	Fructose-2,6-bisphosphate
FAC	Isoform factor
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GCK	Glucokinase, HK IV
GLC	Glucose
GPI	Glucose Phosphate Isomerase
HK	Hexokinase
K, Km	Michaelis menten constant
LAC	Lactic acid

LDH	Lactate Dehydrogenase
MCT	Mono Carboxylate Transporter
my	Growth rate
N	Denominator in rate equations
PFK	Phosphofructokinase
PFKFB	6-Phosphofructo-2-Kinase/Fructose-2,6-Bisphosphatase
PGK	Phosphoglycerate Kinase
PGM	Phosphoglycerate Mutase
pHe	External pH
pHi	Internal (cytosolic) pH
PK	Pyruvate Kinase
qglc	Specific glucose uptake rate
qlac	Specific lactic acid production/consumption rate
rbw	Backward reaction
rfw	Forward reaction
rHK	Rate of HK
rmax	Maximum rate
SLCxxax	Solute Carrier, MCT
TPI	Triose Phosphate Isomerase
TSL	Transcript level factor
VCC	Viable Cell Concentration

Concentrations in mM, rates in mM/h, VCC in E5/ml, my in h-1 unless stated otherwise.

Acknowledgments

Financial support by the Marshall Plan Foundation is gratefully acknowledged. Bhanu Mulukutla is hereby kindly acknowledged for authoring the very first code lines of the mathematical model, without whom this work would not have been possible. The CEMS team at the University of Minnesota is thanked for being supportive, kind and cheerful during the whole duration of the scientific exchange.

Author Contributions

VK designed and performed the simulations, wrote the manuscript, prepared tables, figures and adapted the source code for the simulations, TL helped to explain the mathematical aspects of the code for simulations, gave valuable advice, helped with critical discussions of modeling aspects and was a good comrade during the whole time of the project and beyond, AY shared his code, literature and additional information regarding the mathematical model, RR and NV provided transcriptomic data for integration into the mathematical model, CH supervised and reviewed critically the manuscript, WH supervised, instructed, critically reviewed the manuscript, conceived the study and gave valuable support during the short research exchange at the host University.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Y.-T. Luan, T. C. Stanek, and D. Drapeau, "Controlling lactic acid production in fed-batch cell cultures via variation in glucose concentration; bioreactors and heterologous gene expression," US7429491 B2, 30-Sep-2008.
[2] R. Brent, "A partnership between biology and engineering," *Nat. Biotechnol.*, vol. 22, no. 10, pp. 1211–1214, Oct. 2004.
[3] A. R. Joyce and B. O. Palsson, "Toward whole cell modeling and simulation: comprehensive functional genomics through the constraint-based approach," *Prog. Drug Res. Fortschritte Arzneimittelforschung Prog. Rech. Pharm.*, vol. 64, pp. 265, 267–309, 2007.
[4] P. Wechselberger, P. Sagmeister, H. Engelking, T. Schmidt, J. Wenger, and C. Herwig, "Efficient feeding profile optimization for recombinant protein production using physiological information," *Bioprocess Biosyst. Eng.*, vol. 35, no. 9, pp. 1637–1649, Nov. 2012.
[5] V. Gogvadze, B. Zhivotovsky, and S. Orrenius, "The Warburg effect and mitochondrial stability in cancer cells," *Mol. Aspects Med.*, vol. 31, no. 1, pp. 60–74, Feb. 2010.
[6] O. Warburg, "On the origin of cancer cells," *Science*, vol. 123, no. 3191, pp. 309–314, Feb. 1956.
[7] A. Yongky, J. Lee, T. Le, B. C. Mulukutla, P. Daoutidis, and W.-S. Hu, "Mechanism for multiplicity of steady states with distinct cell concentration in continuous culture of mammalian cells," *Biotechnol. Bioeng.*, p. n/a–n/a, Feb. 2015.
[8] B. C. Mulukutla, M. Gramer, and W.-S. Hu, "On metabolic shift to lactate consumption in fed-batch culture of mammalian cells," *Metab. Eng.*, vol. 14, no. 2, pp. 138–149, Mar. 2012.
[9] A. Lohmeier, T. Thüte, S. Northoff, J. Hou, T. Munro, and T. Noll, "Effects of perfusion processes under limiting conditions on different Chinese Hamster Ovary cells," *BMC Proc.*, vol. 7, no. Suppl 6, p. P64, Dec. 2013.
[10] H. L. Nandita Vishwanathan, "Advancing biopharmaceutical process science through transcriptome analysis," *Curr. Opin. Biotechnol.*, vol. 30, pp. 113–119, 2014.
[11] B. C. Mulukutla, S. Khan, A. Lange, and W.-S. Hu, "Glucose metabolism in mammalian cell culture: new insights for tweaking vintage pathways," *Trends Biotechnol.*, vol. 28, no. 9, pp. 476–484, Sep. 2010.
[12] N. Vishwanathan, A. Yongky, K. C. Johnson, H.-Y. Fu, N. M. Jacob, H. Le, F. N. K. Yusufi, D. Y. Lee, and W.-S. Hu, "Global insights into the Chinese hamster and CHO cell transcriptomes," *Biotechnol. Bioeng.*, p. n/a–n/a, Dec. 2014.
[13] J. Monod, J. Wyman, and J.-P. Changeux, "On the nature of allosteric transitions: A plausible model," *J. Mol. Biol.*, vol. 12, no. 1, pp. 88–118, Mai 1965.
[14] E. Tziampazis and A. Sambanis, "Modeling of cell culture processes," *Cytotechnology*, vol. 14, no. 3, pp. 191–204, Jan. 1994.
[15] K. K. Frame and W.-S. Hu, "Kinetic study of hybridoma cell growth in continuous culture. I. A model for non-producing cells," *Biotechnol. Bioeng.*, vol. 37, no. 1, pp. 55–64, Jan. 1991.
[16] E. L. King and C. Altman, "A Schematic Method of Deriving the Rate Laws for Enzyme-Catalyzed Reactions," *J. Phys. Chem.*, vol. 60, no. 10, pp. 1375–1378, 1956.

[17] M. Aehle, K. Bork, S. Schaepe, A. Kuprijanov, R. Horstkorte, R. Simutis, and A. Lübbert, "Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach," *Cytotechnology*, vol. 64, no. 6, pp. 623–634, Mar. 2012.
[18] S. Craven, N. Shirsat, J. Whelan, and B. Glennon, "Process model comparison and transferability across bioreactor scales and modes of operation for a mammalian cell bioprocess," *Biotechnol. Prog.*, vol. 29, no. 1, pp. 186–196, 2013.
[19] B. Romein, "Mathematical modelling of mammalian cells in suspension culture," *Delft University of Technology*. [Online]. Available: http://www.tnw.tudelft.nl/en/about-faculty/departments/biotechnology/research-groups/cell-systems-engineering/research-projects/romein/. [Accessed: 22-Nov-2013].
[20] N. Vriezen, B. Romein, K. C. A. M. Luyben, and J. P. van Dijken, "Effects of glutamine supply on growth and metabolism of mammalian cells in chemostat culture," *Biotechnol. Bioeng.*, vol. 54, no. 3, pp. 272–286, 1997.
[21] G. W. Hiller, D. S. Clark, and H. W. Blanch, "Cell retention-chemostat studies of hybridoma cells-analysis of hybridoma growth and metabolism in continuous suspension culture in serum-free medium," *Biotechnol. Bioeng.*, vol. 42, no. 2, pp. 185–195, Jun. 1993.
[22] H. Niu, Z. Amribt, P. Fickers, W. Tan, and P. Bogaerts, "Metabolic pathway analysis and reduction for mammalian cell cultures—Towards macroscopic modeling," *Chem. Eng. Sci.*, vol. 102, pp. 461–473, Oktober 2013.
[23] C. . Sanderson, J. . Barford, and G. . Barton, "A structured, dynamic model for animal cell culture systems," *Biochem. Eng. J.*, vol. 3, no. 3, pp. 203–211, Jun. 1999.
[24] Z. Amribt, H. Niu, and P. Bogaerts, "Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures," *Biochem. Eng. J.*, vol. 70, pp. 196–209, Jan. 2013.
[25] J. O. Basch, S. Gangloff, C. E. Joosten, D. Kothari, S. S. Lee, K. Leister, L. Matlock, S. Sakhamuri, B. M. Schilling, and S. G. Zegarelli, "Product quality enhancement in mammalian cell culture processes for protein production," WO2004058944 A2, 15-Jul-2004.
[26] M. K. Jeon, D. Y. Yu, and G. M. Lee, "Combinatorial engineering of ldh-a and bcl-2 for reducing lactate production and improving cell growth in dihydrofolate reductase-deficient Chinese hamster ovary cells," *Appl. Microbiol. Biotechnol.*, vol. 92, no. 4, pp. 779–790, Nov. 2011.
[27] J. C. Rathmell, C. J. Fox, D. R. Plas, P. S. Hammerman, R. M. Cinalli, and C. B. Thompson, "Akt-directed glucose metabolism can prevent Bax conformation change and promote growth factor-independent survival," *Mol. Cell. Biol.*, vol. 23, no. 20, pp. 7315–7328, Oct. 2003.
[28] M. B. Burg, J. D. Ferraris, and N. I. Dmitrieva, "Cellular Response to Hyperosmotic Stresses," *Physiol. Rev.*, vol. 87, no. 4, pp. 1441–1474, Oct. 2007.
[29] F. Franěk and K. Šrámková, "Cell suicide in starving hybridoma culture: survival-signal effect of some amino acids," *Cytotechnology*, vol. 21, no. 1, pp. 81–89, Jan. 1996.
[30] R. Pörtner and T. Schäfer, "Modelling hybridoma cell growth and metabolism — a comparison of selected models and data," *J. Biotechnol.*, vol. 49, no. 1–3, pp. 119–135, Aug. 1996.
[31] "METHOD FOR CONTROLLING pH, OSMOLALITY AND DISSOLVED CARBON DIOXIDE LEVELS IN A MAMMALIAN CELL CULTURE PROCESS TO ENHANCE CELL VIABILITY AND BIOLOGIC PRODUCT YIELD."
[32] M. M. Zhu, A. Goyal, D. L. Rank, S. K. Gupta, T. V. Boom, and S. S. Lee, "Effects of Elevated pCO2 and Osmolality on Growth of CHO Cells and Production of Antibody-Fusion Protein B1: A Case Study," *Biotechnol. Prog.*, vol. 21, no. 1, pp. 70–77, 2005.
[33] R. W. Hutkins and N. L. Nannen, "pH Homeostasis in Lactic Acid Bacteria," *J. Dairy Sci.*, vol. 76, no. 8, pp. 2354–2365, Aug. 1993.
[34] W. F. Boron, "Regulation of intracellular pH," *Adv. Physiol. Educ.*, vol. 28, no. 1–4, pp. 160–179, Dec. 2004.

[35] J. A. Cook and M. H. Fox, "Effects of chronic pH 6.6 on growth, intracellular pH, and response to 42.0 degrees C hyperthermia of Chinese hamster ovary cells," *Cancer Res.*, vol. 48, no. 9, pp. 2417–2420, May 1988.
[36] G. L'Allemain, S. Paris, and J. Pouysségur, "Growth factor action and intracellular pH regulation in fibroblasts. Evidence for a major role of the Na+/H+ antiport.," *J. Biol. Chem.*, vol. 259, no. 9, pp. 5809–5815, May 1984.
[37] M. Ivarsson, T. K. Villiger, M. Morbidelli, and M. Soos, "Evaluating the impact of cell culture process parameters on monoclonal antibody N-glycosylation," *J. Biotechnol.*, vol. 188C, pp. 88–96, Aug. 2014.
[38] E. Trummer, K. Fauland, S. Seidinger, K. Schriebl, C. Lattenmayer, R. Kunert, K. Vorauer-Uhl, R. Weik, N. Borth, H. Katinger, and D. Müller, "Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors," *Biotechnol. Bioeng.*, vol. 94, no. 6, pp. 1033–1044, Aug. 2006.
[39] M. Ivarsson, H. Noh, M. Morbidelli, and M. Soos, "Insights into pH-induced metabolic switch by flux balance analysis," *Biotechnol. Prog.*, p. n/a–n/a, Jan. 2015.
[40] S. Oguchi, H. Saito, M. Tsukahara, and H. Tsumura, "pH Condition in temperature shift cultivation enhances cell longevity and specific hMab productivity in CHO culture," *Cytotechnology*, vol. 52, no. 3, pp. 199–207, Nov. 2006.
[41] M. Magnani, V. Stocchi, N. Serafini, E. Piatti, M. Dachà, and G. Fornaini, "Pig red blood cell hexokinase: Regulatory characteristics and possible physiological role," *Arch. Biochem. Biophys.*, vol. 226, no. 1, pp. 377–387, Oct. 1983.
[42] M. Magnani, V. Stocchi, G. Serafini, and M. Bossù, "Effects of buffers and pH on rabbit red blood cell hexokinase," *Ital. J. Biochem.*, vol. 32, no. 1, pp. 28–35, Feb. 1983.
[43] P. J. Mulquiney and P. W. Kuchel, "Model of the pH-dependence of the concentrations of complexes involving metabolites, haemoglobin and magnesium ions in the human erythrocyte," *Eur. J. Biochem. FEBS*, vol. 245, no. 1, pp. 71–83, Apr. 1997.
[44] P. J. Mulquiney and P. W. Kuchel, "Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: equations and parameter refinement.," *Biochem. J.*, vol. 342, no. Pt 3, pp. 581–596, Sep. 1999.
[45] R. Rekhi and A. A. Qutub, "Systems approaches for synthetic biology: a pathway toward mammalian design," *Comput. Physiol. Med.*, vol. 4, p. 285, 2013.
[46] H. Kitano, "Systems biology: a brief overview," *Science*, vol. 295, no. 5560, pp. 1662–1664, Mar. 2002.
[47] R. Feeney, A. R. Clarke, and J. J. Holbrook, "A single amino acid substitution in lactate dehydrogenase improves the catalytic efficiency with an alternative coenzyme," *Biochem. Biophys. Res. Commun.*, vol. 166, no. 2, pp. 667–672, Jan. 1990.
[48] F. Zülli, R. Schneiter, R. Urfer, and H. Zuber, "Structure and function of L-lactate dehydrogenases from thermophilic and mesophilic bacteria, XI. Engineering thermostability and activity of lactate dehydrogenases from bacilli," *Biol. Chem. Hoppe. Seyler*, vol. 372, no. 5, pp. 363–372, May 1991.
[49] B. C. Mulukutla, A. Yongky, P. Daoutidis, and W.-S. Hu, "Bistability in Glycolysis Pathway as a Physiological Switch in Energy Metabolism," *PLoS ONE*, vol. 9, no. 6, p. e98756, Jun. 2014.
[50] The UniProt Consortium, "UniProt: a hub for protein information," *Nucleic Acids Res.*, Oct. 2014.
[51] U. Wittig, R. Kania, M. Golebiewski, M. Rey, L. Shi, L. Jong, E. Algaa, A. Weidemann, H. Sauer-Danzwith, S. Mir, O. Krebs, M. Bittkowski, E. Wetsch, I. Rojas, and W. Müller, "SABIO-RK--database for biochemical reaction kinetics," *Nucleic Acids Res.*, vol. 40, no. Database issue, pp. D790–796, Jan. 2012.
[52] N. Paczia, A. Nilgen, T. Lehmann, J. Gätgens, W. Wiechert, and S. Noack, "Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms," *Microb. Cell Factories*, vol. 11, no. 1, p. 122, Sep. 2012.
[53] H. J. Cruz, J. L. Moreira, and M. J. Carrondo, "Metabolic shifts by nutrient manipulation in continuous cultures of BHK cells," *Biotechnol. Bioeng.*, vol. 66, no. 2, pp. 104–113, 1999.

[54] A. F. Europa, A. Gambhir, P.-C. Fu, and W.-S. Hu, "Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells," *Biotechnol. Bioeng.*, vol. 67, no. 1, pp. 25–34, Jan. 2000.
[55] P. M. O'Callaghan and D. C. James, "Systems biotechnology of mammalian cell factories," *Brief. Funct. Genomic. Proteomic.*, vol. 7, no. 2, pp. 95–110, Mar. 2008.
[56] G. Seth, S. Charaniya, K. F. Wlaschin, and W.-S. Hu, "In pursuit of a super producer—alternative paths to high producing recombinant mammalian cells," *Curr. Opin. Biotechnol.*, vol. 18, no. 6, pp. 557–564, Dezember 2007.
[57] O. Spadiut, S. Rittmann, C. Dietzsch, and C. Herwig, "Dynamic process conditions in bioprocess development," *Eng. Life Sci.*, vol. 13, no. 1, pp. 88–101, 2013.
[58] J. E. Dowd, K. E. Kwok, and J. M. Piret, "Glucose-based optimization of CHO-cell perfusion cultures," *Biotechnol. Bioeng.*, vol. 75, no. 2, pp. 252–256, Oct. 2001.
[59] D. M. J. D. Jesus, M. Bourgeois, G. Baumgartner, P. Tromba, D. M. Jordan, M. H. Amstutz, and P. F. M. Wurm, "The Influence of pH on Cell Growth and Specific Productivity of Two CHO Cell Lines Producing Human Anti Rh D IgG," in *Animal Cell Technology: From Target to Market*, D. E. Lindner-Olsson, M. N. Chatzissavidou, and D. E. Lüllau, Eds. Springer Netherlands, 2001, pp. 197–199.
[60] M. Gagnon, G. Hiller, Y.-T. Luan, A. Kittredge, J. DeFelice, and D. Drapeau, "High-End pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO Fed-batch cultures," *Biotechnol. Bioeng.*, vol. 108, no. 6, pp. 1328–1337, 2011.
[61] C. D. Smolke and P. A. Silver, "Informing biological design by integration of systems and synthetic biology," *Cell*, vol. 144, no. 6, pp. 855–859, Mar. 2011.
[62] N. Y. Oparina, A. V. Snezhkina, A. F. Sadritdinova, V. A. Veselovskii, A. A. Dmitriev, V. N. Senchenko, N. V. Mel'nikova, A. S. Speranskaya, M. V. Darii, O. A. Stepanov, I. M. Barkhatov, and A. V. Kudryavtseva, "[Differential expression of genes that encode glycolysis enzymes in kidney and lung cancer in humans]," *Genetika*, vol. 49, no. 7, pp. 814–823, Jul. 2013.
[63] K. Sikand, J. Singh, J. S. Ebron, and G. C. Shukla, "Housekeeping Gene Selection Advisory: Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and β-Actin Are Targets of miR-644a," *PLoS ONE*, vol. 7, no. 10, p. e47510, Oct. 2012.
[64] A. Yalcin, B. F. Clem, Y. Imbert-Fernandez, S. C. Ozcan, S. Peker, J. O'Neal, A. C. Klarer, A. L. Clem, S. Telang, and J. Chesney, "6-Phosphofructo-2-kinase (PFKFB3) promotes cell cycle progression and suppresses apoptosis via Cdk1-mediated phosphorylation of p27," *Cell Death Dis.*, vol. 5, p. e1337, 2014.
[65] S. Vora, R. Oskam, and G. E. Staal, "Isoenzymes of phosphofructokinase in the rat. Demonstration of the three non-identical subunits by biochemical, immunochemical and kinetic studies," *Biochem. J.*, vol. 229, no. 2, pp. 333–341, Jul. 1985.
[66] T. A. Rapoport, R. Heinrich, and S. M. Rapoport, "The regulatory principles of glycolysis in erythrocytes in vivo and in vitro. A minimal comprehensive model describing steady states, quasi-steady states and time-dependent processes.," *Biochem. J.*, vol. 154, no. 2, pp. 449–469, Feb. 1976.
[67] J. E. Wilson, "Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function," *J. Exp. Biol.*, vol. 206, no. 12, pp. 2049–2057, Jun. 2003.
[68] L. Miccoli, S. Oudard, F. Sureau, F. Poirson, B. Dutrillaux, and M. F. Poupon, "Intracellular pH governs the subcellular distribution of hexokinase in a glioma cell line.," *Biochem. J.*, vol. 313, no. Pt 3, pp. 957–962, Feb. 1996.
[69] I. Auzat, G. Le Bras, P. Branny, F. De La Torre, B. Theunissen, and J. R. Garel, "The role of Glu187 in the regulation of phosphofructokinase by phosphoenolpyruvate," *J. Mol. Biol.*, vol. 235, no. 1, pp. 68–72, Jan. 1994.
[70] B. Trivedi and W. H. Danforth, "Effect of pH on the kinetics of frog muscle phosphofructokinase," *J. Biol. Chem.*, vol. 241, no. 17, pp. 4110–4112, Sep. 1966.
[71] J. J. Osman, J. Birch, and J. Varley, "The response of GS-NS0 myeloma cells to pH shifts and pH perturbations," *Biotechnol. Bioeng.*, vol. 75, no. 1, pp. 63–73, Oct. 2001.

[72] S. S. Ozturk and B. O. Palsson, "Growth, metabolic, and antibody production kinetics of hybridoma cell culture: 2. Effects of serum concentration, dissolved oxygen concentration, and medium pH in a batch reactor," *Biotechnol. Prog.*, vol. 7, no. 6, pp. 481–494, Dec. 1991.
[73] N. Kurano, C. Leist, F. Messi, S. Kurano, and A. Fiechter, "Growth behavior of Chinese hamster ovary cells in a compact loop bioreactor: 1. Effects of physical and chemical environments," *J. Biotechnol.*, vol. 15, no. 1–2, pp. 101–111, Jul. 1990.
[74] J. P. B. J. D. Jang, "Effect of feed rate on growth rate and antibody production in the fed-batch culture of murine hybridoma cells," *Cytotechnology*, vol. 32, no. 3, 2000.
[75] K. Fischer, P. Hoffmann, S. Voelkl, N. Meidenbauer, J. Ammer, M. Edinger, E. Gottfried, S. Schwarz, G. Rothe, S. Hoves, K. Renner, B. Timischl, A. Mackensen, L. Kunz-Schughart, R. Andreesen, S. W. Krause, and M. Kreutz, "Inhibitory effect of tumor cell-derived lactic acid on human T cells," *Blood*, vol. 109, no. 9, pp. 3812–3819, May 2007.
[76] L. Xie, G. Nyberg, X. Gu, H. Li, F. Möllborn, and D. I. Wang, "Gamma-interferon production and quality in stoichiometric fed-batch cultures of Chinese hamster ovary (CHO) cells under serum-free conditions," *Biotechnol. Bioeng.*, vol. 56, no. 5, pp. 577–582, Dec. 1997.
[77] V. Singh, "On-line measurement of oxygen uptake in cell culture using the dynamic method," *Biotechnol. Bioeng.*, vol. 52, no. 3, pp. 443–448, 1996.
[78] K. Eyer, A. Oeggerli, and E. Heinzle, "On-line gas analysis in animal cell cultivation: II. Methods for oxygen uptake rate estimation and its application to controlled feeding of glutamine," *Biotechnol. Bioeng.*, vol. 45, no. 1, pp. 54–62, 1995.
[79] P. Ducommun, P.-A. Ruffieux, M.-P. Furter, I. Marison, and U. von Stockar, "A new method for on-line measurement of the volumetric oxygen uptake rate in membrane aerated animal cell cultures," *J. Biotechnol.*, vol. 78, no. 2, pp. 139–147, März 2000.
[80] S. Ansorge, G. Esteban, C. Ghommidh, and G. Schmid, "Monitoring Nutrient Limitations by Online Capacitance Measurements in Batch & Fed-batch CHO Fermentations," in *Cell Technology for Cell Products*, R. Smith, Ed. Springer Netherlands, 2007, pp. 723–726.
[81] J. E. Dowd, A. Jubb, K. E. Kwok, and J. M. Piret, "Optimization and control of perfusion cultures using a viable cell probe and cell specific perfusion rates," *Cytotechnology*, vol. 42, no. 1, pp. 35–45, May 2003.
[82] T. Noll and M. Biselli, "Dielectric spectroscopy in the cultivation of suspended and immobilized hybridoma cells," *J. Biotechnol.*, vol. 63, no. 3, pp. 187–198, Aug. 1998.
[83] A. M. Jobé, C. Herwig, M. Surzyn, B. Walker, I. Marison, and U. von Stockar, "Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing," *Biotechnol. Bioeng.*, vol. 82, no. 6, pp. 627–639, Jun. 2003.
[84] S. S. Ozturk, J. C. Thrift, J. D. Blackie, and D. Naveh, "Real-time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor," *Biotechnol. Bioeng.*, vol. 53, no. 4, pp. 372–378, 1997.
[85] H. P. J. Bonarius, C. D. de Gooijer, J. Tramper, and G. Schmid, "Determination of the respiration quotient in mammalian cell culture in bicarbonate buffered media," *Biotechnol. Bioeng.*, vol. 45, no. 6, pp. 524–535, 1995.
[86] M. Jenzsch, S. Gnoth, M. Kleinschmidt, R. Simutis, and A. Lübbert, "Improving the batch-to-batch reproducibility of microbial cultures during recombinant protein production by regulation of the total carbon dioxide production," *J. Biotechnol.*, vol. 128, no. 4, pp. 858–867, März 2007.
[87] G. W. Gould and G. D. Holman, "The glucose transporter family: structure, function and tissue-specific expression.," *Biochem. J.*, vol. 295, no. Pt 2, pp. 329–341, Oct. 1993.
[88] W. Zhou, J. Rehm, and W.-S. Hu, "High viable cell concentration fed-batch cultures of hybridoma cells through on-line nutrient feeding," *Biotechnol. Bioeng.*, vol. 46, no. 6, pp. 579–587, 1995.
[89] "Novel strategy to reduce lactic acid production and control ph in animal cell culture."
[90] F. Lu, P. C. Toh, I. Burnett, F. Li, T. Hudson, A. Amanullah, and J. Li, "Automated dynamic fed-batch process and media optimization for high productivity cell culture process development," *Biotechnol. Bioeng.*, vol. 110, no. 1, pp. 191–205, 2013.

