Structured Kinetic Modeling for Rational Scale-down and Design Optimization of Industrial Fermentations

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Monday 5 June 2023 at 15:00 hours

By

Wenjun TANG

Master of Engineering in Biochemical Engineering,
East China University of Science and Technology, China
Born in Shanghai, China

This dissertation has been approved by the
Promoter: Prof.dr.ir. H.J. Noorman
Copromoter: Dr. W.M. van Gulik

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof.dr.ir. H.J. Noorman	Delft University of Technology, promoter
Dr. W.M. van Gulik	Delft University of Technology, copromoter

Independent members:

Prof.dr.-Ing. R. Takors	University of Stuttgart
Prof.dr. S.A. Wahl	University of Erlangen–Nuremberg
Dr. W. van Winden	DSM-Firmenich
Prof.dr.ir. T. Kiss	Delft University of Technology
Prof.dr. B. Teusink	Free University Amsterdam

Reserve member:

Prof.dr. M. Ottens	Delft University of Technology

The research presented in this thesis was performed at the Bioprocess Engineering section, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, The Netherlands. The research is partially supported by a multi-party research project, (NWO-MoST Joint program, 2013DFG32630) and partially supported by Royal DSM.

This thesis is dedicated to my beloved mother,

Who, unfortunately, can't witness this advanture.

谨以此书纪念我敬爱的母亲

很遗憾，她无法亲自见证这一时刻

Summary

Bioprocesses use the power and versatility of nature via microorganisms to make bio-products from renewable feedstocks. New opportunities lie in front of us, as nowadays microorganisms can very well be engineered to become efficient cell factories. However, the industrialization of those bioprocesses is often hindered by the so called 'scale-up effect', which is apparent as a loss of product yield and/or productivity after transfer to industrial-scale bioreactors (Wang et al., 2015, 2020). Insufficient heat and mass transfer, together with limited mixing, are normally causes of all kinds of gradients in those large bioreactors, which then result in performance loss of the producing strain compared to what was realized in the homogeneous environment of lab-scale reactors. As the first step of solving a problem is to understand it, the first step towards understanding the scale-up effect is to quantitatively reproduce the environmental/physiological heterogeneity of the large-scale bioreactor and study the effect on the performance of the producing strain.

In this thesis, we target a robust and model-based simulation platform to obtain quantitative information on the heterogeneous conditions of the large-scale bioreactor, based on profound biological and physical understanding. Modelling, especially the integration of metabolic kinetic models into a computational fluid dynamics (CFD) framework, is one of the ways to ensure a predictive and lower risk technology transfer, by quantitatively addressing detailed scale-up effects. However, the design of a representative kinetic model for the target microorganism remains the Achilles heel of the concept because a compromise between on one hand sufficient model complexity to ensure proper performance under highly dynamic conditions and on the other hand a manageable computation effort, is required. We used two industrially relevant model strains, *Penicillium chrysogenum* and *Saccharomyces cerevisiae* as examples, to demonstrate the feasibility and power of the integration of metabolic kinetic models and CFD.

In **Chapter 2**, we used *P. chrysogenum* as a model system and developed a metabolically structured kinetic model for growth and penicillin production. By lumping the most important intracellular metabolites in 5 pools and incorporating 4 intracellular enzyme pools, linked by 10 reactions, we succeeded in maintaining the model structure relatively simple, while providing informative insight into the state of the organism. The performance of this 9-pool model was validated with a glucose feast–famine cycle experiment at a time scale of minutes. Comparison of this model and a reported black box

model for this strain showed the necessity of employing a structured model under feast–famine conditions. This model is highlighted by carrying a “cellular memory” via intracellular metabolic pools and being an ideal starting point for integration with a computational fluid dynamics framework.

Continued in **Chapter 3**, the 9-pool model was coupled with a Euler-Lagrange CFD simulation of a large 54 m^3 production bioreactor. The integration provided rich spatial information regarding the performance of the cells, including gradient information of glucose uptake rates and specific penicillin productivity. Furthermore, the couped model successfully reproduced an overall drop in penicillin productivity of 18-50% in the large-scale reactor, depending on the model setup. Based on the analysis of the simulation results we proposed and validated a scale-down design representing the conditions the cells experience in the large tank. By identifying glucose concentration gradients as the main cause of the productivity loss, we proposed to change the position of the feeding port of the large reactor to mitigate the mixing challenge. This resulted in a predicted 50% reduction of the penicillin yield loss. Finally, the coupled 9-pool-CFD model was applied to simulate long-term fed-batch cultivations in the large bioreactor, which showed a good agreement with the available experimental data.