[91] D. Chee Furng Wong, K. Tin Kam Wong, L. Tang Goh, C. Kiat Heng, and M. Gek Sim Yap, "Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures," *Biotechnol. Bioeng.*, vol. 89, no. 2, pp. 164–177, 2005.
[92] R. Stambaugh and D. Post, "Substrate and Product Inhibition of Rabbit Muscle Lactic Dehydrogenase Heart (H4) and Muscle (M4) Isozymes," *J. Biol. Chem.*, vol. 241, no. 7, pp. 1462–1467, Apr. 1966.
[93] S. Merry and H. G. Britton, "The mechanism of rabbit muscle phosphofructokinase at pH8," *Biochem. J.*, vol. 226, no. 1, pp. 13–28, Feb. 1985.
[94] G. A. Dunaway, T. P. Kasten, T. Sebo, and R. Trapp, "Analysis of the phosphofructokinase subunits and isoenzymes in human tissues.," *Biochem. J.*, vol. 251, no. 3, pp. 677–683, May 1988.
[95] R. L. Hanson, F. B. Rudolph, and H. A. Lardy, "Rabbit muscle phosphofructokinase. The kinetic mechanism of action and the equilibrium constant," *J. Biol. Chem.*, vol. 248, no. 22, pp. 7852–7859, Nov. 1973.
[96] M. Otto, R. Heinrich, G. Jacobasch, and S. Rapoport, "A mathematical model for the influence of anionic effectors on the phosphofructokinase from rat erythrocytes," *Eur. J. Biochem. FEBS*, vol. 74, no. 2, pp. 413–420, Apr. 1977.
[97] M. Otto, R. Heinrich, B. Kühn, and G. Jacobasch, "A Mathematical Model for the Influence of Fructose 6-Phosphate, ATP, Potassium, Ammonium and Magnesium on the Phosphofructokinase from Rat Erythrocytes," *Eur. J. Biochem.*, vol. 49, no. 1, pp. 169–178, Nov. 1974.
[98] S. F. Abu-Absi, L. Yang, P. Thompson, C. Jiang, S. Kandula, B. Schilling, and A. A. Shukla, "Defining process design space for monoclonal antibody cell culture," *Biotechnol. Bioeng.*, vol. 106, no. 6, pp. 894–905, 2010.
[99] D. Zalai, C. Dietzsch, and C. Herwig, "Risk-based Process Development of Biosimilars as Part of the Quality by Design Paradigm," *PDA J. Pharm. Sci. Technol. PDA*, vol. 67, no. 6, pp. 569–580, Dec. 2013.
[100] J. W. Locasale and L. C. Cantley, "Metabolic Flux and the Regulation of Mammalian Cell Growth," *Cell Metab.*, vol. 14, no. 4, pp. 443–451, Oktober 2011.
[101] D. Drapeau, Y.-T. Luan, and T. C. Stanek, "Restricted glucose feed for animal cell culture," WO2004104186 A1, 02-Dec-2004.
[102] S. D. Roger and A. Mikhail, "Biosimilars: opportunity or cause for concern?," *J. Pharm. Pharm. Sci. Publ. Can. Soc. Pharm. Sci. Société Can. Sci. Pharm.*, vol. 10, no. 3, pp. 405–410, 2007.
[103] H. Schellekens, "Biosimilar therapeutics—what do we need to consider?," *NDT Plus*, vol. 2, no. Suppl 1, pp. i27–i36, Jan. 2009.
[104] H. Aghamohseni, K. Ohadi, M. Spearman, N. Krahn, M. Moo-Young, J. M. Scharer, M. Butler, and H. M. Budman, "Effects of nutrient levels and average culture pH on the glycosylation pattern of camelid-humanized monoclonal antibody," *J. Biotechnol.*, vol. 186C, pp. 98–109, Jul. 2014.
[105] W.-S. Hu, A. Kantardjieff, and B. C. Mulukutla, "Cell lines that overexpress lacatate dehydrogenase c," WO2012075124 A2, 07-Jun-2012.
[106] R. Noguchi, H. Kubota, K. Yugi, Y. Toyoshima, Y. Komori, T. Soga, and S. Kuroda, "The selective control of glycolysis, gluconeogenesis and glycogenesis by temporal insulin patterns," *Mol. Syst. Biol.*, vol. 9, p. 664, 2013.
[107] E. W. Somberg and M. A. Mehlman, "Regulation of gluconeogenesis and lipogenesis. The regulation of mitochondrial pyruvate metabolism in guinea-pig liver synthesizing precursors for gluconeogenesis," *Biochem. J.*, vol. 112, no. 4, pp. 435–447, May 1969.
[108] S. J. Pilkis and T. H. Claus, "Hepatic Gluconeogenesis/Glycolysis: Regulation and Structure/Function Relationships of Substrate Cycle Enzymes," *Annu. Rev. Nutr.*, vol. 11, no. 1, pp. 465–515, 1991.
[109] N. J. Brown, S. E. Higham, B. Perunovic, M. Arafa, S. Balasubramanian, and I. Rehman, "Lactate Dehydrogenase-B Is Silenced by Promoter Methylation in a High Frequency of Human Breast Cancers," *PLoS ONE*, vol. 8, no. 2, p. e57697, Feb. 2013.

[110] G. Gerber, H. Preissler, R. Heinrich, and S. M. Rapoport, "Hexokinase of Human Erythrocytes," *Eur. J. Biochem.*, vol. 45, no. 1, pp. 39–52, Jun. 1974.
[111] G. Rijksen, G. Jansen, R. J. Kraaijenhagen, M. J. Van der Vlist, A. M. Vlug, and G. E. Staal, "Separation and characterization of hexokinase I subtypes from human erythrocytes," *Biochim. Biophys. Acta*, vol. 659, no. 2, pp. 292–301, Jun. 1981.
[112] G. Rijksen and G. E. Staal, "Regulation of human erythrocyte hexokinase. The influence of glycolytic intermediates and inorganic phosphate," *Biochim. Biophys. Acta*, vol. 485, no. 1, pp. 75–86, Nov. 1977.
[113] G. Fornaini, M. Magnani, A. Fazi, A. Accorsi, V. Stocchi, and M. Dachà, "Regulatory properties of human erythrocyte hexokinase during cell ageing," *Arch. Biochem. Biophys.*, vol. 239, no. 2, pp. 352–358, Jun. 1985.
[114] R. W. Gracy and B. E. Tilley, "Phosphoglucose isomerase of human erythrocytes and cardiac tissue," *Methods Enzymol.*, vol. 41, pp. 392–400, 1975.
[115] S. Kitajima, R. Sakakibara, and K. Uyeda, "Kinetic studies of fructose 6-phosphate,2-kinase and fructose 2,6-bisphosphatase," *J. Biol. Chem.*, vol. 259, no. 11, pp. 6896–6903, Jun. 1984.
[116] M. Kretschmer, W. Schellenberger, and E. Hofmann, "Quasi-stationary concentrations of fructose-2,6-bisphosphate in the phosphofructokinase-2/fructose-2,6-bisphosphatase cycle," *Biochem. Biophys. Res. Commun.*, vol. 131, no. 2, pp. 899–904, Sep. 1985.
[117] D. A. Okar, A. Manzano, A. Navarro-Sabatè, L. Riera, R. Bartrons, and A. J. Lange, "PFK-2/FBPase-2: maker and breaker of the essential biofactor fructose-2,6-bisphosphate," *Trends Biochem. Sci.*, vol. 26, no. 1, pp. 30–35, Jan. 2001.
[118] S. K. Srivastava and E. Beutler, "The effect of normal red cell constituents on the activities of red cell enzymes," *Arch. Biochem. Biophys.*, vol. 148, no. 1, pp. 249–255, Jan. 1972.
[119] E. Beutler, "2,3-diphosphoglycerate affects enzymes of glucose metabolism in red blood cells," *Nature. New Biol.*, vol. 232, no. 27, pp. 20–21, Jul. 1971.
[120] A. H. Mehler, "Kinetic Properties of Native and Carboxypeptidase-altered Rabbit Muscle Aldolase," *J. Biol. Chem.*, vol. 238, no. 1, pp. 100–104, Jan. 1963.
[121] A. H. Mehler and B. Bloom, "Interaction between rabbit muscle aldolase and dihydroxyacetone phosphate," *J. Biol. Chem.*, vol. 238, pp. 105–107, Jan. 1963.
[122] E. E. Penhoet, M. Kochman, and W. J. Rutter, "Ioslation of fructose diphosphate aldolases A, B, and C," *Biochemistry (Mosc.)*, vol. 8, no. 11, pp. 4391–4395, Nov. 1969.
[123] E. E. Penhoet, M. Kochman, and W. J. Rutter, "Molecular and catalytic properties of aldolase C," *Biochemistry (Mosc.)*, vol. 8, no. 11, pp. 4396–4402, Nov. 1969.
[124] E. Strapazon and T. L. Steck, "Interaction of the aldolase and the membrane of human erythrocytes," *Biochemistry (Mosc.)*, vol. 16, no. 13, pp. 2966–2971, Jun. 1977.
[125] D. R. Yeltman and B. G. Harris, "Purification and characterization of aldolase from human erythrocytes," *Biochim. Biophys. Acta*, vol. 484, no. 1, pp. 188–198, Sep. 1977.
[126] I. A. Rose, E. L. O'Connell, and A. H. Mehler, "Mechanism of the Aldolase Reaction," *J. Biol. Chem.*, vol. 240, no. 4, pp. 1758–1765, Apr. 1965.
[127] E. Beutler, *Red Cell Metabolism: A Manual of Biochemical Methods*, 3 Sub edition. Orlando, FL: Grune & Stratton, 1984.
[128] T. H. Sawyer, B. E. Tilley, and R. W. Gracy, "Studies on Human Triosephosphate Isomerase II. NATURE OF THE ELECTROPHORETIC MULTIPLICITY IN ERYTHROCYTES," *J. Biol. Chem.*, vol. 247, no. 20, pp. 6499–6505, Oct. 1972.
[129] A. S. Schneider, W. N. Valentine, M. Hattori, and H. L. Heins, "Hereditary Hemolytic Anemia with Triosephosphate Isomerase Deficiency," *N. Engl. J. Med.*, vol. 272, no. 5, pp. 229–235, Feb. 1965.
[130] O. Meyerhof and R. Junowicz-Kocholaty, "The Equilibria of Isomerase and Aldolase, and the Problem of the Phosphorylation of Glyceraldehyde Phosphate," *J. Biol. Chem.*, vol. 149, no. 1, pp. 71–92, Jul. 1943.

[131] R. W. Gracy, "Triosephosphate isomerase from human erythrocytes," *Methods Enzymol.*, vol. 41, pp. 442–447, 1975.
[132] C. S. Wang and P. Alaupovic, "Glyceraldehyde-3-phosphate dehydrogenase from human erythrocyte membranes. Kinetic mechanism and competitive substrate inhibition by glyceraldehyde 3-phosphate," *Arch. Biochem. Biophys.*, vol. 205, no. 1, pp. 136–145, Nov. 1980.
[133] F. Heinz and B. Freimüller, "Glyceraldehyde-3-phosphate dehydrogenase from human tissues," *Methods Enzymol.*, vol. 89 Pt D, pp. 301–305, 1982.
[134] C. S. Furfine and S. F. Velick, "THE ACYL-ENZYME INTERMEDIATE AND THE KINETIC MECHANISM OF THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE REACTION," *J. Biol. Chem.*, vol. 240, pp. 844–855, Feb. 1965.
[135] C. F. Cori, S. F. Velick, and G. T. Cori, "The combination of diphosphopyridine nucleotide with glyceraldehyde phosphate dehydrogenase," *Biochim. Biophys. Acta*, vol. 4, no. 1–3, pp. 160–169, Jan. 1950.
[136] W. K. G. Krietsch and T. Bücher, "3-Phosphoglycerate Kinase from Rabbit Sceletal Muscle and Yeast," *Eur. J. Biochem.*, vol. 17, no. 3, pp. 568–580, Dec. 1970.
[137] A. Yoshida and S. Watanabe, "Human phosphoglycerate kinase. I. Crystallization and characterization of normal enzyme," *J. Biol. Chem.*, vol. 247, no. 2, pp. 440–445, Jan. 1972.
[138] M. Ali and Y. S. Brownstone, "A study of phosphoglycerate kinase in human erythrocytes. II. Kinetic properties," *Biochim. Biophys. Acta*, vol. 445, no. 1, pp. 89–103, Aug. 1976.
[139] W. J. O. C S Lee, "Properties and mechanism of human erythrocyte phosphoglycerate kinase.," *J. Biol. Chem.*, vol. 250, no. 4, pp. 1275–81, 1975.
[140] M. P, B. W, and K. P, "Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations1: in vivo kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using 13C and 31P NMR," 15-Sep-1999. [Online]. Available: http://www.biochemj.org/bj/342/bj3420567.htm. [Accessed: 16-Nov-2014].
[141] P. J. Mulquiney and P. W. Kuchel, "Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: computer simulation and metabolic control analysis," *Biochem. J.*, vol. 342 Pt 3, pp. 597–604, Sep. 1999.
[142] C. C. Rider and C. B. Taylor, "Enolase isoenzymes in rat tissues: Electrophoretic, chromatographic, immunological and kinetic properties," *Biochim. Biophys. Acta BBA - Protein Struct.*, vol. 365, no. 1, pp. 285–300, Sep. 1974.
[143] F. Wold and C. E. Ballou, "Studies on the enzyme enolase. I. Equilibrium studies," *J. Biol. Chem.*, vol. 227, no. 1, pp. 301–312, Jul. 1957.
[144] L. Garfinkel and D. Garfinkel, "Magnesium regulation of the glycolytic pathway and the enzymes involved," *Magnesium*, vol. 4, no. 2–3, pp. 60–72, 1985.
[145] H. G. Holzhütter, G. Jacobasch, and A. Bisdorff, "Mathematical modelling of metabolic pathways affected by an enzyme deficiency. A mathematical model of glycolysis in normal and pyruvate-kinase-deficient red blood cells," *Eur. J. Biochem. FEBS*, vol. 149, no. 1, pp. 101–111, May 1985.
[146] J. F. Koster, R. G. Slee, G. E. Staal, and T. J. van Berkel, "The influence of glucose I,6-diphosphate on the enzymatic activity of pyruvate kinase," *Biochim. Biophys. Acta*, vol. 258, no. 3, pp. 763–768, Mar. 1972.
[147] S. Ainsworth and N. Macfarlane, "A kinetic study of rabbit muscle pyruvate kinase," *Biochem. J.*, vol. 131, no. 2, pp. 223–236, Feb. 1973.
[148] A. Kahn and J. Marie, "Pyruvate kinases from human erythrocytes and liver," *Methods Enzymol.*, vol. 90 Pt E, pp. 131–140, 1982.
[149] K. R. Albe, M. H. Butler, and B. E. Wright, "Cellular concentrations of enzymes and their substrates," *J. Theor. Biol.*, vol. 143, no. 2, pp. 163–195, Mar. 1990.
[150] E. Rozengurt, L. J. de Asúa, and H. Carminatti, "Some Kinetic Properties of Liver Pyruvate Kinase (Type L) II. EFFECT OF pH ON ITS ALLOSTERIC BEHAVIOR," *J. Biol. Chem.*, vol. 244, no. 12, pp. 3142–3147, Jun. 1969.

[151] V. Zewe and H. J. Fromm, "Kinetic Studies of Rabbit Muscle Lactate Dehydrogenase. II. Mechanism of the Reaction*," *Biochemistry (Mosc.)*, vol. 4, no. 4, pp. 782–792, Apr. 1965.
[152] U. Borgmann, T. W. Moon, and K. J. Laidler, "Molecular kinetics of beef heart lactate dehydrogenase," *Biochemistry (Mosc.)*, vol. 13, no. 25, pp. 5152–5158, Dec. 1974.
[153] C. S. Wang, "Inhibition of human erythrocyte lactate dehydrogenase by high concentrations of pyruvate. Evidence for the competitive substrate inhibition," *Eur. J. Biochem. FEBS*, vol. 78, no. 2, pp. 569–574, Sep. 1977.
[154] M. Uldry and B. Thorens, "The SLC2 family of facilitated hexose and polyol transporters," *Pflüg. Arch. Eur. J. Physiol.*, vol. 447, no. 5, pp. 480–489, Feb. 2004.
[155] A. P. Halestrap and D. Meredith, "The SLC16 gene family-from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond," *Pflüg. Arch. Eur. J. Physiol.*, vol. 447, no. 5, pp. 619–628, Feb. 2004.
[156] C. Juel and A. P. Halestrap, "Lactate transport in skeletal muscle — role and regulation of the monocarboxylate transporter," *J. Physiol.*, vol. 517, no. Pt 3, pp. 633–642, Jun. 1999.

Appendix

This appendix is based on the appendices of [7][49], and unpublished work under review by Yongky et al. with a few optional additions which can be activated or deactivated in the code of the model.

Rate equations

Kinetic rate equations for all the enzymes used in the model have all been previously derived mechanistically and reported in various literature. This section describes these kinetic equations for all the enzymes considered. The steady state kinetics for all the enzymes were based on the King and Altman method [16] employing the known mechanisms and regulations. The rate equations for the enzymes phosphofructokinase (PFK) and pyruvate kinase (PK) employ the Monod-Wyman-Changeux method [13] to model the allosteric effects of various metabolite modulators. Kinetic constants used in the rate equations of the current model were experimentally determined values and obtained from previously reported studies.

Glycolysis

Hexokinase (HK): The rate equation for HK was taken from Mulquiney et al. [44]. The kinetic constants which correspond to those of the isozyme HK2 were adopted from previous literature [110]–[113]. The rate equation employs the partial rapid equilibrium random bi bi mechanism, which is a simplification of the steady state random bi bi system with the assumption that except for the reactive-ternary complexes, all the other steps in the mechanism are fast reactions. The inhibitions by G6P, glucose-1,6-phosphate (G16BP), 2,3-bisphosphoglycerate (23BPG) and glutathione (GSH) were modeled as mixed type of inhibition affecting both the activity (V_{max}) as well as the affinity (K_M) of the enzyme for glucose.

$$r_{HK} = \left(V_{mf}^{HK} \frac{C_{MgATP}^{c} C_{GLC}^{c}}{K_{MgATP}^{HK} K_{GLC}^{HK}} - V_{mr}^{HK} \frac{C_{MgADP}^{c} C_{G6P}^{c}}{K_{i,MgADP}^{HK} K_{G6P}^{HK}} \right) \frac{1}{N_{HK}}$$

$$N_{HK} = \left(1 + \frac{C_{MgATP}^{c}}{K_{i,MgATP}^{HK}} + \frac{C_{G6P}^{c}}{K_{i,G6P}^{HK}} + \frac{C_{GLC}^{c}}{K_{GLC}^{HK}} + \frac{C_{MgATP}^{c} C_{GLC}^{c}}{K_{MgATP}^{HK} K_{GLC}^{HK}} + \frac{C_{MgADP}^{c}}{K_{i,MgADP}^{HK}} + \frac{C_{MgADP}^{c} C_{G6P}^{c}}{K_{i,MgADP}^{HK} K_{G6P}^{HK}} + \frac{C_{GLC}^{c} C_{G6P}^{c}}{K_{GLC}^{HK} K_{G6P}^{HK}} + \frac{C_{GLC}^{c} C_{G16BP}^{c}}{K_{GLC}^{HK} K_{i,G16BP}^{HK}} + \frac{C_{GLC}^{c} C_{23BPG}^{c}}{K_{GLC}^{HK} K_{i,23BPG}^{HK}} + \frac{C_{GLC}^{c} C_{GSH}^{c}}{K_{GLC}^{HK} K_{i,GSH}^{HK}} \right)$$

$V_{mf}^{HK} = 6.38*10^{3} mM\ h^{-1}$
$V_{mr}^{HK} = 41\ mM\ h^{-1}$
$K_{MgATP}^{HK} = 1.0\ mM$
$K_{GLC}^{HK} = 0.1\ mM$
$K_{i,MgADP}^{HK} = 1.0\ mM$
$K_{G6P}^{HK} = 0.47\ mM$
$K_{i,G6P}^{"HK} = 0.47\ mM$
$K_{i,G16BP}^{HK} = 0.03\ mM$
$K_{i,GSH}^{HK} = 3.0\ mM$
$K_{i,23BPG}^{HK} = 4.0\ mM$

Glucose Phosphate Isomerase (GPI): The rate equation for GPI was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [114]. The rate equation employs the steady state uni uni reaction kinetics.

$$r_{GPI} = \frac{V_{mf}^{GPI} \frac{C_{G6P}^{c}}{K_{f}^{GPI}} - V_{mr}^{GPI} \frac{C_{F6P}^{c}}{K_{r}^{GPI}}}{1 + \frac{C_{G6P}^{c}}{K_{f}^{GPI}} + \frac{C_{F6P}^{c}}{K_{r}^{GPI}}}$$

$V_{mf}^{GPI} = 4.8*10^{4} mM\ h^{-1}$
$V_{mr}^{GPI} = 4.0*10^{4} mM\ h^{-1}$
$K_{f}^{GPI} = 0.3\ mM$
$K_{r}^{GPI} = 0.123\ mM$

Phosphofructokinase (PFK): The rate equation for PFK was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [93]–[97]. The rate kinetics was based on the two state allosteric model using ordered bi bi mechanism. The two state model considers that the enzyme can exist in the active or the non-active state as determined by the levels of the activity modulators. These include activators (F6P, F16BP, F26BP, G16BP, AMP etc) and inhibitors (ATP, Mg etc). Some of these activity modulators act on all the isozymes of PFK while others are isozyme specific. For example, F6P acts as a substrate as well as an allosteric activator for all the isozymes of PFK, whereas F16BP only stimulates PFKM and PFKL. The fraction of enzyme in the active state is represented by the nonlinear term N_{PFK} which is a function of the levels of the activity modulators. L_{PFK} represents the equilibrium constant between the two states of the enzyme in the absence of any substrates. The initial velocity expression for the enzyme fraction in the active state was modeled as partial rapid equilibrium random bi bi steady state equation similar to the HK kinetics.

$$r_{PFK} = \frac{\frac{V_f^{PFK} C_{MgATP}^c C_{F6P}^c}{K_{F6P}^{PFK} K_{MgATP}^{PFK}} - \frac{V_r^{PFK} C_{MgADP}^c C_{F16BP}^c}{K_{F16BP}^{PFK} K_{MgADP}^{PFK}}}{\left(\left(1+\frac{C_{F6P}^c}{K_{F6P}^{PFK}}\right)\left(1+\frac{C_{MgATP}^c}{K_{MgATP}^{PFK}}\right)+\left(1+\frac{C_{F16BP}^c}{K_{F16BP}^{PFK}}\right)\left(1+\frac{C_{MgADP}^c}{K_{MgADP}^{PFK}}\right)-1\right)} \frac{1}{N_{PFK}}$$

$$N_{PFK} = 1 + \frac{L_{PFK}\left(1+\frac{C_{ATP}^c}{K_{ATP}^{PFK}}\right)^4\left(1+\frac{C_{Mg}^c}{K_{Mg}^{PFK}}\right)^4\left(1+\frac{C_{2,3BPG}^c}{K_{2,3BPG}^{PFK}}\right)^4}{\left(1+\frac{C_{F6P}^c}{K_{F6P}^{PFK}}+\frac{C_{F16BP}^c}{K_{F16BP}^{PFK}}\right)^4\left(1+\frac{C_{AMP}^c}{K_{AMP}^{PFK}}\right)^4\left(1+\frac{C_{G16BP}^c}{K_{G16BP}^{PFK}}\right)^4\left(1+\frac{C_{Pi}^c}{K_{Pi}^{PFK}}\right)^4\left(1+\frac{C_{F26BP}^c}{K_{F26BP}^{PFK}}\right)^4}$$

$V_f^{PFK} = 15.5*10^2 mM^{-1}h^{-1}$
$V_r^{PFK} = 6.78*10^1 mM^{-1}h^{-1}$
$K_{F6P}^{PFK} = 6*10^{-2}\ mM$
$K_{MgATP}^{PFK} = 6.8*10^{-2}\ mM$
$K_{MgADP}^{PFK} = 0.54\ mM$
$K_{F16BP}^{PFK} = 0.65\ mM$
$K_{F26BP}^{PFK} = 5.5*10^{-3} mM$
$K_{G16BP}^{PFK} = 0.1\ mM$
$K_{ATP}^{PFK} = 0.1\ mM$
$K_{AMP}^{PFK} = 0.3\ mM$
$K_{Mg}^{PFK} = 0.2\ mM$
$K_{Pi}^{PFK} = 30\ mM$
$K_{23BPG}^{PFK} = 0.5\ mM$
$L_{PFK} = 2*10^{-3}$

6-Phosphofructo-2-Kinase/Fructose-2,6-Bisphosphatase (PFKFB): The rate equation for PFKFB and the kinetic constants were taken from previously reported studies [115][116]. PFKFB is a bi-functional enzyme with kinase and bisphosphatase activities, each localized to either terminals of the enzyme and are independent of each other's activity. The kinase domain catalyzes the synthesis of fructose-2,6-bisphosphate (F26BP) from fructose-6-phoshate (F6P) and the bisphosphatase domain mediates the hydrolysis of F26BP to F6P. The reaction kinetics for the kinase domain (r_{PFK2}) follows the ordered bi bi steady state kinetics, with phosphoenolpyurvate (PEP) inhibition of the kinase domain modeled as non-competitive inhibition. The bisphosphatase reaction kinetics ($r_{F2,6BPase}$) was modeled as simple Michaelis-Menten kinetics with non-competitive product inhibition by F6P. Isozymes of PFKFB vary in their kinase to bisphosphatase activity (K/P) [117]. The effect of isozyme (or K/P) was modeled by changing the Vmax of r_{PFK2} and holding $r_{F2,6BPase}$ constant.

$$r_{PFK2}=\frac{V_{f,PFK2}\left(C^c_{ATP}C^c_{F6P}-\frac{C^c_{ADP}C^c_{F26BP}}{K_{eq,PFK2}}\right)}{\left(K^{PFK2}_{i,ATP}K^{PFK2}_{m,F6P}+K^{PFK2}_{m,F6P}C^c_{ATP}+K^{PFK2}_{m,ATP}C^c_{F6P}+\frac{K^{PFK2}_{m,ADP}C^c_{F26B[}}{K_{eq,PFK2}}+\frac{K^{PFK2}_{m,F26B[}C^c_{ADP}}{K_{eq,PFK2}}\right.}$$
$$\left.+C^c_{ATP}C^c_{F6P}+\frac{K^{PFK2}_{m,ADP}C^c_{ATP}C^c_{F26BP}}{K_{eq,PFK2}K^{PFK2}_{i,ATP}}+\frac{C^c_{ADP}C^c_{F26BP}}{K_{eq,PFK2}}+\frac{K^{PFK2}_{m,ATP}C^c_{ADP}C^c_{F6P}}{K^{PFK2}_{i,ADP}}\right.$$
$$\left.+\frac{C^c_{ATP}C^c_{F6P}C^c_{F26BP}}{K^{PFK2}_{i,F26BP}}+\frac{C^c_{ADP}C^c_{F6P}C^c_{F26BP}}{K_{eq,PFK2}K^{PFK2}_{i,F6P}}\right)\left(1+\frac{C^c_{PEP}}{K^{PFK2}_{i,PEP}}\right)$$

$V_{f,PFK2}=41.6\ mM\ h^{-1}$
$K^{PFK2}_{m,ATP}=0.15\ mM$
$K^{PFK2}_{m,F6P}=0.032\ mM$
$K^{PFK2}_{m,F26BP}=0.008\ mM$
$K^{PFK2}_{m,ADP}=0.062\ mM$
$K^{PFK2}_{i,ATP}=0.15\ mM$
$K^{PFK2}_{i,F6P}=0.001\ mM$
$K^{PFK2}_{i,F26BP}=0.02\ mM$
$K^{PFK2}_{i,ADP}=0.23\ mM$
$K^{PFK2}_{i,PEP}=0.013\ mM$
$K^{PFK2}_{eq}=16$

$$r_{F2,6BPase}=\frac{V_{F2,6BPase}C^c_{F26BP}}{\left(1+\frac{C^c_{F6P}}{K^{F2,6BPase}_{i,F6P}}\right)\left(K^{F2,6BPase}_{m,F26BP}+C^c_{F26BP}\right)}$$

$V_{F2,6BPase}=11.78\ mM^{-1}h^{-1}$
$K^{F2,6BPase}_{m,F26BP}=1*10^{-3}\ mM$
$K^{F2,6BPase}_{i,F6P}=25*10^{-3}\ mM$

Aldolase (ALDO): The rate equation for ALDO was taken from Mulquiney et al. [44]. The kinetic constants were adopted or estimated from previous literature [118]–[127]. The reaction kinetics of ALDO follows the ordered uni bi steady state kinetics. Inhibition due to 23BPG as described in the original expression was retained in this study. However, since 23BPG is not a reaction intermediate considered in the model, its concentration was held constant for the purpose of this study.

$$r_{ALD}=\frac{\frac{V^{ALD}_{mf}C^c_{F16BP}}{K^{ALD}_{F16BP}}-\frac{V^{ALD}_{mr}C^c_{GAP}C^c_{DHAP}}{K^{ALD}_{GAP}K^{ALD}_{i,DHAP}}}{\left(1+\frac{C^c_{23BPG}}{K^{ALD}_{i,23BPG}}+\frac{C^c_{F16BP}}{K^{ALD}_{F16BP}}+\frac{K^{ALD}_{DHAP}C^c_{GAP}}{K^{ALD}_{GAP}K^{ALD}_{i,DHAP}}\left(1+\frac{C^c_{23BPG}}{K^{ALD}_{i,23BPG}}\right)+\frac{C^c_{DHAP}}{K^{ALD}_{i,DHAP}}\right.}$$
$$\left.+\frac{K^{ALD}_{DHAP}C^c_{F16BP}C^c_{GAP}}{K^{ALD}_{i,F16BP}K^{ALD}_{GAP}K^{ALD}_{i,DHAP}}+\frac{C^c_{DHAP}C^c_{GAP}}{K^{ALD}_{GAP}K^{ALD}_{i,DHAP}}\right)$$

$V^{ALD}_{mf}=6.75*10^{2}mM\ h^{-1}$
$V^{ALD}_{mr}=2.32*10^{3}mM\ h^{-1}$
$K^{ALD}_{F16BP}=5*10^{-2}mM$
$K^{ALD}_{i,F16BP}=1.98*10^{-2}mM$
$K^{ALD}_{DHAP}=3.5*10^{-2}mM$
$K^{ALD}_{i,DHAP}=1.1*10^{-2}mM$
$K^{ALD}_{GAP}=0.189mM$
$K^{ALD}_{i,23BPG}=1.5mM$

Triose Phosphate Isomerase (TPI): The rate equation for TPI was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [127]–[131]. The rate kinetics of TPI follows a simple steady state uni uni reaction kinetics.

$$r_{TPI} = \frac{V_{mf}^{TPI}\frac{C_{DHAP}^{c}}{K_{f}^{TPI}} - V_{mr}^{TPI}\frac{C_{GAP}^{c}}{K_{r}^{TPI}}}{1+\frac{C_{DHAP}^{c}}{K_{f}^{TPI}}+\frac{C_{GAP}^{c}}{K_{r}^{TPI}}}$$

$$V_{mf}^{TPI} = 5.10*10^{2}\,mM\ h^{-1}$$
$$V_{mr}^{TPI} = 4.61*10^{1}\,mM\ h^{-1}$$
$$K_{f}^{TPI} = 1.62*10^{-1}\,mM$$
$$K_{r}^{TPI} = 4.30*10^{-1}\,mM$$

Glyceraldehyde 3-Phosphate Dehydrogenase (GAPDH): The rate equation for GAPDH was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [132]–[135]. The rate kinetics of GAPDH follows the ter ter (bi uni uni bi ping pong) steady state kinetics.

$$r_{GAPD} = \frac{V_{mf}^{GAPD}\frac{C_{NAD^+}^{c}C_{Pi}^{c}C_{GAP}^{c}}{K_{NAD^+}^{GAPD}K_{i,Pi}^{GAPD}K_{i,GAP}^{GAPD}} - V_{mr}^{GAPD}\frac{C_{13BPG}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,13BPG}^{GAPD}K_{NADH}^{GAPD}}}{\frac{C_{GAP}^{c}}{K_{i,GAP}^{GAPD}}\left(1+\frac{C_{GAP}^{c}}{K_{i,GAP}^{'GAPD}}\right)+\frac{C_{13BPG}^{c}}{K_{i,13BPG}^{GAPD}}\left(1+\frac{C_{GAP}^{c}}{K_{i,GAP}^{'GAPD}}\right)+\frac{K_{13BPG}^{GAPD}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,13BPG}^{GAPD}K_{NADH}^{GAPD}}+\frac{K_{GAP}^{GAPD}C_{NAD^+}^{c}C_{Pi}^{c}}{K_{NAD^+}^{'GAPD}K_{i,Pi}^{'GAPD}K_{i,GAP}^{'GAPD}}}$$
$$+\frac{C_{NAD^+}^{c}C_{GAP}^{c}}{K_{i,NAD^+}^{'GAPD}K_{i,GAP}^{GAPD}}+\frac{C_{Pi}^{c}C_{GAP}^{c}}{K_{i,Pi}^{GAPD}K_{i,GAP}^{GAPD}}\left(1+\frac{C_{GAP}^{c}}{K_{i,GAP}^{'GAPD}}\right)+\frac{C_{NAD^+}^{c}C_{13BPG}^{c}}{K_{i,NAD^+}^{'GAPD}K_{i,13BPG}^{GAPD}}+\frac{K_{13BPG}^{GAPD}C_{Pi}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,Pi}^{GAPD}K_{i,13BPG}^{'GAPD}K_{NADH}^{GAPD}}$$
$$+\frac{C_{GAP}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,GAP}^{'GAPD}K_{i,NADH}^{GAPD}}+\frac{C_{13BPG}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,13BPG}^{'GAPD}K_{NADH}^{'GAPD}}+\frac{C_{NAD^+}^{c}C_{Pi}^{c}C_{GAP}^{c}}{K_{NAD^+}^{'GAPD}K_{i,Pi}^{'GAPD}K_{i,GAP}^{'GAPD}}+\frac{K_{GAP}^{GAPD}C_{NAD^+}^{c}C_{Pi}^{c}C_{13BPG}^{c}}{K_{i,GAP}^{'GAPD}K_{NAD^+}^{'GAPD}K_{i,Pi}^{'GAPD}K_{i,13BPG}^{'GAPD}}$$
$$+\frac{C_{Pi}^{c}C_{GAP}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,Pi}^{'GAPD}K_{i,GAP}^{'GAPD}K_{i,NADH}^{'GAPD}}+\frac{C_{Pi}^{c}C_{13BPG}^{c}C_{NADH}^{c}C_{H^+}^{c}}{K_{i,13BPG}^{'GAPD}K_{NADH}^{'GAPD}K_{i,Pi}^{'GAPD}K_{i,13BPG}^{'GAPD}}$$

$$V_{mf}^{'GAPD} = 5.317*10^{3}\,mMh^{-1}$$
$$V_{mr}^{'GAPD} = 3.919*10^{3}\,mMh^{-1}$$
$$K_{NAD^+}^{'GAPD} = 0.045mM$$
$$K_{i,NAD^+}^{'GAPD} = 0.045mM$$
$$K_{Pi}^{'GAPD} = 3.16mM$$
$$K_{i,Pi}^{'GAPD} = 3.16mM$$
$$K_{GAP}^{'GAPD} = 0.095mM$$
$$K_{i,GAP}^{'GAPD} = 1.59*10^{-16}\,mM$$
$$K_{i,GAP}^{'GAPD} = 0.031mM$$
$$K_{NADH}^{'GAPD} = 0.0033mM$$
$$K_{i,NADH}^{'GAPD} = 0.01mM$$
$$K_{13BPG}^{'GAPD} = 0.00671mM$$
$$K_{i,13BPG}^{'GAPD} = 1.52*10^{-18}\,mM$$
$$K_{i,13BPG}^{'GAPD} = 0.001mM$$
$$K_{eq}^{'GAPD} = 1.9*10^{-8}$$

Phosphoglycerate Kinase (PGK): The rate equation for PGK was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [136]–[139]. The rate kinetics of PGK follows the partial rapid equilibrium random bi bi steady state kinetics.

$$r_{PGK} = \frac{V_{mf}^{PGK}\frac{C_{13BPG}^{c}C_{MgADP}^{c}}{K_{i,MgADP}^{PGK}K_{13BPG}^{PGK}} - V_{mr}^{PGK}\frac{C_{3PG}^{c}C_{MgATP}^{c}}{K_{i,MgATP}^{PGK}K_{3PG}^{PGK}}}{1+\frac{C_{13BPG}^{c}}{K_{i,13BPG}^{PGK}}+\frac{C_{MgADP}^{c}}{K_{i,MgADP}^{PGK}}+\frac{C_{13BPG}^{c}C_{MgADP}^{c}}{K_{i,MgADP}^{PGK}K_{13BPG}^{PGK}}+\frac{C_{3PG}^{c}}{K_{i,3PG}^{PGK}}+\frac{C_{MgATP}^{c}}{K_{i,MgATP}^{PGK}}+\frac{C_{3PG}^{c}C_{MgATP}^{c}}{K_{i,MgATP}^{PGK}K_{3PG}^{PGK}}}$$

$$V_{mf}^{PGK} = 5.96*10^{4} mM\ h^{-1}$$
$$V_{mr}^{PGK} = 2.39*10^{4} mM\ h^{-1}$$
$$K_{MgADP}^{PGK} = 0.1mM$$
$$K_{i,MgADP}^{PGK} = 0.08mM$$
$$K_{13BPG}^{PGK} = 0.002mM$$
$$K_{i,13BPG}^{PGK} = 1.6mM$$
$$K_{MgATP}^{PGK} = 1mM$$
$$K_{i,MgATP}^{PGK} = 0.186mM$$
$$K_{3PG}^{PGK} = 1.1mM$$
$$K_{i,3PG}^{PGK} = 0.205mM$$
$$K_{eq}^{PGK} = 3.2*10^{3}$$

Phosphoglycerate Mutase (PGM): The rate equation for PGM was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [140][141]. The rate kinetics of PGM follows the uni uni steady state kinetics.

$$r_{PGAM} = \frac{V_{mf}^{PGAM}\frac{C_{3PG}^{c}}{K_{3PG}^{PGAM}} - V_{mr}^{PGAM}\frac{C_{2PG}^{c}}{K_{2PG}^{PGAM}}}{1+\frac{C_{3PG}^{c}}{K_{3PG}^{PGAM}}+\frac{C_{2PG}^{c}}{K_{2PG}^{PGAM}}}$$

$$V_{mf}^{PGAM} = 4.894*10^{5} mM\ h^{-1}$$
$$V_{mr}^{PGAM} = 4.395*10^{5} mM\ h^{-1}$$
$$K_{3PG}^{PGAM} = 0.168mM$$
$$K_{2PG}^{PGAM} = 0.0256mM$$
$$K_{eq}^{PGAM} = 0.17$$

Enolase (ENO): The rate equation for ENO was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [142][143][144]. The rate kinetics of ENO follows the partial rapid equilibrium random bi bi steady state kinetics.

$$r_{ENO} = \frac{V_{mf}^{ENO}\frac{C_{2PG}^{c}C_{Mg}^{c}}{K_{i,Mg}^{ENO}K_{2PG}^{ENO}} - V_{mr}^{ENO}\frac{C_{PEP}^{c}C_{Mg}^{c}}{K_{i,Mg}^{ENO}K_{PEP}^{ENO}}}{1+\frac{C_{2PG}^{c}}{K_{i,2PG}^{ENO}}+\frac{C_{Mg}^{c}}{K_{i,Mg}^{ENO}}+\frac{C_{2PG}^{c}C_{Mg}^{c}}{K_{i,Mg}^{ENO}K_{2PG}^{ENO}}+\frac{C_{PEP}^{c}}{K_{i,PEP}^{ENO}}+\frac{C_{Mg}^{c}}{K_{i,Mg}^{ENO}}+\frac{C_{PEP}^{c}C_{Mg}^{c}}{K_{i,Mg}^{ENO}K_{PEP}^{ENO}}}$$

$$V_{mf}^{ENO} = 2.106*10^{4} mM\ h^{-1}$$
$$V_{mr}^{ENO} = 5.542*10^{3} mM\ h^{-1}$$
$$K_{i,Mg}^{ENO} = K_{Mg}^{ENO} = 0.14mM$$
$$K_{PEP}^{ENO} = K_{PEP}^{ENO} = 0.11mM$$
$$K_{2PG}^{ENO} = K_{2PG}^{ENO} = 0.046mM$$
$$K_{eq}^{ENO} = 3.0$$

Pyruvate Kinase (PK): The rate equation for PK was taken from Mulquiney et al. [44]. The kinetic constants were adopted from previous literature [145], [145]–[150]. Like PFK, the rate kinetics of PK was based on the two state allosteric model using the ordered bi bi mechanism. The two state model considers that the enzyme can exist in active or non-active state determined by the levels of the activity modulators. These include activators (F16BP, PEP, PYR etc) and inhibitors (ATP, ALA etc). The fraction of the enzyme in the active state is represented by the nonlinear term N_{PK} which is a function of levels of activity modulators. L_{PK} represents the equilibrium constant between enzymes at the two states in the absence of any substrates. The initial velocity expression for the enzyme fraction in the active state is modeled as partial rapid equilibrium random bi bi steady state equation.

$$r_{PK} = \left(\frac{V_{mf}^{PK} \frac{C_{PEP}^{c}}{K_{PEP}^{PK}} \frac{C_{MgADP}^{c}}{K_{MgADP}^{PK}} - V_{mr}^{PK} \frac{C_{PYR}^{c}}{K_{PYR}^{PK}} \frac{C_{MgATP}^{c}}{K_{MgATP}^{PK}}}{\left(1+\frac{C_{PEP}^{c}}{K_{PEP}^{PK}}\right)\left(1+\frac{C_{MgADP}^{c}}{K_{MgADP}^{PK}}\right)+\left(1+\frac{C_{PYR}^{c}}{K_{PYR}^{PK}}\right)\left(1+\frac{C_{MgATP}^{c}}{K_{MgATP}^{PK}}\right)-1} \right) \frac{1}{N_{PK}}$$

$$N_{PK} = 1 + L_{PK} \frac{\left(1+\frac{C_{ATP}^{c}}{K_{ATP}^{PK}}\right)^{4}\left(1+\frac{C_{ALA}^{c}}{K_{ALA}^{PK}}\right)^{4}}{\left(1+\frac{C_{PEP}^{c}}{K_{PEP}^{PK}}+\frac{C_{PYR}^{c}}{K_{PYR}^{PK}}\right)^{4}\left(1+\frac{C_{F16BP}^{c}}{K_{F16BP}^{PK}}+\frac{C_{G16BP}^{c}}{K_{G16BP}^{PK}}\right)^{4}}$$

$V_{mf}^{PK} = 2.02*10^{4} mM\ h^{-1}$
$V_{mf}^{PK} = 47.5 mM\ h^{-1}$
$K_{PEP}^{PK} = 2.25*10^{-1} mM$
$K_{MgADP}^{PK} = 4.74*10^{-1} mM$
$K_{MgATP}^{PK} = 3\ mM$
$K_{ATP}^{PK} = 3.39\ mM$
$K_{PYR}^{PK} = 4\ mM$
$K_{F16BP}^{PK} = 0.04\ mM$
$K_{G16BP}^{PK} = 1.0*10^{-1} mM$
$L_{PK} = 0.398$
$K_{ALA}^{PK} = 0.02\ mM$

Lactate Dehydrogenase (LDH): The rate equation for LDH and the kinetic constants were adopted from previous literature [44][151]–[153]. The kinetics of LDH was modeled as ordered bi bi steady state kinetics, with substrate inhibition by pyruvate.

$$r_{LDH}=\frac{V_{mf}^{LDH}\frac{C_{NADH}^{c}C_{PYR}^{c}}{K_{i,NADH}^{LDH}K_{PYR}^{LDH}}-V_{mr}^{LDH}\frac{C_{NAD}^{c}C_{LAC}^{c}}{K_{i,NAD}^{LDH}K_{LAC}^{LDH}}}{\left(1+\frac{K_{NADH}^{LDH}C_{PYR}^{c}}{K_{i,NADH}^{LDH}K_{PYR}^{LDH}}+\frac{K_{NAD}^{LDH}C_{LAC}^{c}}{K_{i,NAD}^{LDH}K_{LAC}^{LDH}}\right)\left(1+\frac{C_{PYR}^{c}}{K_{i,PYR}^{"LDH}}\right)+\frac{C_{NADH}^{c}}{K_{i,NADH}^{LDH}}+\frac{C_{NAD}^{c}}{K_{i,NAD}^{LDH}}}$$
$$+\frac{C_{NADH}^{c}C_{PYR}^{c}}{K_{i,NADH}^{LDH}K_{PYR}^{LDH}}+\frac{K_{NAD}^{LDH}C_{NADH}^{c}C_{LAC}^{c}}{K_{i,NAD}^{LDH}K_{i,NADH}^{LDH}K_{LAC}^{LDH}}+\frac{K_{NADH}^{LDH}C_{NAD}^{c}C_{PYR}^{c}}{K_{i,NAD}^{LDH}K_{i,NADH}^{LDH}K_{PYR}^{LDH}}+\frac{C_{NAD}^{c}C_{LAC}}{K_{i,NAD}^{LDH}K_{LAC}^{LDH}}$$
$$+\frac{C_{NADH}^{c}C_{PYR}^{c}C_{LAC}^{c}}{K_{i,NADH}^{LDH}K_{PYR}^{LDH}K_{i,LAC}^{LDH}}+\frac{C_{NAD}^{c}C_{PYR}^{c}C_{LAC}^{c}}{K_{i,NAD}^{LDH}K_{i,PYR}^{LDH}K_{LAC}^{LDH}}$$

$V_{mf}^{LDH}=8.66*10^{3}mM\ h^{-1}$
$V_{mr}^{LDH}=2.17*10^{3}mM\ h^{-1}$
$K_{PYR}^{LDH}=0.137\ mM$
$K_{i,PYR}^{LDH}=0.228\ mM$
$K_{NAD}^{LDH}=0.107\ mM$
$K_{i,NAD}^{LDH}=0.503\ mM$
$K_{LAC}^{LDH}=1.07\ mM$
$K_{i,LAC}^{LDH}=7.33\ mM$
$K_{i,NADH}^{LDH}=5.45*10^{-3}\ mM$
$K_{NADH}^{LDH}=7.43*10^{-3}\ mM$
$K_{i,PYR}^{"LDH}=0.101\ mM$

Transport

Glucose Transporter (GLUT): Glucose transporters mediate transport of glucose across plasma membranes. Till date, fourteen glucose transporters (isozymes) have been identified which perform the same function but have very different kinetic properties [154]. Kinetics of the GLUT1 isozyme was considered in the model and was modeled as uni uni steady state kinetics.

$$r_{GLUT}=\frac{V_{mf}^{GLUT}\frac{C_{GLC}^{e}}{K_{GLC}^{GLUT}}-V_{mr}^{GLUT}\frac{C_{GLC}^{c}}{K_{GLC}^{GLUT}}}{1+\frac{C_{GLC}^{e}}{K_{GLC}^{GLUT}}+\frac{C_{GLC}^{c}}{K_{GLC}^{GLUT}}}$$

$V_{mf}^{GLUT}=7.67\ \mathrm{mM}h^{-1}$
$V_{mr}^{GLUT}=0.767\ \mathrm{mM}h^{-1}$
$K_{GLC}^{GLUT}=1.50\ mM$

Mitochondrial Pyruvate Transporter: Rate equation for pyruvate transport into mitochondrion was modeled as reversible first ordered mass kinetics.