As a summary, from **Chapter 2** and **Chapter 3**, we demonstrated a complete, model-assisted development and optimization approach for industrial fermentations, which starts from metabolic kinetic model development, via coupling with Euler-Lagrange-based CFD simulations of an industrial scale bioreactor and developing rational scale-down systems, finally providing targets for improvement of the industrial-scale bioreactor.

In **Chapter 4**, we investigated another industry-favored model strain, the yeast *Saccharomyces cerevisiae*. Before developing a full-scale yeast model, we first separately studied how the kinetics of storage carbohydrate accumulation and hydrolysis responded to short and long-term dynamic conditions. This was done by combining a kinetic model for glycogen and trehalose storage with a classical black-box model for cell growth, respiration, and maintenance. We clearly demonstrated that for a correct modeling of the formation and release of storage carbohydrate compounds, both gene regulation at enzyme expression level and metabolite regulation are needed to capture the dynamic responses across different time scales. This model also proved to be stable and maintained its dynamic features when integrated with cell lifelines from a 22 m^3 pilot tank. We foresee that the properties of the gene regulation model make it suitable

for integration in a full-scale Euler-Lagrange CFD-CRD simulation or, alternatively, compartment-based large scale bioreactor simulation.

It is, however, clearly not sufficient to describe only the storage response of *S. cerevisiae*. In **Chapter 5**, using a similar concept as for the 9-pool *P. chrysogenum* model, a 7-pool yeast model was developed, focused on reproducing the Crabtree and Pasteur effects which are characteristic for *S. cerevisiae*. The Crabtree effect is the phenomenon whereby the yeast produces ethanol under aerobic conditions at high extracellular glucose concentrations. Also, the effects of the dissolved oxygen concentration, which directly acts at the intracellular redox level, were included in this model: this is the basis of the Pasteur effect, which is the occurrence of ethanol production under oxygen-limited conditions. This model was validated using published data from multiple glucose-limited steady state chemostat cultures, single glucose pulse experiments and periodic glucose oscillations and further tested against highly dynamic lifelines (one-way coupling) obtained from the CFD simulation of the 22 m^3 pilot scale tank. The model showed a stable performance throughout the highly dynamic lifelines. This gave us confidence that this model is suited to be integrated into a complete two-way coupled CFD framework. Nevertheless, this yeast model appeared to have certain shortcomings and could be further improved. With the aim to improve this yeast model, we redesigned the kinetics and present an updated model structure in **Chapter 6**, by taking the lessons learned from the 7-pool model into account, thereby carefully examining the functionality of each part of the metabolic network. The improvements proposed are: 1) insertion of two different glucose uptake mechanisms with gene-regulated enzyme capacities; 2) glycerol and ethanol re-consumption based on environmental threshold values for glucose and oxygen; 3) a lumped gluconeogenesis pathway to enable growth on ethanol, acetate, and glycerol, in the absence of glucose; 4) integration of a kinetic module for storage and re-consumption of trehalose and glycogen. We believe that this extended model structure is still compact enough to allow proper parametrization and CFD simulation and has the capability in predicting yeast performance under various conditions across a wide range of growth rates and external triggers at timescales of seconds to hours.

In **Chapter 6**, we further present an outlook on the future of bioprocess development taking advantage of the possibilities offered by digitalization. The demonstrated application of CFD-CRD, will be the pioneer of the realization and implementation of the so-called digital twin (DT) concept in the future. The DT

is defined as a virtual representation that serves as the real-time digital counterpart of a physical object or process. The challenge of applying the DT concept to a bioprocess is the definition of such a virtual representation. Life is complex and mankind is far from understanding the mystery of life, even for tiny microorganisms like yeast. Therefore, different from the digital twin of a man-made object, e.g. power generator or an engine, the digital twin model of the microorganism, or the bioprocess that involves the microorganism, has to come with a highly simplified and lumped structure. Here the difficulties are how such a simplified model (compared to reality) can be the representation of the process. Therefore, the mindset of "begin with the end in mind", and addressing the modeling objective first, is crucial for keeping the model development path on the right track and preventing the addition of 'unnecessary' complexities to the model. In this sense, questioning the model with respect to its "fitness for purpose" throughout the development phase is required to ensure a proper complexity and fitness to the application target, with still an affordable computation effort.