$$r_{PYRH}=V_{mf}^{PYRH}\left(C_{PYR}^{c}C_{H^{+}}^{c}-C_{PYR}^{m}C_{H^{+}}^{m}\right)$$

$V_{mf}^{PYRH}=6.67*10^{12}\ mM^{-1}\ h^{-1}$

Mono Carboxylate Transporter (MCT)

The kinetics of MCT was modeled as ordered bi bi mechanism. The kinetic constants for MCT were adopted from [155] and [156].

$$r_{MCT} = \frac{V_{m,MCT}\left(C^c_{H^+}C^c_{lac} - C^e_{H^+}C^e_{lac}\right)}{K^{MCT}_{m,laci}K^{MCT}_{i,H^+} + K^{MCT}_{m,H^+}C^c_{lac} + K^{MCT}_{m,laci}C^c_{H^+} + K^{MCT}_{m,H^+}C^e_{lac} + K^{MCT}_{m,laci}C^e_{H^+} + C^c_{H^+}C^c_{lac}}$$
$$C^e_{H^+}C^e_{lac} + \frac{K^{MCT}_{m,H^+}C^e_{lac}C^e_{H^+}}{K^{MCT}_{i,H^+}} + \frac{K^{MCT}_{m,H^+}C^e_{lac}C^c_{H^+}}{K^{MCT}_{i,H^+}} + \frac{C^c_{lac}C^c_{H^+}C^e_{H^+}}{K^{MCT}_{ii,H^+}} + \frac{C^e_{lac}C^c_{H^+}C^e_{H^+}}{K^{MCT}_{ii,H^+}}$$

$V_{MCT} = 2.73*10^3\ mM^{-1}h^{-1}$
$K^{MCT}_{m,H^+} = 10^{-4}\ mM$
$K^{MCT}_{m,laci} = 2.5\ mM$
$K^{MCT}_{i,H^+} = 2*10^{-4}\ mM$
$K^{MCT}_{ii,H^+} = 2*10^{-4}\ mM$

$P_{ATP} = 1 + \frac{C^m_{H^+}}{K_{HATP}} + \frac{C^m_{Mg}}{K_{MgATP}} + \frac{C^m_K}{K_{KATP}}$	$K_{HATP} = 2.57*10^{-7}M$ $K_{Mg4TP} = 1.51*10^{-4}M$ $K_{KATP} = 1.35*10^{-2}M$
$P_{ADP} = 1 + \frac{C^m_{H^+}}{K_{HADP}} + \frac{C^m_{Mg}}{K_{MgADP}} + \frac{C^m_K}{K_{KADP}}$	$K_{HADP} = 3.8*10^{-7}M$ $K_{MgADP} = 1.62*10^{-3}M$ $K_{KADP} = 2.95*10^{-2}M$
$P_{AMP} = 1 + \frac{C^m_{H^+}}{K_{HAMP}} + \frac{C^m_{Mg}}{K_{Mg4MP}} + \frac{C^m_K}{K_{KAMP}}$	$K_{HAMP} = 6.03*10^{-7}M$ $K_{Mg4MP} = 1.38*10^{-2}M$ $K_{KAMP} = 8.91*10^{-2}M$
$P_{GTP} = P_{ATP}$ $P_{GDP} = P_{ADP}$	
$P_{CoASH} = 1 + \frac{C^m_{H^+}}{K_{HCoASH}}$	$K_{HCoASH} = 7.41*10^{-9}M$
$P_{cit} = 1 + \frac{C^m_{H^+}}{K_{Hcit}} + \frac{C^m_{Mg}}{K_{Mgcit}} + \frac{C^m_K}{K_{Kcit}}$	$K_{Hcit} = 2.34*10^{-6}M$ $K_{Mgcit} = 4.27*10^{-4}M$ $K_{Kcit} = 4.58*10^{-1}M$
$P_{scoa} = 1 + \frac{C^m_{H^+}}{K_{Hscoa}}$	$K_{Hscoa} = 1.1*10^{-4}M$
$P_{fATP} = 1 + \frac{C^m_{H^+}}{K_{HATP}} + \frac{C^m_{K^+}}{K_{KATP}}$	
$P_{fADP} = 1 + \frac{C^m_{H^+}}{K_{HADP}} + \frac{C^m_{K^+}}{K_{KADP}}$	
$P_{fAMP} = 1 + \frac{C^m_{H^+}}{K_{HAMP}} + \frac{C^m_{K^+}}{K_{KAMP}}$	
$P_{fGTP} = P_{fATP}$ $P_{fGDP} = P_{fGDP}$	

Mathematical model

The mathematical model for the cellular metabolism consists of material balance equations for each reaction intermediates in glycolysis. The reaction equations are from the mechanistic equations shown in the previous section. The dilution effect on metabolite concentrations caused by the cell growth was neglected considering the difference of at least one order of magnitude between the time constant for growth and specific glucose consumption rate.

1. Glucose: $\frac{dC^c_{GLC}}{dt} = r_{GLUT} - r_{HK}$
2. Glucose 6-phosphate: $\frac{dC^c_{G6P}}{dt} = r_{HK} - r_{GPI}$
3. Fructose 6-phosphate: $\frac{dC^c_{F6P}}{dt} = r_{GPI} - r_{PFK} - r_{PFK2} + r_{F2,6BPase}$
4. Fructose 1,6-bisphosphate: $\frac{dC^c_{F16BP}}{dt} = r_{PFK} - r_{ALD}$
5. Fructose 2,6-bisphosphate: $\frac{dC^c_{F26BP}}{dt} = r_{PFK2} - r_{F2,6BPase}$
6. Dihydroxyacetone phosphate: $\frac{dC^c_{1DHAP}}{dt} = r_{ALD} - r_{TPI}$
7. Glyceraldehyde 3-phosphate: $\frac{dC^c_{GAP}}{dt} = r_{ALD} + r_{TPI} - r_{GAPD}$
8. 1,3-bisphosphoglycerate: $\frac{dC^c_{13BPG}}{dt} = r_{GAPD} - r_{PGK}$
9. 3-phosphoglycerate: $\frac{dC^c_{3PG}}{dt} = r_{PGK} - r_{PGM}$
10. 2-phosphoglycerate: $\frac{dC^c_{2PG}}{dt} = r_{PGM} - r_{EN}$

11. Phosphoenolpyruvate: $\frac{dC^c_{PEP}}{dt} = r_{EN} - r_{PK}$

12. Pyruvate: $\frac{dC^c_{PYR}}{dt} = r_{PK} - r_{LDH} - r_{PYRH}$

13. Lactate: $\frac{dC^e_{lac}}{dt} = r_{LDH} - r_{MCT}$

2.4 Chapter 4

AIChE

A Robust Feeding Strategy to Maintain Set-Point Glucose in Mammalian Fed-Batch Cultures When Input Parameters Have a Large Error

Viktor Konakovsky
Div. of Biochemical Engineering, Inst. of Chemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A 166-4, Vienna 1060, Austria

Christoph Clemens
Boehringer Ingelheim Pharma GmbH & Co. KG Dep. Bioprocess Development, Biberach, Germany

Markus Michael Müller
Boehringer Ingelheim Pharma GmbH & Co. KG Dep. Bioprocess Development, Biberach, Germany

Jan Bechmann
Boehringer Ingelheim Pharma GmbH & Co. KG Dep. Bioprocess Development, Biberach, Germany

Christoph Herwig
Div. of Biochemical Engineering, Inst. of Chemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A 166-4, Vienna 1060, Austria

DOI 10.1002/btpr.2438
Published online 00 Month 2017 in Wiley Online Library (wileyonlinelibrary.com)

Industrial CHO cell cultures run under fed-batch conditions are required to be controlled in particular ranges of glucose, while glucose is constantly consumed and must be replenished by a feed. The most appropriate feeding rate is ideally stoichiometric and adaptive in nature to balance the dynamically changing rate of glucose consumption. However, high errors in biomass and glucose estimation as well as limited knowledge of the true metabolic state challenge the control strategy. In this contribution, we take these errors into account and simulate the output with uncertainty trajectories in silico *in order to control glucose concentration. Other than many control strategies, which require parameter estimation, our assumptions are founded on two pillars: (i) first principles and (ii) prior knowledge about the variability of fed-batch CHO cell culture. The algorithm was exposed to an* in-silico *Design of Experiments (DoE), in which variations of parameters were changed simultaneously, such as clone-specific behavior, precision of equipment and desired control range used. The results demonstrate that our method achieved the target of holding the glucose concentration within an acceptable range. A robust and sufficient level of control could be demonstrated even with high errors for biomass or metabolic state estimation. In a time where blockbuster drugs are queuing up for time slots of their production, this transferable control strategy that is independent of tedious establishment runs may be a decisive advantage for rapid implementation during technology transfer and scale up and decrease in campaign change over time.* © 2017 American Institute of Chemical Engineers *Biotechnol. Prog.*, 000:000–000, 2017
Keywords: CHO cell culture, fed-batch, simulation, dynamic process control, uncertainty

Introduction

Problem statement

In Chinese Hamster Ovary (CHO) cell culture, glucose concentration (GLC) is recognized as a very important and often critical process parameter (CPP) in fermentations, as it affects lactic acid profiles as well as product glycosylation and glycation patterns in mammalian cell cultures.[1–4] For this reason, glucose levels are preferably tightly controlled. Today, most large scale manufacturing operations are run as dynamic fed-batch fermentations. Parameters such as cell count and substrate consumption rate are subject to change during cultivation time, and their estimation often comes with a considerable error. This may lead to processes not running within specifications (i.e., outside a particular glucose concentration range). Optimally, a combination of novel sensors which send regular intervals to a control station may help to monitor and finally enable closed-loop control of a bioprocess. The factors influencing the performance of any feeding strategy, namely cell variability, sensor quality and desired process specifications, were actively taken into account to suggest a stoichiometric feeding strategy which will keep glucose concentration within its designed limits at all time. In this contribution, a

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

Correspondence concerning this article should be addressed to Christoph Herwig at christoph.herwig@tuwien.ac.at

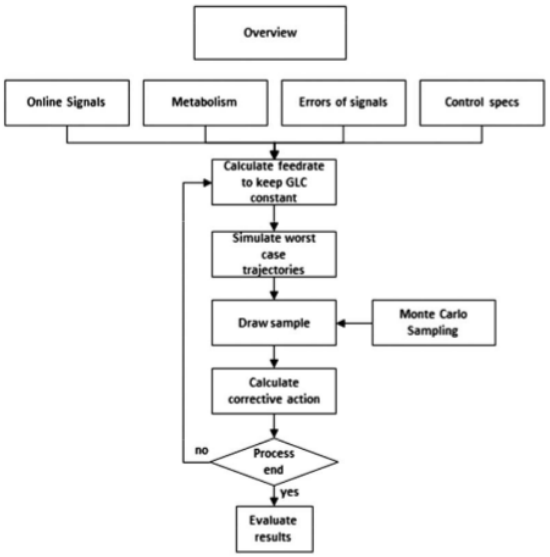


Figure 1. Outline of this work.

feed control strategy was coupled with simulations of typical fed-batch processes, and uncertainty of the abovementioned signals was actively taken into account (Figure 1).

Uncertainty in literature

Uncertainty refers to the fact that something is in the state of uncertainty, doubtful or unknown.[5] Here we limit ourselves mainly to the uncertainty in mammalian cell count estimation using novel sensor technology and uncertainty of metabolism, which are combined and used as lumped errors to predict glucose concentrations. These worst case trajectories then indicate appropriate sampling times, and lead to a correction of the current feedrate, so that future estimations stay better within specified control limits.

Almost 30 years ago, a very comprehensive review was presented by M.B. Beck[6] to examine and tackle the uncertainty in different problem areas, namely: (i) uncertainty about model structure, (ii) uncertainty in the estimation of model parameters, (iii) the propagation of errors due to the uncertainties and (iv) the design of experiments when to reduce the critical uncertainties. The parameters that are subject to uncertainty could furthermore vary over time or be subject to spontaneous disturbances which could change the estimated coefficients in the model profoundly. Uncertainty has to be taken into account in many scientific disciplines where it is an important consideration to prevent critical failures. Examples are found in the aviation industry (flight trajectories),[7] financial technologies (price bidding on the energy market from renewable but unstable energy sources),[8] to classical chemical processes (syngas production).[9] Often, uncertainty is addressed using Monte-Carlo sampling from a probability distribution[10]; in more advanced cases polynomial chaos expansion is applied when dealing with computationally intense problems.[11,12] Adding uncertainty to simulations allows to make better statements about its output and adds a layer of confidence to its results.[13]

In this contribution we dealt with a metabolic state which was constrained within a range that could change with changing environment, over time, or by sudden disturbances in the process (i.e., pH or temperature shifts). We generally assumed that the uncertainty spans the range between two extremes and that the variable of interest could be sampled from a noisy, Quasi-Gaussian distribution that was considered adequate for our problem statement. This resulted in uncertainty trajectories which would evolve over time as the simulation progressed. Similar trajectories as we used them can be generalized in non-linear systems where upper and lower probabilities are represented as fuzzy numbers.[14] When the error ranges are well known, ordinary differential equations (ODEs) can be set up to calculate trajectories which are based on the maximum lumped uncertainty like in guidance systems.[15] Biotechnological applications increasingly consider the propagation of input uncertainty to ensure good modeling practice,[16] to achieve high quality of the developed models even though sensor input is not ideal,[17,18] and to meet end-product specifications by ensuring that the trajectories in which product quality was accepted were never breached.[19]

If distributions are well known, Bayesian approaches may be employed[20,21] to impute meaningful values when high quality data is missing or otherwise unavailable,[22,23] a concept that blends prior expert knowledge with mathematical expressions and allows simulations of experiments *in silico* to get statistical confidence in the models developed. In conclusion, dealing with uncertainty in biological systems using simulations is a natural exercise in risk-minimization. Several models may be consulted to deal with the lack of information[24]; while some are mechanistic in nature and invoke a general principle or law, others draw a conclusion from machine learning, or on the best guess based on prior knowledge or rules.[25,26]

Industrial feeding strategies

Feeding in mammalian cell cultures is a discipline that draws a lot of attention. The way how a culture is fed will affect parameters such as peak cell count, final titer, specific glucose and lactic acid consumption among others.[27] One strategy is to establish bolus feeding profiles which are administered routinely as part of daily sampling. They are considered to be straight-forward and easy to accomplish, if only few vessels are operated. Another strategy is to set a defined, continuous feed rate regardless of culture feedback. Bolus and continuous feeding are easy to realize in both process development and manufacturing environments, however, there are a couple of very significant drawbacks. For instance, the most appropriate feeding regime may need to be optimized in screening experiments, thus often taking up a slot in DoEs, when more relevant effects could be studied instead. Furthermore, a reduction of process characterization steps is always a strong driver in industry. In both early and late development, temperature or pH shifts[28] are introduced, because they can improve productivity. The previously optimized feeding regime is not directly scalable as these variances inherently affect culture metabolism (i.e., by over- or underfeeding), may extend cultivation time, change the feeding pattern and potentially introduce variety in the desired target product quality. Finally, bolus shots themselves are a source of variation as parallel bioreactors will not run at exactly the same cell counts over a time profile, even though in some experiments there may be good reproducibility between reactor vessels,[29] thus the timing for triggering bolus shots will not be the exactly same, which is a problem in parallel cultivations. Adaptive strategies have been implemented to alleviate some of these challenges; a feed rate should increase with cell count, so the most obvious adaptive strategies are based on parameters which allow the estimation of biomass,[30–32] among other more simple strategies such as using historical profiles[3] or an increase of feed with increased vessel volume if no biomass estimation is in place. In a simple scenario, a fixed amount of substrate can be assumed to be consumed by the cells per time. Knowing the cell count, this amount is then replenished and added in the form of feed.[33] However, biomass estimation will come with an error, and the fixed amount that cells consume per time interval may not be constant during the time course of the fermentation, therefore bolus shot safeguards may still be required.

Suggested approach

In this contribution, we established a control strategy for sensors that may have a significant measurement error (>50%), and demonstrate a robust strategy for process control purposes. In the center of attention are those sensors which measure biomass and substrate level and their respective errors, independent of what kind of underlying mechanism was used, i.e., turbidity,[31] capacitance,[34–36] offgas analysis,[37] online analyzer,[38] disposable sensors,[39] RAMAN,[40–42] NIR/MIR,[43–45] FACS,[46] in-situ microscopes[47] among many others.[48,49] A mathematical environment was established in Matlab where continuously estimated online biomass, continuously estimated metabolic state, and discretely measured glucose concentrations are used as inputs in a metabolic model. In this model, lactic acid production or consumption can be described via a clone-specific yield parameter Y L/G (see further below), which can be calibrated to a given clone in early screening experiments, while the specific glucose consumption rate depends on the available glucose concentration (or alternatively can be set to particular values with a user-defined level of uncertainty). The control algorithm aims at feeding a stoichiometric feed rate in real time. This allows holding glucose concentration close to a desired set-point. This set-point is specified in advance and held at all times within a certain, user-defined range. Our algorithm determines the worst-possible deviation by applying the cumulative highest error in both positive and negative sense and suggests appropriate corrective action by simply changing the feeding rate instead of adding large amounts of concentrated feed (Figure 2).

Several checkpoints were integrated into the strategy during development to ensure close representation of a real process:

i. Different, real biomass profiles established in process development and manufacture were used as input for the online estimated biomass profile.

ii. A mechanistic relationship between GLC concentration and specific glucose consumption rate was built into the model and coupled with a measured historical lactic acid per glucose yield (YL/G) relationship to simulate a realistic lactic acid trend for the simulations.

iii. The specific glucose uptake rate qs could be set to particular values, instead of having to rely on mechanistically derived values, if prior knowledge was available. In our contribution, both conditions (known qs and mechanistic qs) were successfully tested for robustness.

iv. The metabolic model was connected to a feeding strategy which took regular samples with the help of a sensor with identified error. The offset between the measurements and set-point was used to slightly correct the stoichiometric feeding rate. Because no real samples are taken, we sampled *in silico* values between the maximum possible errors from a noisy Quasi-Gaussian distribution.

v. *In-silico* DoEs were run by varying the selected mechanistic parameters and assumptions simultaneously to find potentially interesting clone behavior and conditions under which our algorithm may struggle to meet the defined control specifications.

Goal

The goal of this study was to develop a process-independent feeding strategy which can immediately utilize the output from novel online-capable sensors even though the level of uncertainty may be high. The challenge was to test robustness and scalability of the algorithm so that a transfer to new processes is possible even though equipment, material, clone or even the control specifications may vary. We demonstrate the functionality of our algorithm on a dataset that was not varied in its estimated model parameters but rather in its fundamental assumptions. To do so, a multitude of conditions was thus tested via *in-silico* DoEs. This approach allowed the exploration of an unusually large set of combinations, which circumvented the requirement of obtaining a very large dataset from wet-lab experiments.

The novelty of this approach is the consideration of the measurement error of the utilized signals to control the culture; optimally, every cell is supplied with exactly the amount of glucose required at any given time. However, such a strategy relies on the knowledge of two important parameters: (i) the amount of cells (viable cell concentration, VCC) in the reactor and (ii) the stoichiometric glucose consumption rate (qs), in almost real time.

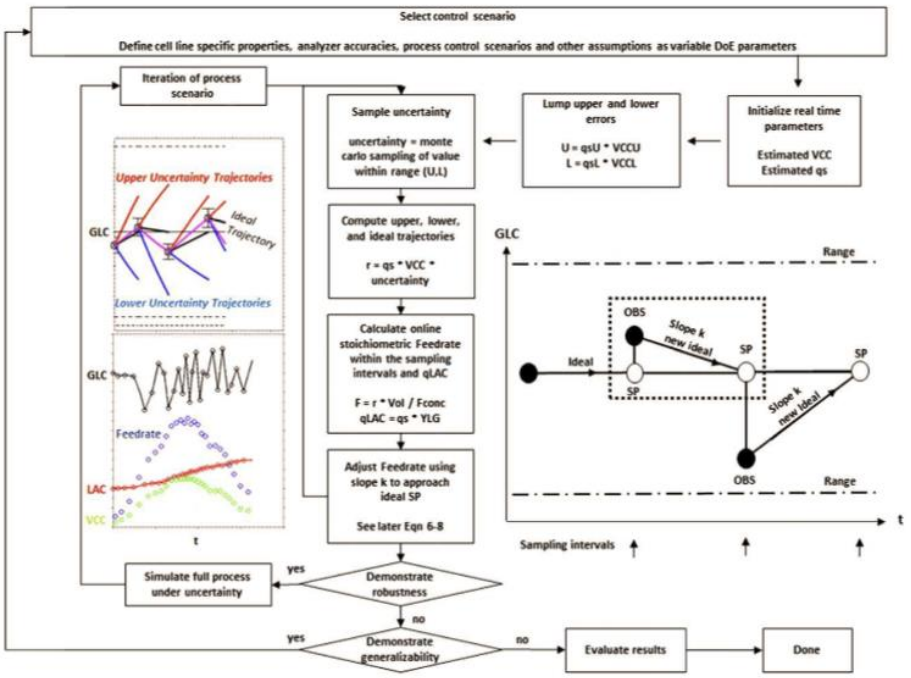


Figure 2. High level overview of control strategy with *in-silico* simulations under uncertainty.

Experimental Section

Materials

Bioreactor data was recorded at Vienna University of Technology (VUT) and at Boehringer Ingelheim Biberach, Germany (BI). A biomass profile together with the metabolic characteristics from one of our CHO clones in process development (2L scale) and manufacturing runs (12,000L) was used in our simulations. A balanced feed platform exists and could be used to supply the concentration of the quantitatively most important carbon source glucose along with all other nutrients stoichiometrically. Alternatively, a separate glucose feed may be set up for the purpose of simply keeping the process within a defined glucose range. The metabolic model was constructed and tested in an *in-silico* environment using Matlab (Mathworks, Natick, MA, USA). Modde (Umetrics, Sweden) was used to construct and interpret a DoE, which was used to suggest conditions for a series of experiments where several parameters were varied at once. The biomass profile of one real process development (2L scale) experiment was used to run 350 *in-silico* experiments where extreme conditions, that are known to affect control performance, were tested rigorously. Additionally, industrial scale data (12,000L) was used to test its performance using a different biomass profile and volumetric input to demonstrate scale-up capability and robustness *in-silico* in two independent simulations in 100-fold repetition. The algorithm did not fail once to meet the specified control requirements in any of the over 550 cases developed.

Process parameters

The minimum amount of data to test the algorithm was any mammalian culture's biomass profile, represented as the viable cell concentration over time together with realistic clone- and equipment-specific parameters as well as the expected error of biomass estimation and measurement in the process.

One of the biomass profiles obtained in wet-lab experiments in our labs was used to fuel a part of the simulations. Since the metabolic state of a culture dictates how much substrate the culture consumes and requires for a stable glucose concentration, this parameter had to be estimated somehow—we chose two paths to do accomplish this—either by (i) historic process knowledge of the typical range in which both the highly variable specific substrate consumption rate qs (and in which range it is converted into lactic acid), and (ii) by first principles, i.e., establishing a mechanistic link between glucose concentration and its uptake by cells. The ratio at which conversion of glucose into lactic acid was observed (Y L/G), was further extended by changing the slope of the relationship. This allowed the *in-silico simulation* of different metabolic behaviors, i.e., as if other clones were studied.

Instead of using a very detailed model with many model parameters, for instance for simulating intracellular chemical gradients, we decided to use a parsimonious model which would still represent the highly process-dependent variable qs[33,50] as close to reality as required. This allowed us to focus on the real task ahead, which was the demonstration of our control strategy rather than a detailed simulation of mammalian CHO cells metabolism.

qs. In a very simple metabolic model representation, qs was dependent on the external glucose concentration (GLC), the maximum substrate uptake rate it could achieve, and its uptake into the cell depended on a cell-specific substrate specificity, represented here as the Monod model parameter Ks,[51] see Eq. 1. In this contribution we use two approaches to estimate qs, one is a mechanistic approach where GLC dictates qs, while the other is an expert knowledge approach where qs is known to fall within a specified range from processes observed in the past (Eq. 2 and Eq. 3 in Table 1). The metabolic model we used is based on the Monod equation, which we extended among other parameters with uncertainties in its glucose set-point. Different glucose sensors may be selected for the task of measuring GLC, which may lead to differences in measurement accuracy, represented here as GLCerror. As substrate S, the glucose set-point (GLCSP) selected and the error of its measurement GLCerror will result in a variance of qs in a non-linear fashion (Figure 3).