This study provides several examples that not only show the power of using simplified metabolic kinetic models coupled with a CFD framework in bioprocess application but also present guidelines in how to construct a lumped kinetic model that is fit for its purpose. This part of work is the foundation and one of the cornerstones of realizing the DT concept in the bioprocess industry, enabling a promising future of smart biomanufacturing.

Samenvatting

Bioprocessen maken gebruik van de veelzijdigheid van de natuur om producten te maken uit hernieuwbare grondstoffen. De hedendaagse ontwikkelingen betreffende het modificeren van micro-organismen tot efficiënte 'cellulaire fabrieken' bieden vele nieuwe kansen voor de bioprocestechnologie. De ontwikkeling van industriële bioprocessen wordt echter vaak bemoeilijkt door opschalingseffecten, in de vorm van een minder efficiënte conversie van grondstof naar product, of een lagere productiesnelheid (Wang et al., 2015, 2020). Deze opschalingseffecten komen vaak voort uit de aanwezigheid van gradiënten in procescondities in industriële bioreactoren, met als onderliggende oorzaak limitaties in warmte- en massaoverdracht en menging op grote schaal. De gradiënten worden door micro-organismen ervaren als tijdsvariabele procescondities, die afwijken ten opzichte van de homogene condities waarin een micro-organisme op labschaal is getest. Om de impact van opschalingseffecten te begrijpen is het essentieel om de invloed van heterogene condities in industriële bioreactoren op het metabolisme van micro-organismen te begrijpen.

De doelstelling van dit proefschrift is het ontwikkelen van een robuust, model gedreven simulatieplatform waarmee kwantitatieve informatie kan worden verkregen betreffende heterogene condities in een industriële bioreactor, gebaseerd op onderliggende biologische en fysische kennis. Het toepassen van een dergelijk simulatieplatform bij industriële procesontwikkeling kan tot een betrouwbaarder opschalingsproces leiden. Door gebruikt te maken van de combinatie van een biokinetisch model voor het micro-organisme, en Computational Fluid Dynamics (CFD) voor transportverschijnselen in de reactor, kan kwantitatief inzicht worden verschaft in opschalingseffecten. Het meest complexe aspect van een dergelijk model is het opstellen van een representatief biokinetisch model voor het micro-organisme: voor accurate voorspellingen van de celdynamiek in variërende omgevingen is een bepaalde modelcomplexiteit noodzakelijk, terwijl efficiënte integratie met CFD om een simpel model vraagt: een compromis tussen deze aspecten is noodzakelijk. In dit werk wordt voor twee micro-organismen een biokinetisch model opgesteld: *Penicillium chrysogenum* en *Saccharomyces cerevisiae*. Deze organismen dienen als voorbeeld om de haalbaarheid en potentie van de integratie van biokinetische modellen in CFD-simulaties aan te tonen.

In **hoofdstuk 2** wordt *P. chrysogenum* als modelorganisme toegepast. Voor dit organisme is een gestructureerd biokinetisch model opgesteld waarmee groei

en penicilline-productie worden beschreven. Het samenvoegen van de relevante intracellulaire componenten in 5 metaboliet- en 4 enzympools, gekoppeld door 10 reacties, zorgt voor een relatief simpele modelstructuur wat integratie met CFD-modellen mogelijk maakt. De kwantitatieve prestaties van dit 9-pool model zijn getoetst tegen experimentele 'feast-famine' data, waarin de glucoseconcentratie op een tijdsschaal van minuten werd gevarieerd. De 9 intracellulaire pools dienen praktisch gezien als een 'cellulair geheugen', waarmee de aanpassing van de cel aan de omgevingscondities wordt gemodelleerd. Door de resultaten van het 9-pool model te vergelijken met een ongestructureerd model wordt de noodzaak van het bijhouden van dit geheugen aangetoond.

Het 9-pool model wordt in **hoofdstuk 3** gekoppeld aan een Euler-Lagrange CFD-simulatie van een industriële productiereactor (54 m^3). Deze koppeling geeft diepgaand inzicht in de interactie tussen micro-organismen en een ruimtelijk heterogene productieomgeving. Het gekoppelde model voorspelt een afname in de specifieke penicilline-productiecapaciteit van 18-50%, ten gevolge van de heterogene verdeling van glucose in het reactorvat. Op basis van de simulatiedata is een ontwerp voor een neerschaalreactor opgezet en gevalideerd, die de op industriële schaal door micro-organismen ervaren condities op labschaal nabootst. Verder wordt het model gebruikt om de invloed van de voedingslocatie op de glucoseverdeling penicillineproductie te bepalen: een gunstiger locatie leidt tot 50% minder verlies in de penicillineopbrengst. Tot slot is het model gebruikt voor de analyse van een fed-batch productieproces, met goede overeenkomst tussen simulatie en gemeten productiedata. Samenvattend wordt in **hoofdstuk 2 en 3** een compleet, model-gedreven ontwikkelings- en optimalisatieproces voor bioprocessen gepresenteerd, middels koppeling van een biokinetisch model en CFD-model, waarmee opschaaleffecten in bioreactoren inzichtelijk worden gemaakt.

In **hoofdstuk 4** wordt de dynamische respons van een ander veelgebruikt micro-organisme, *S. cerevisiae,* onderzocht. Alvorens een volwaardig biokinetisch model op te stellen, wordt eerst de dynamiek van de opslagmetabolieten onder fluctuerende extracellulaire condities bestudeerd, door een kinetisch model voor opname- en afbraak van trehalose en glycogeen te combineren met een ongestructureerd model voor celgroei, ademhaling en onderhoud. Om de dynamiek op alle relevante tijdsschalen te vangen is het van belang om zowel genetische regulatie als metabole regulatie in het model te vangen. Hoewel het resulterende biokinetisch model nog niet direct in CFD-simulaties is

geïntegreerd, is wel getoetst hoe het model reageert op variabele glucoseconcentraties-tijdsseries ('lifelines'); aangezien het model een stabiele, realistische dynamische respons vertoonde, voorzien we geen problemen met integratie in CFD-simulaties (of compartimentenmodellen) van bioreactoren.

Voor een volwaardig model van *S. cerevisiae* is meer nodig dan een dynamisch model voor opslagmetabolieten; in **hoofdstuk 5** wordt op basis van de methodologie uit hoofdstuk 2 een 7-pool model voor *S. cerevisiae* ontwikkeld, met nadruk op de beschrijving van het Crabtree- en het Pasteureffect, karakteristiek voor dit organisme. Het Crabtree-effect betreft de productie van ethanol onder aerobe condities met een hoge glucoseconcentratie, het Pasteur-effect de productie van ethanol onder zuurstof-gelimiteerde condities. Om beide effecten mee te kunnen nemen, houdt het model rekening met variaties in zowel de extracellulaire glucose- als zuurstofconcentratie. Het model is gevalideerd met verschillende gepubliceerde datasets, die zowel constante als gepulseerde en oscillerende condities omvatten. Daarnaast is het model toegepast op 'lifelines' uit de CFD-simulatie van een 22m^3 pilot-schaal fermentatieproces. Ondanks de sterke dynamiek in deze CFD-data gaf het model een stabiele respons, wat de weg vrijmaakt voor toekomstige integratie in een CFD-simulatie. Het ontwikkelde 7-pool model bevatte nog enige tekortkomingen; in **hoofdstuk 6** wordt de doorontwikkeling van het 7-pool model beschreven, waarbij de functionaliteit van de verschillende onderdelen van het metabole netwerk wordt onderzocht. Dit heeft tot de volgende modificaties geleid: 1) implementatie van twee verschillende glucose opnamemechanismen met gereguleerde enzymcapaciteiten, 2) heropname van glycerol en ethanol op basis van de extracellulaire glucose- en zuurstofconcentratie, 3) toevoegen van gluconeogenese om groei in afwezigheid van glucose te modelleren, 4) integratie van een module voor opslag en consumptie van trehalose en glycogeen. Hoewel de modelcomplexiteit hiermee toeneemt, is de verwachting dat het model nog steeds voldoende compact is voor integratie met CFD-simulaties, en dat het model in staat is de biokinetische respons van *S. cerevisiae* op verschillende tijdsschalen en onder variërende extracellulaire condities te voorspellen.