GLCerror is the range in which glucose measurements are accurately obtained, and i.e., an error of 4% means that a value of 0.04 is added to the setpoint. It will add to the uncertainty of knowing qs exactly with GLC alone, and will have an upper (U) and a lower (L) value depending on the direction of the measurement offset. This will create a band rather than a single point in which its real value may lie. Several substrate transporters have individual Ks values,[52,53] but it shall be noted that the Ks we used in our contribution is a macroscopic observation and its purpose in our model is to describe clonal variance, i.e., different substrate uptake behavior at the same GLC concentration. The range of industrially encountered Ks may be historically around 0.2 g/L (or higher), which might imply that controlling GLC is relevant for qs control when the glucose levels may decrease below 1 g/L. However, as cell line development progresses, this statement may not hold true for processes of the future. The specified range in which GLC is attempted to be held is represented by GLCrange, therefore in our simple metabolic model, qs will change with GLC concentration within this band, rather than being a one-point minimum or maximum. Alternatively, and very importantly, qs could be estimated from historical data sets to fall within a particular range. This is could be very useful when qs is known to approach a narrow range in a process phase, effectively reducing uncertainty about qs. Regardless which model for qs is selected, the value will definitely vary during the process, resulting in an upper and lower estimate as suggested in Eq. 3.

Table 1. Calculation of qs for Simulated Trajectories in Mechanistic and Expert Knowledge Scenarios

qs Estimation Using the Monod Model	
$qs = qs_{max} \cdot \frac{S}{K_s + S}$	Eq. 1
qs estimated via metabolic model with GLC uncertainty	
$qs_{ideal} = qs_{max} \cdot \frac{GLCSP}{K_s + GLCSP}$	Eq. 2a
$qs_{high} = qs_{max} \cdot \frac{GLCSP \cdot (1 + GLCerror)}{K_s + GLCSP \cdot (1 + GLCerror)}$	Eq. 2b
$qs_{low} = qs_{max} \cdot \frac{GLCSP \cdot (1 - GLCerror)}{K_s + GLCSP \cdot (1 - GLCerror)}$	Eq. 2c
qs estimated via metabolic model dependent on the selected setpoint range	
$qs_{ideal} = f(qs_{ideal}, GLCSP)$	Eq. 3a
$qs_{upper\ range} = f(qs_{high}, \text{upper GLCrange})$	Eq. 3b
$qs_{lower\ range} = f(qs_{low}, \text{lower GLCrange})$	Eq. 3c
qs estimated via expert knowledge within historical range	
$qs_{upper\ range} = f(\text{upper historical range})$	Eq. 3d
$qs_{lower\ range} = f(\text{lower historical range})$	Eq. 3e
$qs_{ideal} = f(\text{setpoint for qs})$	Eq. 3f

Example Calculation. Introducing a glucose range rather than a single set-point with measurement error will further increase the possible upper and lower range for qs, for instance holding 1 g/L with a measurement error of 0% and a range of +- 20% will mean the set-point will be somewhere within 0.8 and 1.2 g/L, and if the measurement error increases, the potential glucose range will increase correspondingly with it.

As can be seen in the simple example in Figure 3, the normalized qs at 1 g/L GLC and Ks 1.0 is 0.5. However, due to the uncertainty about GLC levels, the normalized qs may vary within 0.42 and 0.57. This variance is even more pronounced at lower GLC levels (see Supporting Information) and shows the importance of a stringent set-point control. More information can be found in the supplement.

qLAC. In-house knowledge as well as earlier contributions from other researchers,[33,54,55] confirm that the specific glucose uptake rate correlates very well with specific lactic acid production in mammalian fed-batch processes. According to the

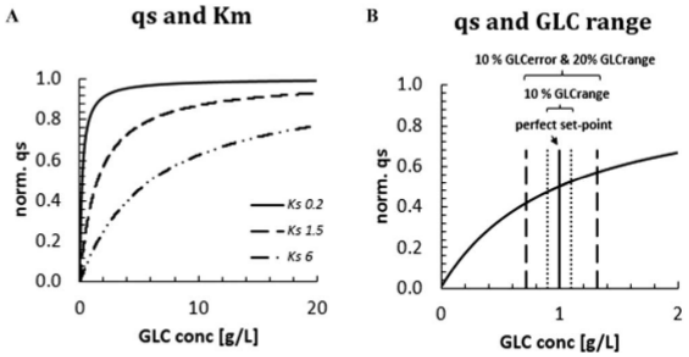


Figure 3. (A): qs vs. Ks at different Ks values (0.2–6), (B): qs at various GLC set-points, with 0–10% GLCerror and 0–20% GLCrange.

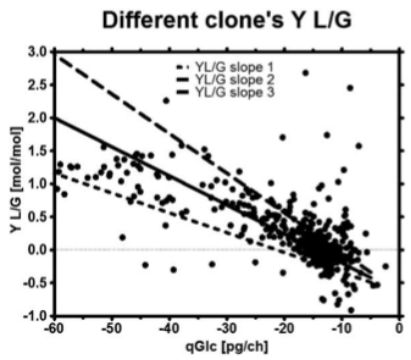


Figure 4. Specific glucose uptake rate is plotted versus the metabolic conversion of glucose to lactic acid.

A different control strategy was already used in wet-lab experiments for setting up and controlling a desired Y L/G of the culture. The clone-specific Y L/G in our experimental data measured in our labs ("Y L/G 2," full line) is plotted as central line. Two extreme metabolic conditions ("Y L/G 3," dashed line, and "Y L/G 1," dotted line), which were not measured in our labs, extended the possible metabolic states which could be simulated *in silico*.

clone's genetic background,[56–58] the yield of GLC to LAC (Y L/G) may be higher or lower at the same qs,[59] which is represented here as a difference in YLGslope, which then yields qLAC (Eq. (4)). For clarity, the qs which is used in the equation is drawn within an interval which is specified as minimum and maximum, and the process of drawing the sample will be described further below.

$$qLAC = qs \cdot YLGslope \tag{4}$$

In Figure 4, we show the yield of glucose conversion to lactic acid (Y L/G). In mammalian cell culture bioprocesses, a low yield is often targeted to reduce lactic acid accumulation which is often achieved by running the process at very low GLC levels.[60–62]

Biomass. Finally, various methods are available to estimate biomass. For the purpose of this contribution, we have selected two biomass profiles to enter our simulations as quasi-online estimated biomass profiles. One was taken from process development experiments in our labs at VUT, the other from a manufacturing scale process at BI. The datasets can be also found in the Supporting Information. The error between the biomass currently present in the bioreactor and its estimation is represented by Xest.

Others. Other process parameters that are required for a feed calculation, for instance working volume, feed concentration etc. were assumed to be known in the process in real time and given reasonable starting values. While some model parameters are approximately or very well known (i.e., the analyzer's error, biomass estimation error), others are difficult to estimate at all, in particular the cell-specific parameter Ks and with it qs, as they may change over time. However, all of them can be assigned plausible ranges so that the feeding strategy could be tested for its robustness in a series of simulation. In this contribution, the desired GLC setpoint range, assigned GLC and Xest errors and observed Ks or YLG values were treated as variables and used as input in our simulations within robustness and DoE studies.

Description of the algorithm with examples

The intention of this work is to develop an algorithm, which is based on uncertainty from glucose measurement, viable cell concentration estimation and metabolic state. Its purpose is to keep the glucose concentration within a previously specified range. The algorithm does this by feeding glucose to a growing cell culture stoichiometrically. This means that if the consumption rate of each cell was known, and exactly this amount was added by feeding in real time, then the change of glucose over time between two successive measurements would be exactly zero because the culture would remove exactly the same amount of substrate that is added. However, cell count does indeed change and neither cell count nor metabolic state are precisely known, and non-reference measurements, which might be very useful for immediate control, i.e online, inline or at-line, sometimes have a considerable error.

Two potential sources of uncertainty for lumped errors are described with the following example: if cell count was underestimated with a stoichiometric feed, i.e., by a factor of 1.1 (10%), then glucose would start accumulating. The same is true if qGLC was not known exactly, i.e., by overestimating the real consumption rate by a factor of 1.25 (25%). In both cases the factor of the base case would have been 1.0. If both effects occurred at the same time, i.e., underestimation of cell count and overestimation of specific uptake rate, then a lumped error must be calculated to account for both effects. This was done by calculating a lower lumped error by multiplying both errors with each other, and using this factor in the calculation of the upper bound of the worst possible expected value, i.e., $1.25 \cdot 1.1 = 1.375$, where the feed rate was 38% higher than it should have been to supply the culture with stoichiometric feed. Both directions of the possible deviation were calculated and show the worst case predictions based on expected errors (and in our simulation, also beyond). For viable cell count, the measurement error could be either the measurement error or estimation error of the method, i.e., 20% uncertainty for Xest (Figure 5). For qGLC, the uncertainty comes from our simple metabolic model, where GLC is not controlled to a set-point but stays within a range, thus giving qGLC also a range to fall into. Alternatively, historical knowledge is very useful to define the most probable value for the whole culture time or for particular phases of the culture together with an upper and lower error (see Table 2 and our industrial case study further below).

One important aspect of this strategy is not to exceed the desired GLC range, which is represented in an example as 1 g/L ± 20% of the glucose set-point. However, the worst-case estimation of GLC in Figure 5D shows a breach of the specified range. The developed algorithm deals with such cases by taking an earlier sample to take corrective action. This is done by correcting the current feed rate (more feed when GLC is below SP, less when above the SP). The worst-case measurement can fall only within the furthest edge of acceptance and lies always within the specified desired range, thus the required specifications are always met by design.

A sample is drawn after some time (here, after 12 h). The black line represents an ideal glucose trend, the blue line

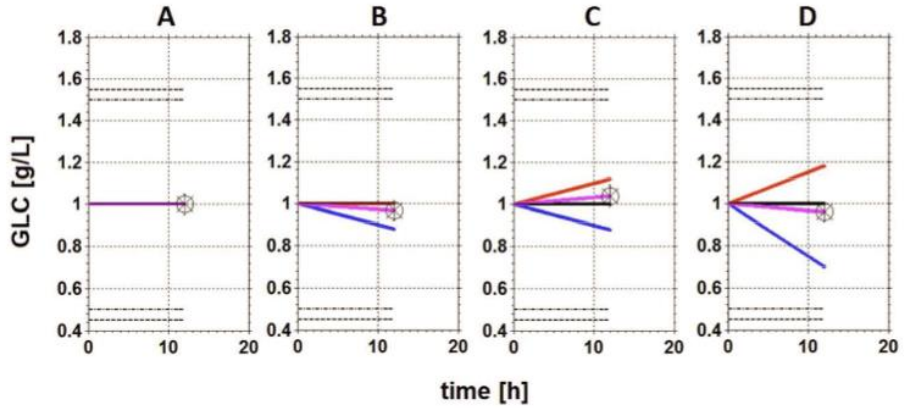


Figure 5. Uncertainty trajectories for predicted GLC [g/L] levels over time [h] after 12 h with active stoichiometric feeding in Matlab: (A) no errors, (B) overestimation of biomass (Xest 50%) and randomly sampled GLC point (x) within this interval (C) both 50% over- and underestimation of biomass (Xest 50%) and randomly sampled GLC point (x) within this interval, (D) over- and underestimation of biomass (Xest 50%), randomly sampled GLC point (x) within this interval, and an asymmetric error for qs, as function of GLCerror and GLCrange which increase the maximum lumped error in both directions.

Color code: red symbolizes an increasing trend for glucose concentration, blue a decreasing trend. Black is the ideal, calculated glucose profile which approached the set-point without errors and purple is the observed (simulated) path, which lies between the two possible maximum uncertainty trajectories. The circle represents the simulated glucose sample measurement.

Table 2. List of Parameters With Ranges Used in This Simulation and Its Dependency on Cell Line, process, Analyzer or Biomass Model

Parameter	Simulation Range	Industrial Conditions	Conditions Set in Script	Depends On
Ks [g/L]	0.2–5	Our clone's estimated Ks: < 2	0.2, 0.5, 0.8, 1, 1.5, 2, 5	Cell line/Clone
YLGslope [-]	−0.06 to −0.03	Our clone's estimated YLG: −0.044	−0.06, −0.044, −0.03	Cell line/Clone
GLC [g/L]	0.2–5	Our process' relevant SP: 2 g/L	0.2, 0.5, 1, 1.5, 2, 3, 5	Process
GLCrange [%]	10–75%	Our process' relevant SP: 50%	0.1, 0.2, 0.5, 0.75	Process
GLCerror [%]	4–20%	Our analyser's error: 5% or better	0.04, 0.1, 0.2	Analyzer model
Xest [%]	5–50%	Our estimate's error: 25% or better	0.05, 0.1, 0.15, 0.33, 0.5	Biomass model

represents a positive lumped error, which leads to a decrease of GLC due to a deviation from the intended stoichiometric feed. The red line represents a negative lumped error resulting in underestimation of both VCC and/or qs, leading to an increase of the glucose concentration. The GLC trend line corresponding to that sampled measurement is represented in purple. As the sampling is randomized within both errors, the simulation will be slightly different each time it is run with identical parameters, however due to its robustness, the outcome will be shown a little later to be very similar.

How to draw a random sample

The "sampled" GLC concentration in this simulation was drawn between both two maximum error ranges following a randomized, Quasi-Gaussian distribution with $n = 3$ degrees of freedom and $k = 100$ number of random draws. The higher n is selected to be (i.e., $n = 10$), the less likely a value close to the tails of a constrained Gaussian is sampled (an example is shown in Figure 6). The chosen sampling distribution allows almost any value, high and low, to be chosen and is the reason for a strong noise in our data. We purposely embraced the strong noise as this prepares our algorithm for the most realistic conditions in which measurements come with a high error.

Typical control action during one run

Corrective Action. If the GLC range was breached (1st and 2nd sample drawn in Figure 7A), corrective action is initiated. As a first step, it matters whether the uncertainty trajectories were still within the upper or lower bands or would already fall beyond the permitted GLC range around the GLCSP. If any of both uncertainty trajectories touched the permitted range, this time point was recorded and passed on to the system to sample earlier to avoid falling outside of the specifications or later to avoid oversampling.

The algorithm adapted the sampling time dynamically by following the following rule; (i) if a breach of the range would occur based on the uncertainty trajectories, the sampling interval was reduced and a sample needed to be taken earlier by a compression factor of 0.66, i.e., for a sampling interval of 6 h this would have been a reduction to sample after 4 h. Alternatively, if no breach occurred and the measurement was within the correct range, then the sampling time was extended from (in this example) 6 h by a stretching factor of 1.33 to 7.8 h (see illustration of the 3rd and 4th sample in simulation Figure 7A). A rationale and reasoning behind selecting those values is provided in the appendix. However, the reason for designing our algorithm to learn from the current sampling requirements was based on our

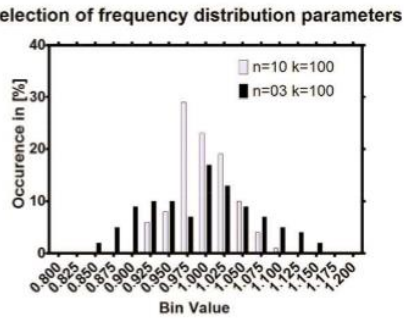


Figure 6. In order to sample between the maximum possible error, a function which created a distribution was required that allowed sampling from, in this particular example, the bin value of lumped errors between 1.2 and 0.8.

A new random distribution with n degrees of freedom and k number of samples was seeded every time that sampling was required, which fulfilled its purpose of adding randomness to the sampling process. A higher n enabled more frequent sampling from the mean, while a lower n allowed for more frequent sampling from the more extreme tails of the *in-silico* sample. In our simulations $n = 03$ and $k = 100$ was employed.

observation that simply keeping the sampling time constant would result in unnecessary oversampling of a well-controlled process——this would be the case at the beginning and end of a process—or undersampling and not guaranteeing that the process is truly within control limits around peak VCC. Increasing or decreasing the current sampling interval by 33% fulfilled the purpose of modifying the static sampling interval automatically without operator interventions and resulted in a robust and elegant solution which worked well for all simulations presented in this work.

After modifying the sampling time, another important control is triggered: the feed itself is very gently corrected by a slope k which either increases or decreases the current stoichiometric rate (Figure 8). First, the difference between the ideal GLCSP and the sampled GLC is calculated as delta (Eq. 5).

$$delta\left[\frac{g}{L}\right] = GLCSP - sampledGLC \tag{5}$$

This delta can be a positive or a negative value. If the value was positive, the matter would be more complicated as the feed rate would have to be reduced as there is no way or removing the excess glucose. We decided to walk a different path by adapting (in this case reducing) the feed rate slightly instead of introducing metabolic perturbations which could affect our process negatively: this is done by calculating a

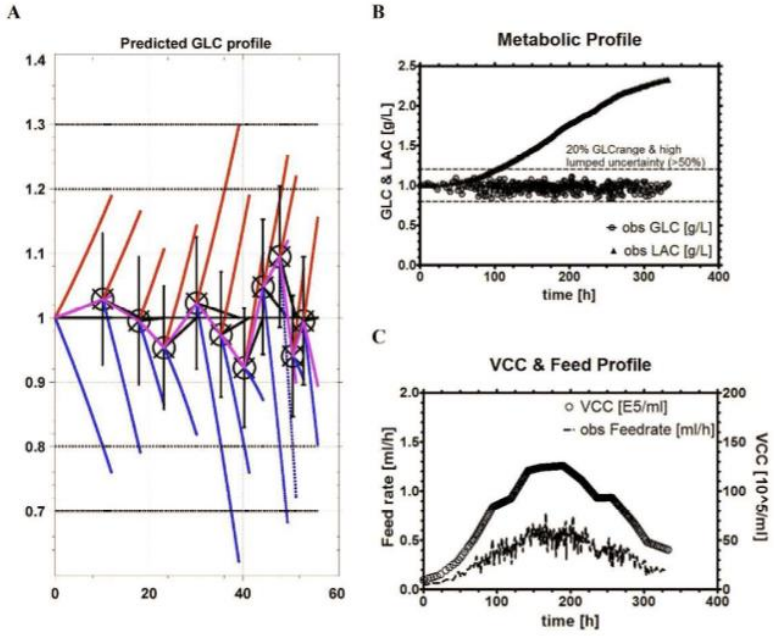


Figure 7. Exemplarily, the feeding algorithm is shown in its first 10 iterations (A) and in full length (B-C).

A mechanistic link between qs and GLC is assumed and the observed, noisy qs lies in average around −20. The observed change in GLC and LAC are shown using the aforementioned YLGslope relationship between qs and qLac **(B)** and the observed stoichiometric feeding rate, which was calculated for the estimated biomass (VCC) and plotted in **(C)**.

corrective factor k (Eq. (6)), which is delta divided by the sampling interval to obtain a linear trajectory from the measured GLC point to get back to the desired GLC set-point within the calculated sampling time as slope k that acts correctively on the stoichiometric feedrate (Eqs. (7) and (8)).

$$Corrective\ k\left[\frac{g}{Lh}\right] = \frac{delta}{samplinginterval} \tag{6}$$

$$Stoichiometric\ Feedrate\left[\frac{mL}{h}\right] = \frac{VCC\left[\frac{c}{mL}\right] \cdot Vol\,[mL] \cdot (-qs)\left[\frac{g}{ch}\right]}{Feed\ concentration_{GLC}\left[\frac{g}{mL}\right]} \tag{7}$$

$$\begin{aligned} &New\ stoichiometric\ Feedrate\left[\frac{mL}{h}\right] \\ &= \frac{VCC\left[\frac{c}{mL}\right] \cdot Vol[mL] \cdot (-qs)\left[\frac{g}{ch}\right]}{Feed\ concentration_{GLC}\left[\frac{g}{mL}\right]} \\ &+ \frac{k\left[\frac{g}{mlh}\right] \cdot Vol\,[mL]}{Feed\ concentration_{GLC}\left[\frac{g}{mL}\right]} \end{aligned} \tag{8}$$

If this corrective factor was zero, both the measured GLC and GLCSP were the same value and the feeding rate would continue perfectly horizontally from the currently measured GLC with a stoichiometric feed. However, once measured GLC would not be identical with the GLCSP concentration, without a correction of the feed rate, the trend would also continue horizontally, therefore the feed would require some sort of modification to get back to the desired GLCSP. In other words, k is an important parameter and a means to get closer to stoichiometric alignment. It corrects our assumptions of stoichiometric feeding, and adds the extra grams of glucose (in form of mL/h of concentrated feed in a control application) when the measurements is below the set point, or decreases the feed slightly, if too much glucose was measured. The feed rate increases when the measured offset is below the SP, otherwise it decreases. If there is no offset, the feedrate does not change. The modification can be considered a variable to add or remove from an otherwise stoichiometric feed of GLC.

Typically what is done to get back to the GLCSP is to add a bolus shot, which would be a step change in the GLC concentration profile and continue feeding the same amount of substrate. This strategy would result in GLC values always below the set-point and would require very intense sampling during peak cell density. When the measured GLC concentration is above the set-point, there is currently no convincing mechanistic corrective action; if the feed rate would be for instance halved, or reduced by 10%, or even turned off, starvation might be risked. If no action is taken, GLC accumulation might continue. In either case, such a strategy will definitely result in fluctuations of the GLC profile, which we strive to avoid here. The example chosen uses a very high error on biomass (50%) in combination with qs uncertainty and a GLC measurement error of 10%, which makes the observations noisy. We have deliberatively chosen such an extreme example to illustrate the mechanism behind the algorithm we are using and to demonstrate that our algorithm can cope with poorly estimated input variables (i.e., VCC, GLC measurement, metabolic state).

We wish to address the rising need for control strategies once adequate sensors are available and in place. Assuming that VCC can be already estimated continuously[33,63] and that glucose can be measured in discrete intervals by a sensor (i.e., reasonably not less than 30 min but this shall not post a restriction for our control algorithm), only qs needs to be estimated. Candidates for potential glucose sensors can be found in a number of articles specialized in this field,[64–67] and a small selection which we consider to be relevant for industrial control purposes is listed in the appendix (Appendix 1). The metabolic state may be assessed either by calculation with a model using online information such as, i.e., glucose and lactic acid concentration, or by estimation, i.e., knowledge of clone behavior over time. We may assume that it follows a first-principle relationship, for instance a very simple Monod-kinetics relationship between GLC and qs. Alternatively, qs could be also known by experience to lie within a particular range (typically between −8 and −60 pg/cell per hour, based on our own observations in mammalian CHO cell culture). Both scenarios were discussed above. The knowledge of Y L/G may be already well known for particular cell lines or clones. Any method, mechanistic or historical knowledge, could be used to determine and potentially confine the specific glucose consumption rate for the simulation within a particular range. We tested both cases independently of each other in our simulations. Finally, qs might be also calculated during the fermentation and assumed not to change its value until the next measurement when it needs to be re-evaluated. From these possibilities to deduct qs, we focus in this contribution only on the first two as the latter one introduces another element of variability due to the nature of derivative calculations.

Results and Discussion

Reviewing critical assumptions in a DoE-aided simulation

With our suggested control strategy we have exemplarily shown how to keep glucose within a particular range, yet this was done only under very particular conditions. Typically, the conditions and parameters one may be interested in to set up may vary, and with it the necessary amount of control. This means that the amount of sampling to successfully hold glucose at the desired set-point may be very high and therefore uninteresting for industrial purposes. For this reason, we tested several cases (see Table 2) in a DoE approach to determine under which circumstances our control strategy could unfold its fullest potential. The designed DoE as well as the simulation results can be found as Excel template files in the Supporting Information.

In industrial bioprocesses, glucose measurements are currently not taken continuously but in discrete time steps. This means that the time until a sample is taken, analyzed and automatically recorded must be considered.