Hoofstuk 6 bevat verder een visie betreffende de toekomst van bioprocesontwikkeling in een wereld van toenemende digitalisering. Het gebruik van CFD-simulaties met geïntegreerde biokinetiek is een eerste stap naar het opzetten en implementeren van een zogenoemde 'digital twin' van een bioreactor – een virtuele reactor die simultaan aan de werkelijke

procesvoering de staat van het proces voorspelt. Digital twins worden reeds toegepast voor verscheidene apparaten, zoals motoren en stroomgeneratoren. Voor deze processen is typisch een goede fysische modelbeschrijving beschikbaar. Bioprocessen zijn echter bijzonder complex, en onze mechanistische kennis van levensprocessen is zelfs voor 'simpele' micro-organismen zeer beperkt. Dit zorgt voor unieke uitdagingen bij het opstellen van digital twins voor bioprocessen, waarbij een sterk versimpelde modelopzet nodig is om tot een effectieve procesbeschrijving te komen, die desondanks het procesverloop onder relevante condities voldoende accuraat moet beschrijven. Hiervoor wordt de filosofie "begin bij het eindpunt" toegepast; beginnen met heldere modeldoelstellingen is essentieel om het model in de juiste richting te ontwikkelen, en om toevoeging van onnodige complexiteiten te voorkomen waardoor de benodigde rekenkracht acceptabel blijft.

Hiermee beschrijft dit proefschrift de mogelijkheden van versimpelde gestructureerde biokinetische modellen geïntegreerd binnen CFD-simulaties, maar worden ook richtlijnen gegeven over hoe dergelijke modellen met een specifieke doelstelling in gedachten kunnen worden ontwikkeld. Hiermee is dit werk een fundamenteel onderdeel voor het realiseren van digital twins in de bioprocesindustrie, en een stap richting in de richting van de veelbelovende toekomst van 'smart biomanufacturing'.

Nomenclature

PGI	phosphoglucoisomerase
PFK	phosphofructokinase
ALDO	fructose-bisphosphate aldolase
TPI	triosephosphate isomerase
GAPDH	glyceraldehyde phosphate dehydrogenase
PGK	phosphoglycerate kinase
PGM	phosphoglycerate mutase
ENO	enolase
PK	pyruvate kinase
PDH	pyruvate dehydrogenase
CS	citrate synthase
ICDH	isocitrate dehydrogenase
α-KGDH	alpha-ketoglutarate dehydrogenase
PDC	pyruvate decarboxylase
HXT	hexose transporters
G6P	glucose-6-phosphate
F6P	fructose-6-phosphate
FBP	fructose-1,6-bisphosphate
Pyr	pyruvate
ATP	adenosine triphosphate
ADP	adenosine diphosphate
AMP	adenosine monophosphate
NADH	nicotinamide adenine dinucleotide (reduced)
NAD^+	nicotinamide adenine dinucleotide (oxidized)
ACA	acetaldehyde
ADH	alcohol dehydrogenase
TCA	tricarboxylic acid cycle
ETC	electron transfer chain
NGAM	non-growth associated maintenance
GAM	growth associated maintenance
AA	amino acids pool
ATP	adenosine triphosphate
PAA	phenylacetic acid
PP	pentose phosphate
TCA	citric acid cycle

Contents

Summary I
Samenvatting V
Nomenclature IX
Chapter 1 1
Chapter 2 27
Chapter 3 53
Chapter 4 83
Chapter 5 121
Chapter 6 157
Publications 181
Acknowledgement 183
Curriculum Vitae 187

Chapter 1

Introduction

Adapted from

Wang, G., **Tang, W.**, Xia, J., Chu, J., Noorman, H., & Gulik, W. M. (2015). Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. Engineering in Life Sciences, 15(1), 20–29. https://doi.org/doi:10.1002/elsc.201400172

Wang, G., Haringa, C., **Tang, W.**, Noorman, H., Chu, J., Zhuang, Y., & Zhang, S. (2019). Coupled Metabolic-hydrodynamic Modeling Enabling Rational Scale-up of Industrial Bioprocesses. Biotechnology and Bioengineering, 117(3), 844-867. https://doi.org/10.1002/bit.27243

Background

Over the last decades, large-scale industrial fermentations using a wide range of microorganisms, plant cells, and mammalian cells have been extensively employed for the production of commercial products such as food ingredients, nutraceuticals, biomaterials, vaccines, pharmaceuticals, biofuels and pigments (Festel, 2010; Otero and Nielsen, 2010).

A leading example has been the industrial penicillin manufacturing, which was initiated more than 70 years ago, and the productivity of the current industrial strain is 100,000 times higher than the original strain, by means of successive rounds of strain improvement, bioreactor design, and process optimization (Rokem et al., 2007). From a stoichiometric point of view, the maximum theoretical yield of penicillin on glucose (mole/mole) was calculated as 0.18 by VanGulik et al. (2001) and 0.5 by Jørgensen et al. (1995). However, the current yield of penicillin is still far away from either of those values. As an accepted fact, fermentation performance is largely driven by both strain characteristics and the environmental conditions of the cells. It is, therefore, common practice to engineer the strain and optimize the process together. To pinpoint potential targets for metabolic engineering, however, becomes at some stage the limiting step for getting closer to the theoretical maximum product yield (van Gulik et al., 2000). In this respect, it is of added value to get a thorough knowledge of the pathway kinetics. Pulse response experiments as of now are a very fast and useful instrument to gain knowledge of *in vivo* microbial kinetics. It also aids in identification of response mechanisms as well as metabolic bottlenecks for metabolic engineering (Spadiut et al., 2013).

It is known that time constants of enzyme-catalyzed reactions have nothing to do with the vessel size with which mixing time is positively correlated. In large-scale high-cell density fed-batch bioprocesses, environmental concentration gradients such as of substrate, pH, CO_2, and dissolved oxygen are prone to be induced, as long as the mixing time of the fermentor is larger than the relevant cellular reaction time. As a consequence, large-scale bioprocesses are often characterized by strong dynamics, with which cells are forced to cope (Lara et al., 2006). In many cases, substrate gradients lead to the formation of undesired byproducts and decreased product titers, but in contrast other cases proved that external perturbations, such as pulsed feeding, caused a substantial increase of the productivity (Bhargava et al., 2003; Bylund et al., 1998; Jazini and Herwig, 2014).

To understand these scale-up effects, the scale-down method is often advocated to efficiently take the local information of a large-scale bioprocess into consideration. After studies in the lab, an optimized scheme is then transferred to the production scale (Noorman, 2011). Computational fluid dynamics (CFD) is a powerful tool to simulate the detailed flow field of the fermentor and has already been applied for fermentor design and process optimization (Yang et al., 2012; Zou et al., 2012). As an example, in a 12 m^3 cephalosporin C production case, the effect of different impeller configurations on the cell morphology and cephalosporin C production was simulated in detail using CFD. It revealed that a novel impeller combination decreased power consumption and enhanced mass transfer as compared to the conventional configuration (Yang et al., 2012). The results of CFD simulations were also helpful in identifying the main cause of undesired scale-up effects. By using CFD for real-time fluid dynamic simulations of 50L and 132 m^3 fermentors applied for the erythromycin fermentation, it was found that the decrease of the oxygen transfer rate largely impaired cellular metabolism and production formation (Zou et al., 2012). Similar results were also observed in the production of glucoamylase from *Aspergillus niger* where extra benefits were gained via improved impeller settings (Tang et al., 2015).

As a direction for future development, the insight from a proper CFD model is essential for people to realize that a complete prediction and detailed description of industrial bioprocess requires combined efforts in microbial kinetic and CFD modeling.