In an earlier contribution, our research group evaluated the sampling interval by the signal (biological activity) to noise (precision) ratio of the measurement. These elements were established in a different simulation and resulted in a recommendation of 2 h for the sampling intervals using a state of the art analyzer or 24 h for a membrane-based device. That simulation used a peak cell density which was roughly five times lower than the peak cell density in this contribution, therefore it made no sense to measure even more frequently when a bolus shot strategy was employed.[68]

Here we established a continuous feeding strategy which countered the glucose decrease by a stoichiometric feed to achieve a more consistent glucose concentration profile between samples. The ideal sampling interval using a continuous feeding strategy depends mainly on cell concentration,

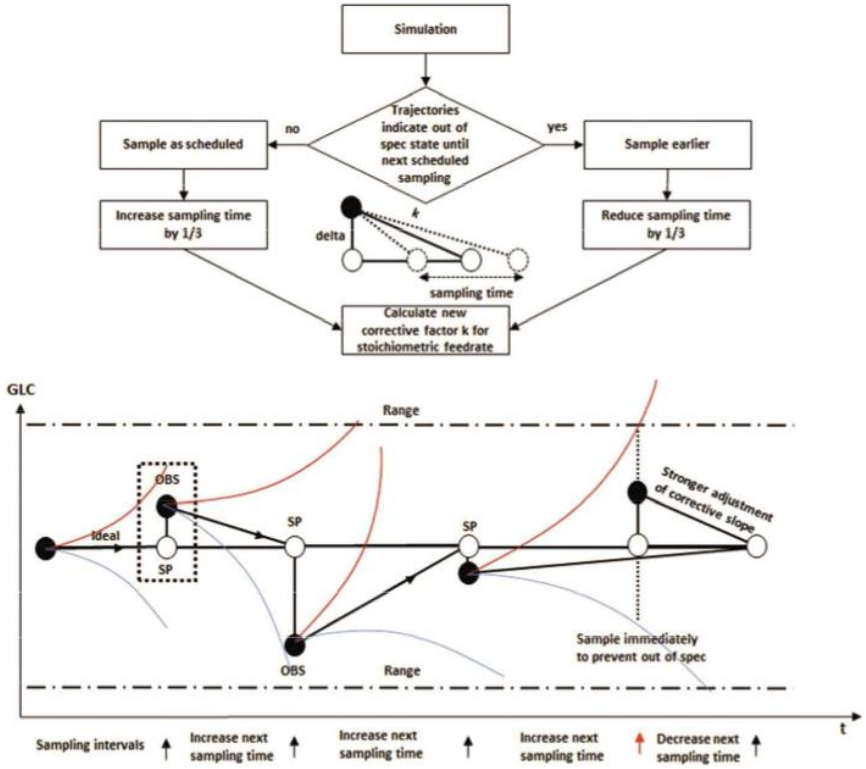


Figure 8. Automatic control action in simulation to prevent out of specification.

the metabolic demand, and the range within which the glucose concentration is to be held. Because both metabolic demand and viable cell count were assigned uncertainties in this contribution, the necessary sampling interval is variable and tailored to the process specifications and its expected errors.

There are practical limitations concerning the maximum frequency of the sampling interval. For instance, a sampling interval of 30min corresponds to a maximum of 672 samples taken in a 336 h process. This is still a rather large amount of measurements when compared to an industrial process with often less than 30 sampling events and could be in practice too high to truly pay off, although with new sensors, this could quickly change. Some analyzers might not be able to acquire the signal in such frequent time intervals because the analysis steps take time, even though they are automated; cleaning/priming lines, multiple analytes instead of just glucose, automatic re-dilution, recalibrations, and other analyzer-dependent variables can impact the total amount of time until the signal is logged and effectively translated into a control action. Therefore, even though it is technically possible, we restricted the evaluation of our DoE to scenarios with less than 300 sampling events. For a sensor that may be capable of measuring on a minute basis, this is a rather low number, while for membrane-based analyzers, the reactor might be "sampled dry" with such a large amount of samples withdrawn, depending on how much sample volume is necessary for analysis. Cases in which qs is in average below −5 pg/ch (picogram per cell per hour) are uninteresting for further investigation and are removed from the dataset, as such a low case of metabolic activity would mean the cell line might be growing very slowly, perhaps too slowly in comparison with CHO cells. As product formation is mainly a function of cell count, we currently focus on cells which show a strong growth, thus high substrate uptake, as these are among other features the characteristics of industrially relevant cell lines.

The statistical model yielded model coefficients for each parameter, which were used to describe the target variable, in this case the sampling number (Figure 9). The highest p-value was used to remove non-significant parameters from the model one by one until all parameters were significantly contributing to the model fit. In general, a high GLCSP, a high range for this set-point and a better biomass estimation decreased the sampling intervals. Control at a very low GLCSP and a very stringent control of the glucose range was possible but came with a higher sampling number cost. A high Ks in our model yielded interesting results. The idea

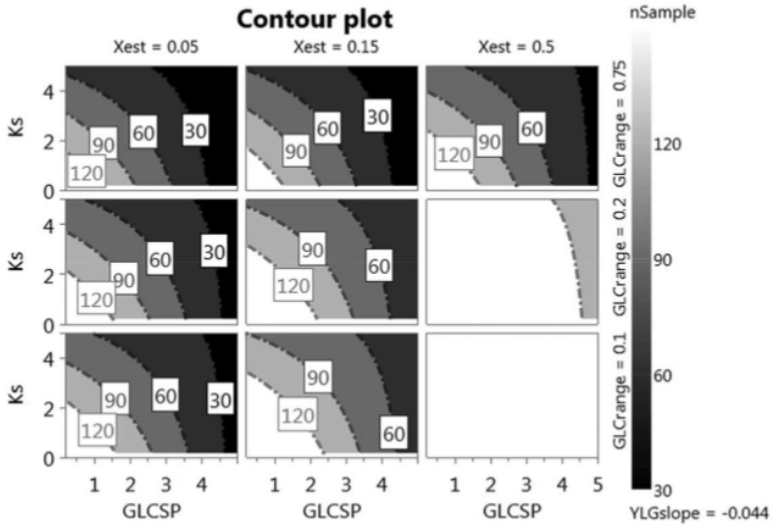


Figure 9. Model fit for two Principal Components (Comp1 and Comp2): $R^2 = 0.72$, $Q^2 = 0.65$, for more details on coefficient contribution see the Supporting Information.

Contour plot of target variable (sampling number) in relation to glucose set-point (GLCSP), Ks, glucose range (GLCrange) and biomass estimation (Xest) simulating a cell line with a YLGslope of −0.044. White areas indicate more than 150 sampling events while black areas require less than 40 to control the desired GLC SP in a specified range. Higher glucose set-points are in general easier to hold, as well as areas in which Xest is good. For control purposes, a high Ks at the same GLCSP leads to easier process control, but in practice the cell line's lumped Ks may not be so high in typical industrial cell lines. The model fit with an example can be found in the appendix.

was to simulate how a cell line would behave in a glucose-rich environment. As a high Ks goes hand in hand with a low qs, the reduced amount of glucose taken up by the cells decreases the need to feed very fast, resulting in better control of a process.

Robustness with a mechanistic qs approach

Next, simulations were run under one of the conditions from the DoE. As mentioned, this control strategy has a certain element of randomness. This happens because not actual GLC samples were taken, but a random sample was drawn from a distribution which can be anywhere between the highest and lowest lumped error in a simulation. To demonstrate robustness, we ran 100 iterations of a full fed-batch and summarized how our algorithm coped with the high variation. Our targeted qs was −30 pg/ch but could vary in a mechanistic relationship as GLC constantly changed. The metabolic behavior was selected to resemble a clone in its early stage of growth on fresh medium, without perturbations or shifts in its environment, which could affect its lactic acid behavior (Figure 10). The chosen range of glucose was very large, because the range which was attempted to being controlled lied between 1 ± 0.5 g/L, which allowed qs to vary considerably between 143% and 53% of its original set-point and made it interesting for our robustness study. For more information about this experiment please refer to the supplement.

Naturally, the feed rate (Figure 10A) increased along with VCC, while the control strategy kept GLC within its specified borders. GLC was randomly sampled, therefore the observed qs was rather noisy due to a derivative calculation. Because qs was rather stable, so was the observed LAC accumulation. As VCC increased, the power of the corrective action increased (Figure 10F), because more cells could change the environment faster. Tighter control of the GLC level by more sampling events was a natural consequence. This was also true if GLC levels increased, as the mechanistic model connected high GLC with a high qs.

Robustness with a known qs range including a metabolic step change

Some cell lines display a behavior that cannot be explained with simple Monod kinetics. For instance, GLC level may stay constant, while the cells consume less and less GLC. We know from historical runs that after a high qs phase of approximately 4 days, qs assumes a relatively constant value.[33] This information is useful because it could be fed into the control strategy to improve control of the process using *a priori* experience. All that needs to be done is to specify two ranges for qs error, one for the higher qs and one for the lower qs. In the beginning of the process, qs is high, yet there are just a low number of cells that could reduce the glucose concentration in the bioreactor, therefore not much corrective action is required. In the later phases, there are many cells which have a reduced glucose consumption rate, and those are actually easy to control because qs is almost constant. So what does that mean for our strategy, can it cope with step-like changes? We asked ourselves this question and tested our algorithm with the following case:

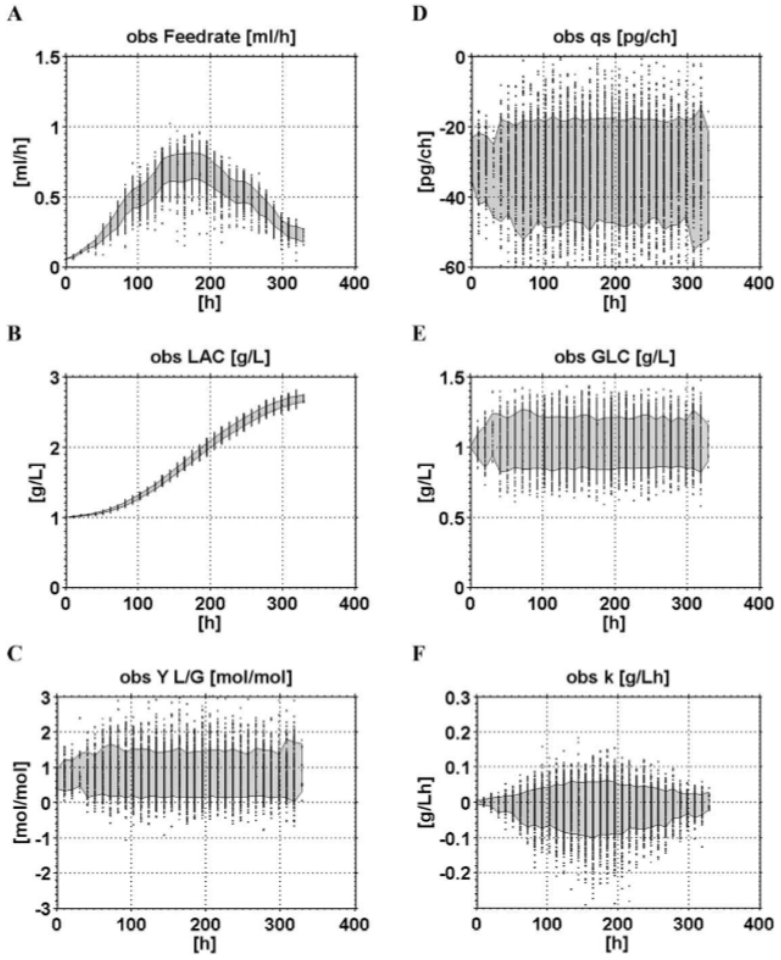


Figure 10. All Figures: black scatter points: original measurements binned at particular time-points; grey band of 90th and 10th percentile of all 100 simulations taken at binned time points.

(**A**): Feeding rate, (**B**): predicted LAC profile with a given YLGslope, (**C**): observed Y L/G for these conditions, (**D**): observed qs when the rate was calculated from a randomized GLC sample, (**E**): observed GLC concentration, (**F**): required corrective factor *k* of the current feed rate in absolute units. As can be seen, qs, the yield and GLC were subject to constant variance for 100 independent simulations, while the feed rate and the correction factor that were required to keep a dynamic culture within specifications were changing together with cell count.

We assumed that our targeted qs of −40 pg/ch lied between −20 and −60 pg/ch for up to four days, a typical observation in our cultures, afterwards qs lied ideally around −12 pg/ch between −6 and −18 pg/ch with a qs error of ± 50% and a biomass estimation error of ± 20%. Often particular substrate depletion events,[62,69,70] pH[71–73] and temperature adjustments or shifts[74,75] occur in industrial processes which aim at a decrease in LAC accumulation for an increase of culture longevity and productivity. We did the same *in silico;* we assumed that qs changed, but not necessarily because of GLC. As lactic acid metabolism is described by a yield relationship with qs, LAC accumulation should change when qs switches suddenly a low level, as will be shown below (Figure 11). From a practical point of view, a metabolic switch of qs is not a trivial problem for feed formulations; when only one 'balanced' feed is employed, the reduced need to feed stoichiometrically with glucose will result also in all the other components to be fed at a reduced rate, which might lead to undesirable multiple limitations. A practical solution would be to use a separate GLC feed to control the GLC level and using multiple feeds to deliver other required nutrients to the culture.

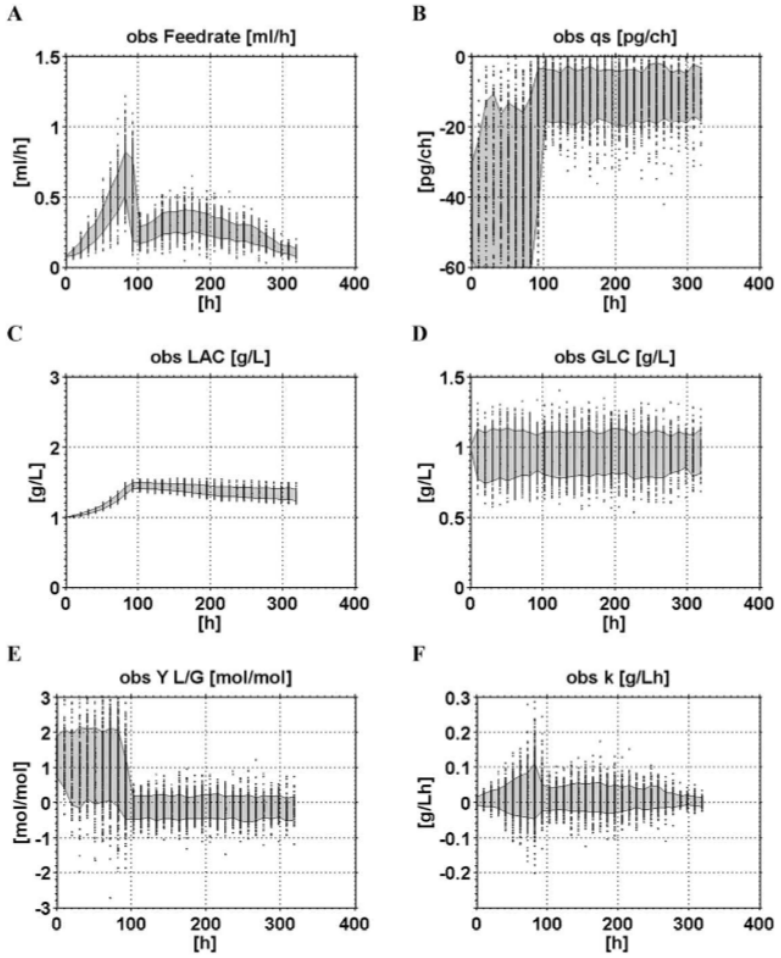


Figure 11. All Figures: black scatter points: original measurements binned at particular time-points; grey band of 90th and 10th percentile of all 100 simulations taken at binned time points.

(**A**): Feeding rate, (**B**): predicted LAC profile with a given YLGslope, (**C**): observed Y L/G for these conditions, (**D**): observed qs when the rate was calculated from a randomized GLC sample, (**E**): observed GLC concentration, (**F**): required corrective factor k of the current feed rate in absolute units.

As can be seen, qs, the yield and LAC undergo through a shift in qs, which had an impact on the required stoichiometric feed rate and the observed control action, but not at all on GLC concentration (as this effect may be elicited by other factors, such as temperature decrease etc.). GLC could be still well controlled in 100 independent iterations, while confinement of qs resulted in a favorable and constant lactic acid profile which may lead to more efficient substrate utilization and productivity.[76]

Industrial case study

In the final example, industrial manufacturing data at full scale (12 000L) was used as input (Figure 12); the objective is to control GLC within 4 ± 2 g/L as good as possible with the following level of uncertainty: online biomass accuracy (Xest) is within 20%, GLCerror 5%, and qs was known to be within -20pg/ch $\pm$ 50%. VCC and Volume data for this case was provided by BI (see Supporting Information). As the manufacturing data was not generated with our suggested feed rate and did not sample the same GLC values that were sampled in these 100 iterations, the obtained profiles are to be understood to represent a simulated output based on our inputs. We decided to add those profiles nevertheless as if these values were observed in the manufacturing runs allow comparison of all observations with each other.

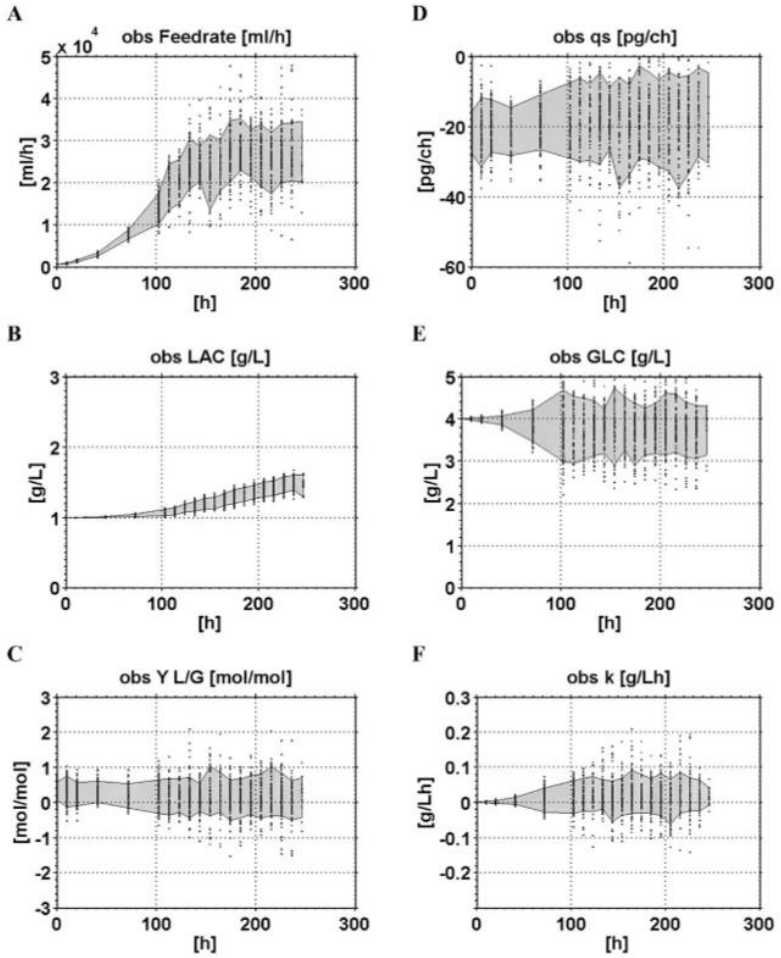


Figure 12. All Figures: black scatter points: original measurements binned at particular time-points; grey band of 90th and 10th percentile of all 100 simulations taken at binned time points.

Manufacturing scale data served as input for simulations where the goal was to control GLC within 4 ± 2 g/L as good as possible. Note: fewer sampling is required when the setpoint range is chosen to be relatively large, and therefore easy to control within its boundaries (see Table 3).

As can be seen in Figure 12E, the high GLCSP and range allow the algorithm to control the set-point very well, especially in the beginning of the process. The largest expected uncertainty trajectories are actually very small which suggests that a process simply would not require frequent sampling so early after inoculation. In average, a sample is required every 15 h, which would make this strategy eventually even suitable for offline-obtained GLC samples, although the suggested concept aimed at utilizing online measured samples for control purposes. Such an application might be for instance a much more stringent control requirement; the same process was further tested with a glucose corridor of 1 ± 0.5 g/L and resulted in an average required sampling frequency of 4 h (see Supporting Information for more details).

Interpretation of results

All values were within 95% confidence of being in range of the mean and its standard deviation (Table 3). In our simulations testing robustness, LAC end concentration could be reduced by ca. 50% from 2.7 to 1.33 g/L by applying a split-feeding strategy in which qs could be affected directly. Y L/G was in average around 0.66 over the whole process

Table 3. Comparison of Target Parameters of Interest

	LAC End [g/L]			Obs. YLG [mol/mol]			Obs. qs [pg/ch]			Obs. Sampling n [−]		
Case	μ	SD	Within	μ	SD	Within	μ	SD	Within	μ	SD	Within
Mechanistic	2.70	0.04	95%	0.657	0.053	95%	−29.49	1.21	95%	94.17	3.85	95%
Split-range	1.33	0.061	95%	−0.052	0.035	95%	−13.35	0.79	95%	73.93	2.43	95%
Industrial 4 ± 2 g/L	1.54	0.086	95%	0.215	0.061	95%	−19.43	1.38	95%	17.15	1.01	95%
Industrial 1 ± 0.5 g/L	1.55	0.05	95%	0.184	0.045	95%	−18.72	1.02	95%	62.45	1.97	95%

Statistical robustness is demonstrated for 100 repetitions per case.
$N = 100$ per case, all values fall into 95% within $\mu \pm$ SD.

time in the mechanistic approach, and in average around 0 during the split range (higher YLG before, lower YLG afterwards). LAC could increase in only 4 days to 1.5 g/L, and would have continued to rise without the split-range intervention. It also shows the importance of controlling qs and with it lactic acid production early on,[55] for instance as soon as all residual GLC from the inoculation culture is consumed. After the shift, qs could be reduced from −29.5 to −13.35 pg/ch which is a decrease to ca. 45% of its high initial value, which resulted in a reduction of required sampling to hold GLC from 94 to 74 which is roughly 22%. From an industrial perspective, the required sampling could ideally coincide with regular IPC (In Process Control) measurements, which are taken off-line manually on a daily basis. However, if the sampling should be based on the risk of breaching a defined set-point with uncertainty trajectories calculated in real time, then the required measurements are unlikely to be available in worker-friendly time intervals, unless cell count is low or the control range is very broad. Optimally, sampling itself does not require operator intervention at all but is executed automatically by the availability of a not necessarily very accurate GLC sensor. IPC could be then used to verify the satisfactory state of control, without having to act immediately to close a control loop. Compared to our process development data, the industrial process is shorter (260 vs. 340 h), cell counts differ (peak viable cell density 100 vs. $125 \cdot 10^5$ cells/mL), and the control scenario is a different one (better Xest, lower GLCerror, higher/lower GLCSP) which makes direct comparison only possible within both cases (mechanistic vs. split-range, industrial cases). Due to the nature of the simulations, all metabolic profiles are generated *in-silico*, as the run was iterated with random GLC measurements every time a 'sample' in the industrial process was required.

Conclusions

Summary

The goal of the study was to develop a feeding strategy that allowed glucose to be held within a customizable range in a dynamic fed-batch culture while considering uncertainty in different forms: variations in clone metabolism, process range, analyzer accuracy, quality of biomass estimation, and non-mechanistic behavior of qs. We have shown for the first time that a dynamic fed-batch process can be easily controlled even if the metabolic state (qs) and the estimated biomass are only roughly known. This was possible by calculating uncertainty trajectories and resorting to automatic sampling shortly before that worst case would occur. The feeding rate could then be corrected before the process would go outside of its specifications with the trade-off that more frequent sampling was required. Some sensors however do not require physical sampling at all, which makes this strategy so attractive. Even though the sensors might come with a lower accuracy than the conventional equipment, where offline samples are required, the online availability of the signal might be a much more important factor in designing a suitable specifications in a control loop. We found useful operational ranges for our control strategy for both cases, which allows very robust operations under uncertainty considerations in biological systems and allowed us to screen robust control specifications which could be used as starting point for wet-lab experiments (Figure 13).

Robustness

Robustness of the control strategy was demonstrated; first, the initial assumptions were varied in a DoE approach to verify that the algorithm still managed to control the process. All results were within their specified set-point range as this level of control was built into the process by design. As a benefit of running the DoE at different conditions, estimates for typical sampling number, metabolic state, yield, lactic acid end concentration and feeding rate could be obtained and can be found in the supplement. Even a highly variable qs could be shown to be controllable within the desired range. Finally, the algorithm was shown to cope reliably with a sudden metabolic shift, where a large step-change type of disturbance was introduced, which would not have been easily possible by ordinary set-point controllers such as PID controllers. Furthermore, the oscillating behavior of such a controller around the set-point could not be forced to fall within a specified range, unless the optimal PID settings were used. This might change with equipment/clone/desired range as tested in our DoE. In contrast, the herewith suggested approach was tested with data under industrially relevant conditions. Robustness was achieved in all cases and this algorithm met all of the imposed control specifications.