Environmental gradients: causes and consequences

The production of secondary metabolites, including antibiotics, is often repressed in the presence of excess glucose, which is called catabolite repression. A fed-batch mode is thus introduced as an effective way to bypass this obstacle and it's also applied to avoid oxygen limitation in large-scale high-cell density fed-batch processes. The feed nutrient, often glucose, is then supplied at a growth limiting rate. To minimize dilution of the broth, this feed solution is fed in a highly concentrated form. Scale-up at constant variables such as volumetric mass transfer coefficient is performed, leading to heterogeneities at industrial scales (Bannari et al., 2008). Alternatively, if the liquid circulation time at production scales is kept the same as for the bench scale, an incredible power input would be required(Lara et al., 2006). Therefore, it is rather difficult to realize successful scale-up due to limitations of power input at industrial scale and a lack of knowledge about cellular response mechanisms. As is shown

in Table 1 and Table 2, in large-scale bioprocesses, the mixing time is tens or even hundreds of seconds (Canelas et al., 2008; Heijnen, 2010; Jolicoeur et al., 1992; Kiss et al., 1994; Langheinrich et al., 1998; Lara et al., 2006; Leckie et al., 1991; Nasution et al., 2006a; Nasution et al., 2006b; Schügerl, 1993; Sweere et al., 1987; Taymaz-Nikerel et al., 2009; Vrabel et al., 2000). This is longer than the relevant cellular reaction time; especially with respect to the intermediates of central metabolism, having turnover times in the order of seconds (Nasution, 2007). As a result, heterogeneities in the system inevitably occur. Previous studies revealed that substrate concentrations in the feed zone can reach values of dozens to even hundreds of times higher than in the low-concentration zones, and this difference may increase with higher biomass concentration (Bylund et al., 1998; Larsson et al., 1996). In this case, cells are periodically forced to circulate through famine and feast regions where substrate availability changes lead to rapid metabolite, flux and growth rate responses. Amounts, qualities, and yields of biomass and products will be severely affected and stress responses will also be induced (Enfors et al., 2001).

To effectively study the influence of environmental gradients, scale-down of industrial conditions is highly advocated for its convenience and efficiency (Noorman, 2011). As is shown in Table 3, a vast majority of scale-down studies has dealt with microbial fermentation (Amanullah et al., 2001; Baez et al., 2011; Junne et al., 2011; Lorantfy et al., 2013; Nienow et al., 2013; Sandoval-Basurto et al., 2005). Generally, scale-down of industrial practices is performed either through a combination of scale-down bioreactors or a special feed regime. With respect to scale-down devices, one-compartment (STR) and two-compartment systems (STR-STR/PFR, where PFR stands for plug flow reactor) are of great use in investigating the influence of external stimuli mimicked from those at production scales (Neubauer and Junne, 2010). As a case in point, using a two-compartment system (STR-PFR), oscillating dissolved oxygen and substrate concentration at the production-scale fermentations with *Bacillus subtilis* were simulated (Junne et al., 2011).The results indicated that the decrease of amino acid synthesis was due in large part to a metabolic shift toward ethanol formation. Käß et al. (2014) used a two-compartment scale-down system to study the influence of oxygen supply and substrate oscillations on cell metabolism. Broth cycling from an aerobic STR to an anaerobic PFR can be employed to simulate oxygen feast and famine zones at large scales and the residence time in the PFR can be adapted to simulate circulation times of large fermentors. It was found that oscillations within a minute range exerted insignificant impact on the metabolism of *Corynebacterium glutamicum*. In this

strain, a futile cycle of side products formation and subsequent consumption was found to enhance its metabolic robustness against process inhomogeneities, which facilitated its use in large-scale fed-batch applications (Käß et al., 2014). Apart from scale-down devices, applying a cyclic feed regime is frequently used for scale-down studies of large-scale gradients. As an example, the influence of a substrate concentration gradient on penicillin formation was investigated via a 6-min on-off feed cycle in a chemostat cultivation system. De Jonge et al. (2011) imposed this intermittent feed regime to study the influence of substrate gradients on process performance and cell metabolism. It was found that penicillin production was almost reduced by a factor of 2 because of fluctuations in energy levels in response to the glucose perturbations as compared with constant feed cultivations.