Observations

- The precision of the sensor that will be ultimately used did not have a big impact on this control strategy. The only thing that changes with high imprecision is that the possible range band around a particular set-point increases in both directions, making it only harder to hold 'a' particular concentration because every single measurement has already high variance.
- Different clones may be easier to control when the Ks of the cells is rather high because of a decrease in the cell's consumption rate, compared to typical clones, which may have a high consumption rate at the same concentration of substrate, a fact that was not yet explored much in industry.
- Our observations were made in high cell density CHO cell cultures (qs −40 to −10 pg/ch at $>100 \cdot 10^5$ VCC/mL), however, there are also other cell lines and clones which

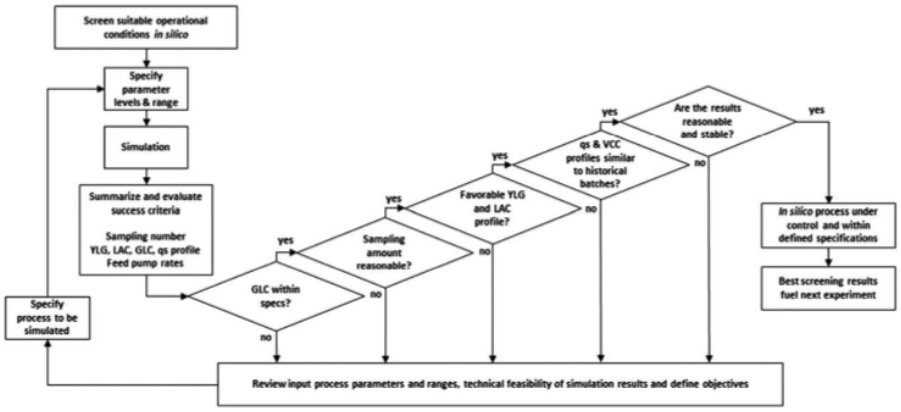


Figure 13. Iterative screening for suitable process control specifications until desired results are obtained for a given problem statement *in silico*.

may have markedly different requirements and benefit from our proposed strategy. Some requirements may be actually easier to fulfill when dealing with non-CHO cell cultures. For instance, Ks, peak VCC and qs may differ by an order of magnitude, which results in easier process control, regardless of high uncertainty. Examples of such cases may be, i.e., human mesenchymal stem cells (qs around −40 pg/ch at $2 \cdot 10^5$ cells/mL[77–79]) or insect cells (qs around −10 pg/ch at $40 \cdot 10^5$ cells/mL[54,80]), which do not have the strong glucose metabolism or do not reach the high cell densities that CHO cell cultures do.

- What may be a little surprising is that even feeding stoichiometrically with a biomass estimation that is usually thought as being terribly inaccurate (50% error) can lead to perfectly stable glucose profiles when this seemingly bad measurement is combined with a sensor signal with similarly high errors (50%) because uncertainty trajectories enable very subtle corrections of the current control output.
- The only 'cost' of using low signal quality is the need for more frequent sampling, which might not be a problem at all if the sensor is a probe or a disposable patch which is in-line or in-situ of the process stream. This could have huge implications for researchers in different fields because it means that even though they may strive for developing the best possible sensor product, the seemingly low precision at hand may be already good enough for current control scenarios.
- Based on our simulations, we conclude that even if model quality of sensor equipment currently employed in pharmaceutical industries could be better (i.e., some of the spectroscopy-based methods, but also novel in-situ sensors), it may be enough for the current problem statement. Cell culture applications ranging from late stage processes at full scale to novel process applications in R&D may benefit equally from an algorithm to keep glucose within defined boundaries.

Benefit and outlook

There is a great need in industry to reduce the amount of necessary pre-calibrations, and this thought can be extended to necessary parameter estimation when first-principle models are employed. In this contribution an algorithm is presented which is intended to directly use sensor signals with their errors to calculate a robust stoichiometric feed rate that ensures that glucose stays reliably within a narrow control range by design. The core of the proposed method relies on first principles, where approximate knowledge of two parameters, estimated biomass and estimated metabolic state, can be very well extrapolated to new conditions for control purposes using uncertainty trajectories. During scale-up and technology transfer, the conditions, equipment and the precision sometimes undergo considerable changes, which makes it challenging to properly transfer an established process to production scale. Some of these conditions were turned into modifiable variables and tested rigorously in a large range using a DoE approach to ensure robustness and scalability of the algorithm, which did not even once fail to stay within specifications. The algorithm itself can be considered to be generalizable for immediate application with novel sensors and processes in a plug-and-play manner for which control strategies have not been written yet.

We showed the high reward when inherently high uncertainties in biological systems are embraced instead of ignored, as one of our most important finding is that perfect control of a bioprocess in narrow operational ranges is possible even when equipment was not very precise. This was possible because a combination of imprecise measurements was better than no measurements at all. Our proposed control algorithm was shown to detect and slightly correct when a process was going out of specifications in real time and we hope that this contribution and the methodology behind it will stimulate researchers to develop even more sophisticated control strategies for manufacturing processes in the near future.

Author Contributions

V.K. designed, planned and performed the experiments at the Vienna University of Technology, built the model and designed the algorithm, set up, analyzed and interpreted the DoE, analyzed the results, performed statistical analysis, prepared Tables, Figures, additional files, drafted and wrote the manuscript. C.C., M.M., and J.B. shared valuable insights

about feeding strategies in an industrial environment, provided manufacture scale data to challenge the developed algorithm and critically reviewed the manuscript. C.H. reviewed and improved the manuscript, supervised, designed, reviewed and helped to conceive the study by supporting a research grant application. All authors read and approved the final manuscript. Partial financial support from a Scholarship from Vienna University of Technology to conduct the *in-silico* experiments, is gratefully acknowledged.

Abbreviations

μ = Growth rate [1/h]
CHO = Chinese Hamster Ovary
Comp1 = Principal Component 1
Comp2 = Principal Component 2
CPP = Critical Process Parameter
delta = Difference between glucose setpoint and measured glucose concentration [g/L]
DoE = Design of Experiments
F = Feedrate [mL/h]
Fconc = Glucose feed concentration [g/L]
GLC = Glucose concentration [g/L]
GLCerror = Analyzer-specific error of GLC measurement [%]
GLCrange = Process-specific range around the glucose setpoint [g/L]
GLCSP = Process-specific glucose setpoint [g/L]
IPC = In process control
k = Corrective factor k in g glucose per liter and hour [g/Lh], application in the same units as F [mL/h]
Ks = Clone-specific Monod constant for glucose affinity [g/L]
L = Lower range
LAC = Lactic acid concentration [g/L]
N = Number of observations
OBS = Observation, here: simulated observation
ODE = Ordinary Differential Equation
PID = Proportional-integrative-derivative controller
Q^2 = Variable prediction coefficient
qLAC = Specific lactic acid uptake/production rate [pg/(cellh)]
qs, qGLC = Specific substrate uptake rate [pg/(cellh)], here: glucose
r = Rate [g/h]
R^2 = Regression coefficient
S = Substrate, here: glucose [g/L]
SD = Standard deviation
SP = Set-point [g/L]
spl = Samples
U = Upper range
VCC = Viable Cell Concentration [cells/mL]
Vol = Volume [mL]
Xest = Model/Analyzer-specific biomass estimation error [%]
Y L/G = Yield lactate per glucose [mol/mol]
YLGslope = Clone- specific yield lactate per glucose [mol/mol]

Literature Cited

1. Hayter PM, Curling EM, Baines AJ, Jenkins N, Salmon I, Strange PG, Tong JM, Bull AT. Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-gamma. *Biotechnol Bioeng*. 1992;39:327–335.
2. Wlaschin KF, Hu WS. Fedbatch culture and dynamic nutrient feeding. *Adv Biochem Eng Biotechnol*. 2006;101:43–74.
3. Lin H, Bezaire J. Pre-programmed non-feedback controlled continuous feeding of cell cultures, WO2013040444 A1, 2013.
4. Yuk IH, Zhang B, Yang Y, Dutina G, Leach KD, Vijayasankaran N, Shen AY, Andersen DC, Snedecor BR, Joly JC, Controlling glycation of recombinant antibody in fed-batch cell cultures. *Biotechnol Bioeng*. 2011;108:2600–2610,
5. Merriam-Webster Online Dictionary, "Definition of UNCERTAINTY." 2016, [Online]. Available at: http://www.merriam-webster.com/dictionary/uncertainty. Accessed 2 Mar 2016.
6. Beck MB. Water quality modeling: a review of the analysis of uncertainty. *Water Resour Res*. 1987;23:1393–1442.
7. Eele AJ, Maciejowski JM. Sequential Monte Carlo Optimisation for Air Traffic Management, *TECHNICAL REPORT: CUED/F-INFENG/TR.693*, 2015, [Online]. Available at: https://arxiv.org/pdf/1506.02869.pdf. Accessed 12 Nov 2016.
8. Ferruzzi G, Cervone G, Delle Monache L, Graditi G, Jacobone F. Optimal bidding in a Day-Ahead energy market for Micro Grid under uncertainty in renewable energy production. *Energy*. 2016;106:194–202.
9. Duong PLT, Ali W, Kwok E, Lee M. Uncertainty quantification and global sensitivity analysis of complex chemical process using a generalized polynomial chaos approach. *Comput Chem Eng*. 2016;90:23–30.
10. Farrance I, Frenkel R. Uncertainty in measurement: a review of Monte Carlo simulation using microsoft excel for the calculation of uncertainties through functional relationships, including uncertainties in empirically derived constants. *Clin Biochem Rev*. 2014;35:37–61.
11. Mandur J, Budman H. Robust optimization of chemical processes using Bayesian description of parametric uncertainty. *J Process Control*. 2014;24:422–430.
12. Streif S, Kim KKK, Rumschinski P, Kishida M, Shen DE, Findeisen R, Braatz RD, Robustness analysis, prediction and estimation for uncertain biochemical networks. *IFAC Proc Vol*. 2013;46:1–20.
13. T. Barth. A brief overview of uncertainty quantification and error estimation in numerical simulation, 2005, [Online]. Available at: http://web.stanford.edu/group/cits/pdf/lectures/barth.pdf. Accessed 9 Nov 2016.
14. Măceş DA, Stadtherr MA. Computing fuzzy trajectories for nonlinear dynamic systems. *Comput Chem Eng*. 2013;52:10–25.
15. Soylu S, Proctor AA, Podhorodeski RP, Bradley C, Buckham BJ. Precise trajectory control for an inspection class ROV. *Ocean Eng*. 2016;111:508–523.
16. Sin G, Gernaey KV, Lantz AE. Good modeling practice for PAT applications: propagation of input uncertainty and sensitivity analysis. *Biotechnol Prog*. 2009;25:1043–1053.
17. Singh R, Gernaey KV, Gani R. Model-based computer-aided framework for design of process monitoring and analysis systems. *Comput Chem Eng*. 2009;33:22–42.
18. Singh R, Gernaey KV, Gani R. ICAS-PAT: a software for design, analysis and validation of PAT systems, *Comput Chem Eng*. 2010;34:1108–1136.
19. Lopez-Montero EB, Wan J, Marjanovic O. Trajectory tracking of batch product quality using intermittent measurements and moving window estimation. *J Process Control*. 2015;25:115–128.
20. Theodoridis S, Pikrakis A, Koutroumbas K, Cavouras D. Chapter 1—Classifiers based on bayes decision theory. In: *Introduction to Pattern Recognition*, Boston: Academic Press; 2010:1–27.
21. J. Kruschke. *Doing Bayesian Data Analysis: A Tutorial Introduction with R*, 1st ed. Academic Press; 2010.
22. Soley-Bori M. Dealing with missing data: key assumptions and methods for applied analysis. *Tech Rep*. 2013.
23. Hodgson BJ. Hybrid modelling of bioprocesses, *Thesis Submitt. Degree Eng. Dr. Biochem. Eng. Univ. Lond.*, 2013.
24. Hastie T, Tibshirani R, Friedman J, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Auflage: 2nd ed. 2009. Corr. 7th printing 2013. New York, NY: Springer; 2011.
25. Gustafsson C, Vallverdú J. The best model of a cat is several cats. *Trend Biotechnol*. März 2016;34:207–213.
26. Bayrak ES, Wang T, Cinar A, Undey C. Computational modeling of fed-batch cell culture bioreactor: hybrid agent-based approach. *IFAC-Pap*. 2015;48:1252–1257.
27. CMC Biotech Working Group, "A-Mab: a Case Study in Bioprocess Development," 2009.
28. M. Liu, Kimura R, Jayapal K, Wu P, Trautwein M, Mayer-Bartschmid A, Jockwer A, Barrett S, Fed-batch cell culture

process development: implementing a novel nutrient additive for a robust, high-titer, scalable process, *BioProcess Int*, 11 Sep 2014. [Online]. Available at: http://www.bioprocessintl.com/upstream-processing/bioreactors/fed-batch-cell-culture-process-development-implementing-novel-nutrient-additive-robust-high-titer-scalable-process/. Accessed 27 Feb 2016.
29. Delouvroy F, Le Reverend G, Fessler B, Mathy G, Harmsen M, Kochanowski N, Malphettes L. Evaluation of the advanced micro-scale bioreactor (ambr^{TM}) as a highthroughput tool for cell culture process development. *BMC Proc*. 2013;7:P73.
30. Lu F, Toh PC, Burnett I, Li F, Hudson T, Amanullah A, Li J. Automated dynamic fed-batch process and media optimization for high productivity cell culture process development. *Biotechnol Bioeng*. 2013;110:191–205.
31. Zhou W, Hu WS. On-line characterization of a hybridoma cell culture process. *Biotechnol Bioeng*. 1994;44:170–177.
32. Eyer K, Oeggerli A, Heinzle E. On-line gas analysis in animal cell cultivation: II. Methods for oxygen uptake rate estimation and its application to controlled feeding of glutamine. *Biotechnol Bioeng*. 1995;45:54–62.
33. Konakovsky V, Clemens C, Müller MM, Bechmann J, Berger M, Schlatter S, Herwig C. Metabolic control in mammalian fed-batch cell cultures for reduced lactic acid accumulation and improved process robustness. *Bioengineering*. 2016;3:5.
34. Cannizzaro C, Gügerli R, Marison I, von Stockar U. On-line biomass monitoring of CHO perfusion culture with scanning dielectric spectroscopy. *Biotechnol Bioeng*. 2003;84:597–610.
35. Opel C, Li FJ, Amanullah A. Quantitative modeling of viable cell density, cell size, intracellular conductivity, and membrane capacitance in batch and fed-batch CHO processes using dielectric spectroscopy. *Biotechnol Prog*. 2010;26:1187–1199.
36. 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. *Biotechnol Prog*. 2014;30:479–487.
37. Ducommun P, Ruffieux PA, Furter MP, Marison I, von Stockar U. A new method for on-line measurement of the volumetric oxygen uptake rate in membrane aerated animal cell cultures. *J Biotechnol*. März 2000;78:139–147.
38. Dietzsch C, Spadiut O, Herwig C. On-line multiple component analysis for efficient quantitative bioprocess development. *J Biotechnol*. 2013;163:362–370.
39. K, Dhara, Stanley J, Ramachandran T, Nair BG, Sathish Babu TG. Pt-CuO nanoparticles decorated reduced graphene oxide for the fabrication of highly sensitive non-enzymatic disposable glucose sensor. *Sens Actuators B Chem*. Mai 2014;195:197–205.
40. Whelan J, Craven S, Glennon B. In situ Raman spectroscopy for simultaneous monitoring of multiple process parameters in mammalian cell culture bioreactors. *Biotechnol Prog*. 2012;28:1355–1362.
41. Berry B, Moretto J, Matthews T, Smelko J, Wiltberger K. Cross-scale predictive modeling of CHO cell culture growth and metabolites using Raman spectroscopy and multivariate analysis. *Biotechnol Prog*. 2015;31:566–577.
42. 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.
43. Capito F, Zimmer A, Skudas R. Mid-infrared spectroscopy-based analysis of mammalian cell culture parameters. *Biotechnol Prog*. März 2015;31:578–584.
44. Hoehse M, Alves-Rausch J, Prediger A, Roch P, Grimm C. Near-infrared spectroscopy in upstream bioprocesses. *Pharm Bioprocess*. 2015;3:153–172.
45. Hakemeyer C, Strauss U, Werz S, Jose GE, Folque F, Menezes JC. At-line NIR spectroscopy as effective PAT monitoring technique in Mab cultivations during process development and manufacturing. *Talanta*. 2012;90:12–21.
46. Sitton G, Srienc F. Mammalian cell culture scale-up and fed-batch control using automated flow cytometry. *J Biotechnol*. 2008;135:174–180.
47. Bluma A, Höpfner T, Lindner P, Rehbock C, Beutel S, Riechers D, Hitzmann B, Scheper T. In-situ imaging sensors for bioprocess monitoring: state of the art. *Anal Bioanal Chem*. 2010;398:2429–2438.
48. Biechele P, Busse C, Solle D, Scheper T, Reardon K. Sensor systems for bioprocess monitoring. *Eng Life Sci*. 2015;15:469–488.
49. Zhao L, Fu HY, Zhou W, Hu WS. Advances in process monitoring tools for cell culture bioprocesses. *Eng Life Sci*. 2015;15:459–468.
50. Mulukutla BC, Gramer M, Hu WS. On metabolic shift to lactate consumption in fed-batch culture of mammalian cells. *Metab Eng*. 2012;14:138–149.
51. Changeux JP. 50 years of allosteric interactions: the twists and turns of the models. *Nat Rev Mol Cell Biol*. 2013;14:819–829.
52. Gould GW, Holman GD. The glucose transporter family: structure, function and tissue-specific expression. *Biochem J*. 1993;295:329–341.
53. Hu W-S, Zhou W, Seth G, Ozturk S, and Zhang C. *Cell Culture Bioprocess Engineering*. Minnesota, Minn: Wei-Shou Hu; 2012.
54. Ozturk S and Hu W-S, *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*, 1st ed. Boca Raton: CRC Press; 2005.
55. Le H, Kabbur S, Pollastrini L, Sun Z, Mills K, Johnson K, Karypis G, Hu WS. Multivariate analysis of cell culture bioprocess data–lactate consumption as process indicator. *J Biotechnol*. 2012;162:210–223.
56. Konakovsky V, Le T, Yongky A, Raju R, Vishvanathan N, Herwig C, Hu WS. Improved metabolic process control by analysis of genetic clone background in mammalian cell culture, *Marshall Plan Foundation*, 2015, [Online]. Available at: http://www.marshallplan.at/images/Konakovsky.pdf. Accessed 27 May 2015.
57. Hu W-S, Kantardjieff A, Mulukutla BC. Cell lines that overexpress lacatate dehydrogenase c, WO2012075124 A2, 2012.
58. Becker TC, Han H-P, Newgard CB, Wilson JE. Methods and compositions for inhibiting hexokinase, WO/1997/026322, 1997.
59. Toussaint C, Henry O, Durocher Y. Metabolic engineering of CHO cells to alter lactate metabolism during fed-batch cultures. *J Biotechnol*. 2016;217:122–131.
60. Drapeau D, Luan Y-T, Stanek TC. Restricted glucose feed for animal cell culture. WO2004104186 A1, 2004.
61. Luan Y-T, Stanek TC, Drapeau D. Controlling lactic acid production in fed-batch cell cultures via variation in glucose concentration; bioreactors and heterologous gene expression, US7429491 B2, 2008.
62. Cruz HJ, Moreira JL, Carrondo MJ. Metabolic shifts by nutrient manipulation in continuous cultures of BHK cells. *Biotechnol Bioeng*. 1999;66:104–113.
63. Konakovsky V, Yagtu AC, Clemens C, Müller MM, Berger M, Schlatter S, Herwig C. Universal capacitance model for real-time biomass in cell culture. *Sensors*. 2015;15:22128–22150.
64. Poddar R, Andrews J, Shukla TP, Sen P. Non-invasive glucose monitoring techniques: a review and current trends. Review Article, arXiv.org, 2008, Available at: https://arxiv.org/abs/0810.5755. Accessed 22 Dec 2016.
65. Bosch ME, Sánchez AJR, Rojas FS, Ojeda CB. Recent development in optical fiber biosensors. *Sensors*. 2007;7:797–859.
66. Waynant RW, Chenault VM. Overview of non-invasive fluid glucose measurement using optical techniques to maintain glucose control in diabetes mellitus, 1998, [Online]. Available at: http://photonicssociety.org/newsletters/apr98/overview.htm. Accessed 13 Mar 2016.
67. Zhu Z, Garcia-Gancedo L, Flewitt AJ, Xie H, Moussy F, Milne WI. A critical review of glucose biosensors based on carbon nanomaterials: carbon nanotubes and graphene. *Sensors*. 2012;12:5996–6022.
68. Herwig C, Assessment of data quality and Know-Why for a scalable QbD Approach, 2013, [Online]. Available at: https://lifescience.roche.com/wcsstore/CBCatalogAssetStore/Articles/Biolprod-DUS_ Herwig_20130924.pdf. Accessed 31 Jul 2016.
69. Zagari F, Jordan M, Stettler M, Broly H, Wurm FM. Lactate metabolism shift in CHO cell culture: the role of mitochondrial oxidative activity. *New Biotechnol*. 2013;30:238–245.
70. Camila CA, Wilkens A. Comparative metabolic analysis of lactate for CHO cells in glucose and galactose. *PLoS One* 2011;16, no. 4:714–724.

71. Osman JJ, Birch J, Varley J. The response of GS-NS0 myeloma cells to pH shifts and pH perturbations. *Biotechnol Bioeng*. 2001;75:63–73.
72. Aghamohseni H, Ohadi K, Spearman M, Krahn N, Moo-Young M, Scharer JM, Butler M, Budman HM. Effects of nutrient levels and average culture pH on the glycosylation pattern of camelid-humanized monoclonal antibody. *J Biotechnol*. 2014; 186C:98–109.
73. Brunner M, Fricke J, Kroll JP, Herwig C. Investigation of the interactions of critical scale-up parameters (pH, pO2 and pCO2) on CHO batch performance and critical quality attributes. *Bioprocess Biosyst Eng*. 2016; 1–13.
74. Trummer E, Fauland K, Seidinger S, Schriebl K, Lattenmayer C, Kunert R, Vorauer-Uhl K, Weik R, Borth N, Katinger H, Müller D. Process parameter shifting: part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors. *Biotechnol Bioeng*. 2006;94:1033–1044.
75. Kaufmann H, Mazur X, Fussenegger M, Bailey JE. Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. *Biotechnol Bioeng*. 1999;63:573–582.
76. Zhou W, Rehm J, Europa A, Hu WS. Alteration of mammalian cell metabolism by dynamic nutrient feeding. *Cytotechnology*. 1997;24:99–108.
77. Raghunand N, Dale BE. Alteration of glucose consumption kinetics with progression of baculovirus infection in *Spodoptera frugiperda* cells. *Appl Biochem Biotechnol*. 1999; 80:231–242.
78. Higuera G, Schop D, Janssen F, van Dijkhuizen-Radersma R, van Boxtel T, van Blitterswijk CA. Quantifying in vitro growth and metabolism kinetics of human mesenchymal stem cells using a mathematical model. *Tissue Eng Part A*. 2009;15:2653–2663.
79. Pattappa G, Heywood HK, de Bruijn JD, Lee DA. The metabolism of human mesenchymal stem cells during proliferation and differentiation. *J Cell Physiol*. 2011;226:2562–2570.
80. Reuveny S, Kim YJ, Kemp CW, Shiloach J. Production of recombinant proteins in high-density insect cell cultures. *Biotechnol Bioeng*. 1993;42:235–239.
81. Druhmann D, Reinhard S, Schwarz F, Schaaf C, Greisl K, Nötzel T. Utilizing Roche Cedex Bio analyzer for in process monitoring in biotech production. *BMC Proc*. 2011;5:P106.
82. Klein T, Reichelt WN, Herwig C. Cedex Bio HT Analyzer as a tool for high-throughput analysis for fast physiological characterization of microbial bioprocesses. Application Note, 2014, Available at: http://custombiotech.roche.com/content/dam/internet/dia/custombiotech/custombiotech_com/en_GB/pdf/Custom Biotech_Cedex_Bio-HT_Analyzer_ApplicationNote.pdf. Accessed 22 Dec 2016.
83. Larson TM, Gawlitzek M, Evans H, Albers U, Cacia J. Chemometric evaluation of on-line high-pressure liquid chromatography in mammalian cell cultures: analysis of amino acids and glucose. *Biotechnol Bioeng*. 2002;77:553–563.
84. Kurokawa H, Park YS, Iijima S, Kobayashi T. Growth characteristics in fed-batch culture of hybridoma cells with control of glucose and glutamine concentrations. *Biotechnol Bioeng*. 1994; 44:95–103.
85. Hartlep M, Künnecke W. Online -Controlling of Glucose in Cultivations and Fermentations, 2014, Available at: https://www.trace.de/fileadmin/bilder/pdf/Poster_A0_C2_Control_04_2014.pdf. Accessed 22 Dec 2016.
86. Derfus G, Abramzon E, Tung D, Chang M, Kiss DR, Amanullah A. Cell culture monitoring via an auto-sampler and an integrated multi-functional off-line analyzer. *Biotechnol Prog*. 2010;26:284–292.
87. Gleason B, Dowe N, Klingensmith J, Miller W. Rapid measurement of xylose and glucose for monitoring corn stover fermentation in bioethanol production, 2015, Available at: https://www.ysi.com/File%20Library/Documents/Application%20Notes/YSI_LS_NREL_Poster_Biofuel_37SBFC.pdf. Accessed 22 Dec 2016.
88. Senzime A. "BioSenz™ vs Nova Biomedical 100+ vs Roche Cedex Bio HT. *Senzime AB*, 2016, Available at: http://www.senzime.com/storage/app/media/uploaded-files/AN%20BioSenz%20Alexion.pdf. Accessed 22 Dec 2016..
89. Rhiel M, Ducommun P, Bolzonella I, Marison I, von Stockar U. Real-time in situ monitoring of freely suspended and immobilized cell cultures based on mid-infrared spectroscopic measurements. *Biotechnol Bioeng*. 2002;77:174–185.
90. Moretto J, Smelko JP, Cuellar M, Berry B, Doane A, Ryll T, Wiltberger K. Process Raman spectroscopy for in-line CHO cell culture monitoring. *Am Pharm Rev*. Apr. 2011;14:18–25.
91. Mehdizadeh H, Lauri D, Karry KM, Moshgbar M, Procopio-Melino R, Drapeau D. Generic Raman-based calibration models enabling real-time monitoring of cell culture bioreactors. *Biotechnol Prog*. 2015;31:1004–1013.
92. Jeon WY, Choi YB, Kim HH. Disposable non-enzymatic glucose sensors using screen-printed nickel/carbon composites on indium tin oxide electrodes. *Sensors*. 2015;15:31083–31091.
93. So CF, Choi KS, Wong TK, Chung JW. Recent advances in noninvasive glucose monitoring. *Med Devices Auckl NZ*. 2012; 5:45–52.
94. Bauer I, Spichinger S, Spichinger-Keller UE, and John GT. Novel single-use sensors for online measurement of glucose. *BioProcess Int*. 2012. [Online]. Available at: http://www.bioprocessintl.com/upstream-processing/cell-culture-media/novel-single-use-sensors-for-online-measurement-of-glucose-334669/. Accessed 28 Feb 2016.
95. Stefan Spichiger, *In-line Glucose Sensor CITSens Bio for the use in disposable bioreactors*, youtube, 2013. Available at: https://www.youtube.com/watch?v=oeFUNP1mZGs. Accessed 22 Dec 2016.
96. Moser I, Jobst G. Pre-calibrated biosensors for single-use applications. *Chem Ing Tech*. 2013;85:172–178.
97. JangMarkowitz SK, Li H. A new approach to present a non-invasive optical glucose sensor using advanced opto-electronic technology. *ASEE Am Soc Eng Educ*.
98. Geddes CD, Lakowicz JR. Topics in Fluorescence Spectroscopy, Glucose Sensing, Vol 11, Springer US; 2006. Available at: https://www.amazon.com/Glucose-Sensing-Topics-Fluorescence-Spectroscopy/dp/1489993444. Accessed 13 Mar 2016.

Appendix 1: Sensors Capable of Automatic Glucose Measurement for Potential Control Purposes in Fermentation

Technology	Example	Analysis Time	Reference
Metabolic analyzers	CuBiAn XC	Minutes	33,38
	Cedex BioHT	Minutes	81,82
	HPLC	Minutes	83,84
	Trace	Minutes	85
	Nova BioProfile FLEX	Minutes	86
	YSI Biochemistry Analyzer	Minutes	87
	BioSenz	Minutes	88
Spectroscopic probes	MIR	Instant	89
	NIR	Instant	44,90
	RAMAN	Instant	41,42,91
Novel sensors	Non-Enzymatic Sensor	Instant	92
	Non-Enzymatic Sensor	Instant	39,93
	CITSens	Instant	94,95
	Disposable Biosensor	Instant	96
	Optical	Instant	65,97,98

All Online Analyzers require a membrane to withdraw sample liquid, spectroscopic-based inline probes require multivariate data processing and potentially recalibrations, but do not reduce culture volume, while novel sensors may detect glucose levels without either but might not yet be as accurate as established techniques due to interferences with culture fluid.

Appendix 2: Definition of Metabolic State

One metabolic state is defined as the state of balanced growth condition. Hence, the cell is in equilibrium with its environment

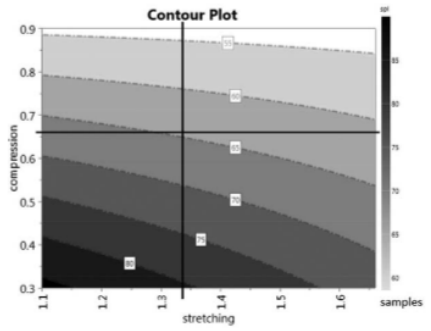


Figure A1. Screening for alternative stretching and compression factors which alter the sampling interval automatically. In all simulations of this contribution, 1.33 for stretching and 0.66 for compression were selected (indicated by a cross in the plot).

Notes: In the author's opinion, a simulation is only credible if all assumptions are named, all equations supplied and the calculation steps described in the necessary detail to reproduce results, so that other researchers may find them useful for their own purposes. Equations were avoided where possible in this document and replaced with figures which illustrate the most important concepts. In addition, this contribution will provide the full algorithm written in Matlab in the Supporting Information files upon publication alongside with the dataset obtained in the simulations.

and exhibits one defined growth stoichiometry. Metabolic states are for example oxidative growth without metabolite over flow, lactate production or lactate uptake. Those states are generally independent from the processing mode, e.g., continuous, fed-batch or batch as not concentrations matter but the physiological specific rates, which are triggered by the media environment. Of course in cell culture the transition from one to another metabolic state is also a function of time, and cannot be set so clearly as in microbial systems.

Appendix 3: Feeding Individual Cells vs. the Whole Population

Our approach targets bioprocess development and manufacturing needs. It assumes homogeneously mixed bioreactor, as it is true for small scale bioreactors, which are used in process development. Inhomogeneities as they may occur in large scale are not scope of this study, those elements should be targeted by scale down models, as we target them currently in our labs. For the assumption of well mixed conditions, we can convert volumetric feeds via VCC into specific rates, such as a specific substrate uptake rate in the unit pg/ch.

Appendix 4: Compression and Stretching Factors

In our proposed control strategy, compression and stretching factors were introduced which influence the automatically selected sampling interval. This indirectly affects the power of the corrective slope k, which slightly adapts the stoichiometric feedrate. We found that a modification of the current sampling interval by 33% (up or down) worked well in all systems in this study, but acknowledge that they might be optimized in future studies with different growth profiles or clones with different metabolic properties. We challenged therefore our choice one more time, using one of the simulated datasets as benchmark. The industrial process with tighter control limits at 1 ± 0.5 g/L GLC was chosen, as it could be potentially more sensitive to changes in the sampling intervals.

The stretching factors screened were 1.1, 1.33, 1.6, while the compression factors were 0.3, 0.66, 0.9. Factors very close to 1.0 are generally not recommended, as these may lead to over- or undersampling in different process phases. All combinations were studied with 5-fold replicates and three center points in a full factorial DoE and can be found in the supplement. A response surface plot using PLS regression on sampling events was created and visualized in Appendix Figure 1. The regression coefficients can be found in the supplement. As expected, the results fall in an area of the variation for our benchmark simulation (61 ± 2) Varying the factors seemed to influence the total sampling number slightly, yet the room for improvement is rather low (spl = 55 vs. 61, an improvement of ca. 10%).

Considering the potential application of this algorithm, where it would have to be generalizable to a score of new processes without such an optimization screen, we believe that even though in this particular screen the compression factor could be increased to reduce sampling events, we would recommend not to change them as they were working very well in all of the different cases studied.

Manuscript received Nov. 20, 2016, and revision received Jan. 11, 2017.

3 Summary & Conclusions

Summary

Online process monitoring and control in our labs was implemented to improve productivity, robustness and predictability at different scales. Predictive models for cell physiology and sensor signals using PAT tools were established, combining our own and industrial data sets at lab (2L) and pilot plant scale (80L). The developed models proved to be scale-independent and transferable to other CHO clones, and processes with limited prior information were demonstrated to be effectively controlled in our labs. Feed control in fed-batch mode was achieved by setting the specific glucose consumption rate within narrow ranges, using an inline capacitance probe for control and a metabolic analyser for process monitoring. Not only could we show that statistical biomass models generated in other processes were transferable, but we also held stable glucose, lactic acid and pH profiles, improving productivity and robustness of the platform process with scale-free parameters. Novel mechanistic models were combined with transcriptomic and fermentation process data inputs *in*-silico to gain novel insights into the root cause of lactic acid overflow metabolism. The results suggest interesting targets for cell line development and imply that timelines may be shortened using novel *in-silico* bioreactor screening runs, which utilize the available genetic information as input parameter. Simpler versions of the mechanistic model allowed predictive run forecast at 2L and 12000L scale to hold glucose levels tightly with uncertainty trajectories, using novel PAT sensor technology. The developed control strategy was developed, tested and validated on both lab and industrial scale datasets, where it was shown in over 500 *in silico* simulations to adequately cope with high uncertainty and error in key process variables, such as biomass, glucose and metabolic state.

Conclusions

All established methodologies were developed to facilitate and improve process transfer and scale-up of an industrial platform CHO cell process. This was realized by implementing PAT-capable control loops in the studies to further develop the platform, and using historical industrial datasets to demonstrate the applicability of the developed methods for large scale cultures. In summary, M^3C using PAT technology was employed throughout this work and, where no measurements were available yet, industrial and process development data was mimicked as input for *in-silico* simulations. Reliable and robust process control could be demonstrated both in theory and in practice, and additional supplementary material, including data, can be found on the publisher's websites. Novel control strategies were established with the intended purpose to overcome challenges as more PAT sensor data will become available in the near future. In conclusion, this thesis aims at a facilitated scale-up of industrial mAb platform processes through advanced process monitoring, modelling and control for the benefit of patients who depend on medicines, which are produced robustly and predictably to the highest quality standards by design, as demanded by regulators from the FDA and EMA.

4 Technical Experience

- Launched Mammalian Cell Culture 2012 at the Research Division Biochemical Engineering, Herwig Group, Vienna University of Technology
- Experience with adherent and suspension cell lines in shake flasks and bioreactors
- Tech transfer and establishment of an industrial mammalian cell culture scale-down model
- Extensive hands-on experience in specifying, operating, monitoring and controlling mammalian cell culture bioprocesses from cryo-vial to bioreactor (Infors Bioreactor)
- Experience with scripting process management software (Lucullus PIMS, Securecell) for automated process control
- Experience with online-capable sensors and equipment for process development (pH, pO2, pCO2, Turbidity, Capacitance, Offgas, Dialysis, Online Analyzer)
- Development of new HPLC methods (biogenic amines) and transfer of known methods (amino acids) in research division
- Prepared SOPs for autoclaving, amino acid measurement, cell measurement, technical operation of an enzymatic analyzer in online mode, seed train, and bioreactor preparation and operation.
- Instructing and training students in autoclaving, safe handling of liquid nitrogen, sterile working in a cell culture laboratory and the art of operating bioreactors
- Programming, data analysis, MVDA and modeling in MATLAB, R, Excel, Spotfire, Modde, Simca and related software
- Upstream mammalian cell culture process monitoring and control using soft sensors and PAT technology

5 Abbreviations

μmax	Maximum growth rate
μ	Growth rate
a_(1…n)	Parameters in a multivariate model
BI	Böhringer Ingelheim
c_(1…n)	Coefficients in a multivariate model
CHO	Chinese Hamster Ovary Cells
DoE	Design of Experiment
EMA	European Medicines Agency
ESBES	European Society of Biochemical Engineering Science
FDA	Food and Drugs Administration
GLC	Glucose
GLN	Glutamine
GLU	Glutamate
HK	Hexokinase
IgG	Immunoglobulin G, mostly referred to as product
K_s	Substrate affinity coefficient
LAC	Lactic acid
LDH	Lactate dehydrogenase
M^3C	Measurement, Modelling, Monitoring and Control
MVDA	Multivariate Data Analysis
NAD/NADH	Nicotinamide adenine dinucleotide
NH_4	Ammonia
PAT	Process Analytical Technology
PFK	Phosphofructokinase
PFKFB	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase
PIMS	Process Information Management System
PK	Pyruvate kinase
QbD	Quality by Design
S	Substrate, here mostly GLC
SOP	Standard Operating Procedure
VK	Viktor Konakovsky
VUT	Vienna University of Technology
y	Predicted variable in a multivariate model

6 Scientific contributions

	Year	Title	Ref
Poster contributions			
	2014	*Real-time process control with PAT tools*	[27]
	2015	*Online monitoring, modeling and process control in mammalian cell culture*	[47]
Oral presentations			
	2014	*Lactic acid consumption without collinear effects*	[28]
	2015	*Bridging transcriptomics, modeling and process control in a systems approach for mammalian cell culture*	[48]
	2015	*BioPro World Talent Campus case study - Full scale fermentation heterogeneity*	[46]
	2016	*Metabolic Control Of Mammalian Fed-Batch Processes With Transferable Biomass Models*	[49]
	2017	*Universal Capacitance Model for Real-Time Biomass in Cell Culture*	[50]
Publications			
	2015	*Improved metabolic process control by analysis of genetic clone background in mammalian cell culture*	[12]
	2015	*Universal Capacitance Model for Real-Time Biomass in Cell Culture*	[10]
	2016	*Metabolic Control in Mammalian Fed-Batch Cell Cultures for Reduced Lactic Acid Accumulation and Improved Process Robustness*	[11]
	2017	*A robust feeding strategy to maintain set-point glucose in mammalian fed-batch cultures when input parameters have a large error*	[13]

7 Tables

8 Figures

9 References

[1] M. Aehle, A. Kuprijanov, S. Schaepe, R. Simutis, and A. Lübbert, “Increasing batch-to-batch reproducibility of CHO cultures by robust open-loop control,” *Cytotechnology*, vol. 63, no. 1, pp. 41–47, Jan. 2011.

[2] M. Gagnon, G. Hiller, Y.-T. Luan, A. Kittredge, J. DeFelice, and D. Drapeau, “High-End pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO Fed-batch cultures,” *Biotechnol. Bioeng.*, vol. 108, no. 6, pp. 1328–1337, 2011.

[3] R. Luttmann *et al.*, “Soft sensors in bioprocessing: a status report and recommendations,” *Biotechnol. J.*, vol. 7, no. 8, pp. 1040–1048, Aug. 2012.

[4] D. Paulsson, R. Gustavsson, and C.-F. Mandenius, “A Soft Sensor for Bioprocess Control Based on Sequential Filtering of Metabolic Heat Signals,” *Sensors*, vol. 14, no. 10, pp. 17864–17882, Sep. 2014.

[5] P. Biechele, C. Busse, D. Solle, T. Scheper, and K. Reardon, “Sensor systems for bioprocess monitoring,” *Eng. Life Sci.*, vol. 15, no. 5, pp. 469–488, Jul. 2015.

[6] C.-F. Mandenius and N. J. Titchener-Hooker, Eds., *Measurement, Monitoring, Modelling and Control of Bioprocesses*, 2013 edition. Springer, 2015.

[7] K. V. Gernaey *et al.*, “Monitoring and control of microbioreactors: An expert opinion on development needs,” *Biotechnol. J.*, vol. 7, no. 10, pp. 1308–1314, Oct. 2012.

[8] J. Glassey *et al.*, “Process analytical technology (PAT) for biopharmaceuticals,” *Biotechnol. J.*, vol. 6, no. 4, pp. 369–377, 2011.

[9] G. Striedner and K. Bayer, “Considerations in upstream bioprocess monitoring and statistical data analysis in the context of process analytical technology and quality by design,” *Pharm Bioprocess*, vol. 1, no. 2, pp. 159–166, Jun. 2013.

[10] V. Konakovsky *et al.*, “Universal Capacitance Model for Real-Time Biomass in Cell Culture,” *Sensors*, vol. 15, no. 9, pp. 22128–22150, Sep. 2015.

[11] V. Konakovsky *et al.*, “Metabolic Control in Mammalian Fed-Batch Cell Cultures for Reduced Lactic Acid Accumulation and Improved Process Robustness,” *Bioengineering*, vol. 3, no. 1, p. 5, Jan. 2016.

[12] V. Konakovsky *et al.*, “Improved metabolic process control by analysis of genetic clone background in mammalian cell culture,” *Marshall Plan Foundation*. [Online]. Available: http://www.marshallplan.at/images/Konakovsky.pdf. [Accessed: 27-May-2015].

[13] V. Konakovsky, C. Clemens, M. M. Müller, J. Bechmann, and C. Herwig, “A robust feeding strategy to maintain set-point glucose in mammalian fed-batch cultures when input parameters have a large error,” *Biotechnol. Prog.*, p. n/a-n/a, Jan. 2017.

[14] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[15] Thomas Natschläger and Birgit Zauner, “Fused stage-wise lasso – a waveband selection algorithm for spectroscopy - SCCH - DE.” [Online]. Available: http://www.scch.at/de/publikationen/publication_id/802. [Accessed: 19-Aug-2014].

[16] V. Nandita, L. Huong, T. Le, and H. Wei-Shou, “Advancing biopharmaceutical process science through transcriptome analysis,” *Curr. Opin. Biotechnol.*, vol. 30, pp. 113–119, 2014.

[17] FDA, “Guidance for Industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance.” [Online]. Available: http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf. [Accessed: 02-Nov-2016].

[18] EMA, “ICHQ8(R2).” Sep-2015.

[19] ICH, “Quality Guidelines : ICH.” [Online]. Available: http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html. [Accessed: 14-Mar-2017].

[20] A. S. Rathore and H. Winkle, “Quality by design for biopharmaceuticals,” *Nat. Biotechnol.*, vol. 27, no. 1, pp. 26–34, Jan. 2009.

[21] MATLAB, Inc., “Robust local regression using weighted linear least squares in matlab (RLOWESS), http://de.mathworks.com/help/curvefit/smoothing-data.html#bq_6ys3-3.” .

[22] MATLAB, Inc., “Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) - MATLAB pchip - MathWorks Deutschland.” [Online]. Available: https://de.mathworks.com/help/matlab/ref/pchip.html. [Accessed: 12-Mar-2017].

[23] MATLAB, Inc., “Savitzky-Golay filtering - MATLAB sgolayfilt - MathWorks Deutschland.” [Online]. Available: https://de.mathworks.com/help/signal/ref/sgolayfilt.html. [Accessed: 12-Mar-2017].

[24] H. Lohninger, “H. Lohninger, (2000) Datalab 3.5, A programme for statistical analysis, http://datalab.epina.at/.” .

[25] J. Monod, J. Wyman, and J.-P. Changeux, “On the nature of allosteric transitions: A plausible model,” *J. Mol. Biol.*, vol. 12, no. 1, pp. 88–118, Mai 1965.

[26] J.-P. Changeux, “50 years of allosteric interactions: the twists and turns of the models,” *Nat. Rev. Mol. Cell Biol.*, vol. 14, no. 12, pp. 819–829, Dec. 2013.

[27] V. Konakovsky, M. M. Müller, M. Berger, S. Schlatter, and C. Herwig, “Real-time process control with PAT-Tools.” PAT & QBD Forum Göttingen, Feb-2014.

[28] V. Konakovsky, M. M. Müller, M. Berger, S. Schlatter, and C. Herwig, “Lactic acid consumption without collinear effects.” 10th European Symposium on Biochemical Engineering Sciences and 6th International Forum on Industrial Bioprocesses in collaboration with ACS, Sep-2014.

[29] G. L’Allemain, S. Paris, and J. Pouysségur, “Growth factor action and intracellular pH regulation in fibroblasts. Evidence for a major role of the Na+/H+ antiport.,” *J. Biol. Chem.*, vol. 259, no. 9, pp. 5809–5815, May 1984.

[30] M. Brunner *et al.*, “The impact of pH inhomogeneities on CHO cell physiology and fed-batch process performance – two-compartment scale-down modelling and intracellular pH excursion,” *Biotechnol. J.*, Feb. 2017.

[31] M. Magnani, V. Stocchi, G. Serafini, and M. Bossù, “Effects of buffers and pH on rabbit red blood cell hexokinase,” *Ital. J. Biochem.*, vol. 32, no. 1, pp. 28–35, Feb. 1983.

[32] P. J. Mulquiney and P. W. Kuchel, “Model of the pH-dependence of the concentrations of complexes involving metabolites, haemoglobin and magnesium ions in the human erythrocyte,” *Eur. J. Biochem. FEBS*, vol. 245, no. 1, pp. 71–83, Apr. 1997.

[33] B. C. Mulukutla, S. Khan, A. Lange, and W.-S. Hu, “Glucose metabolism in mammalian cell culture: new insights for tweaking vintage pathways,” *Trends Biotechnol.*, vol. 28, no. 9, pp. 476–484, Sep. 2010.

[34] B. C. Mulukutla, M. Gramer, and W.-S. Hu, “On metabolic shift to lactate consumption in fed-batch culture of mammalian cells,” *Metab. Eng.*, vol. 14, no. 2, pp. 138–149, Mar. 2012.

[35] B. C. Mulukutla, A. Yongky, P. Daoutidis, and W.-S. Hu, “Bistability in Glycolysis Pathway as a Physiological Switch in Energy Metabolism,” *PLoS ONE*, vol. 9, no. 6, p. e98756, Jun. 2014.

[36] A. Yongky, “Analysis of central metabolic pathways in cultured mammalian cells,” Oct. 2014.

[37] A. Yongky, J. Lee, T. Le, B. C. Mulukutla, P. Daoutidis, and W.-S. Hu, “Mechanism for multiplicity of steady states with distinct cell concentration in continuous culture of mammalian cells,” *Biotechnol. Bioeng.*, p. n/a-n/a, Feb. 2015.

[38] T. C. Becker, H.-P. Han, C. B. Newgard, and John E Wilson, “METHODS AND COMPOSITIONS FOR INHIBITING HEXOKINASE,” WO/1997/026322, 24-Jul-1997.

[39] B. Szperalski, C. Jung, Z. Shao, A. Kantardjieff, and W.-S. Hu, “LDH-C can be differentially expressed during fermentation of CHO cells,” *BMC Proc.*, vol. 5, no. Suppl 8, p. P107, Nov. 2011.

[40] W.-S. Hu, A. Kantardjieff, and B. C. Mulukutla, “Cell lines that overexpress lacatate dehydrogenase c,” WO2012075124 A2, 07-Jun-2012.

[41] H. Le *et al.*, “Multivariate analysis of cell culture bioprocess data--lactate consumption as process indicator,” *J. Biotechnol.*, vol. 162, no. 2–3, pp. 210–223, Dec. 2012.

[42] W. Zhou, J. Rehm, A. Europa, and W.-S. Hu, “Alteration of mammalian cell metabolism by dynamic nutrient feeding,” *Cytotechnology*, vol. 24, no. 2, pp. 99–108, Jul. 1997.

[43] S. S. Ozturk and B. O. Palsson, “Growth, metabolic, and antibody production kinetics of hybridoma cell culture: 2. Effects of serum concentration, dissolved oxygen concentration, and medium pH in a batch reactor,” *Biotechnol. Prog.*, vol. 7, no. 6, pp. 481–494, Dec. 1991.

[44] S. S. Ozturk, J. C. Thrift, J. D. Blackie, and D. Naveh, “Real-time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor,” *Biotechnol. Bioeng.*, vol. 53, no. 4, pp. 372–378, 1997.

[45] S. Ozturk and W.-S. Hu, *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*, 1 edition. Boca Raton: CRC Press, 2005.

[46] A. Bockisch, A. O. Chatzivasileiou, S. Kang, V. Konakovsky, and P. Skou, “BioPro WTC case study - Full scale fermentation heterogeneity.” BioPro World Talent Campus, Aug-2015.

[47] V. Konakovsky and C. Herwig, “Online monitoring, modeling and process control in mammalian cell culture.” BIOPRO World Talent Campus, Sorø, Aug-2015.

[48] V. Konakovsky *et al.*, “Bridging transcriptomics, modeling and process control in a systems approach for mammalian cell culture.” 11. Minisymposium Verfahrenstechnik, Wien; ISBN: 978-3-200-04069-4; 142 - 147., Apr-2015.

[49] V. Konakovsky, C. Clemens, M. M. Müller, J. Bechmann, M. Berger, and C. Herwig, “Metabolic Control Of Mammalian Fed-Batch Processes With Transferable Biomass Models.” European Symposium on Biochemical Engineering (ESBES), Dublin, Sep-2016.

[50] V. Konakovsky, C. Clemens, M. M. Müller, J. Bechmann, M. Berger, and C. Herwig, “Universal Capacitance Model for Real-Time Biomass in Cell Culture.” BIT’s 5th Annual Congress of AnalytiX, Fukuoka, Mar-2017.