After all, it is the *in vivo* kinetics of relevant enzymes that determine which parameters need to be altered to obtain a desired change in a system (Teusink and Smid, 2006). It is, therefore, necessary to understand the mechanism of strain responses upon an external stimulus, in order to find potential targets for metabolic engineering. It is highly recommended that dynamic strategies such as shifts, ramps, pulses, and oscillations should be used for fast characterization of cell metabolism and process optimization (Spadiut et al., 2013). Pulse response experiments are typically carried out in a time frame of seconds to a few minutes. In this short time window, it can be assumed that the enzyme levels do not change. Therefore, changes in both intra- and extracellular metabolites are ascribed to rapid enzyme-metabolite interactions only. Carrying out pulse response experiments is thus an efficient and effective approach for obtaining *in vivo* kinetic information and estimating kinetic parameters for kinetic metabolic models (Visser et al., 2004).

Table 1. Reported mixing times in bioreactors.

Type of reactor	Cell line	Mixing time (s)	Reference
Cell culture			
5-L STR	CHO	2-5	(Kenty et al., 2005)
8.5-L STR	Plant cells	3.6	(Leckie et al., 1991)
11-L STR Helical ribbon	Plant cells	18-25	(Jolicoeur et al., 1992)
20-L STR	CHO	20-80	(Kenty et al., 2005)
8m^3 STR	Namalwa cells	40-200	(Langheinrich et al., 1998)
10m^3 STR	Plant cells	20-200	(Doran, 1993)
12m^3 STR	Mammalian cells	120-360	(Kiss et al., 1994)
10-L STR with spin filter hydrofoil impeller, 20 rpm	CHO	120	(Jem et al., 1994)
250-L STR with spin filter Hydrofoil impeller, 80 rpm	CHO	120	(Jem et al., 1994)
250-L STR with spin filter Pitched blade impeller, 80 rpm	CHO	1630	(Jem et al., 1994)
1m^3 STR with spin filter Hydrofoil impeller / mixing via spin filter	CHO	3120	(Jem et al., 1994)
15-L bubble column	Plant roots	2400	(Curtis, 2000)
10m^3 Airlift	Plant cells	200-1000	(Doran, 1993)
Microbial cultures			
12m^3 STR, equipped with 3 Rushton-type impellers	Microorganisms	10-50	(Vrabel et al., 2000)
12m^3 STR, equipped with 3 Scaba-type impellers	Microorganisms	10-30	(Vrabel et al., 2000)
30m^3 STR, equipped with 3 Rushton-type impellers	Microorganisms	125-250	(Vrabel et al., 2000)
30m^3 STR, equipped with 3 Scaba-type impellers	Microorganisms	70-110	(Vrabel et al., 2000)
2m^3 Bubble column	Microorganisms	18	(Schügerl, 1993)
2m^3 Airlift	Microorganisms	80	(Schügerl, 1993)
4m^3 Airlift tower loop	Baker's yeast	100-175	(Schügerl, 1993)
40m^3 Bubble column	Microorganisms	80	(Schügerl, 1993)
40m^3 Airlift	Microorganisms	101	(Schügerl, 1993)
150m^3 Bubble column	Baker's yeast	10-1000	(Sweere et al., 1987)

Table 2. Intracellular metabolite concentrations and turnover time in glucose-limited aerobic cultures of several organism (*Saccharomyces cerevisiae* from (Canelas et al., 2008), *Penicillium chrysogenum* from (Nasution et al., 2006b; Nasution et al., 2006a) and *Escherichia coli* from (Taymaz-Nikerel et al., 2009))

Metabolites	Intracellular level (µmol/gDW)			Turnover time (s)		
	P. chrysogenum	*S. cerevisiae*	*E. coli*	*P. chrysogenum*	*S. cerevisiae*	*E. coli*
Central metabolites						
G6P	4.64	5.2	1.42	23.3	17	3.6
F6P	0.71	1.4	0.38	5.7	7.3	1.2
T6P	0.55		0.13	47.8		
M6P	1.95		0.48			
6PG	0.25	0.48	0.1	3.7	4.5	1.1
Mannitol-1P			0.99			
G3P		0.13	0.17		57	13.1
FBP	0.9	0.64	0.82	7.2	3.2	2.5
F2,6bP	0.01		0.35			
2PG+3PG	0.59	2.8	1.65	2.3	6.6	2.5
PEP	0.24	2.3	1.61	0.9	5.7	2.7
Pyruvate	0.22	1.1	0.75	0.9	1.7	1.5
α-Ketoglutarate	2.05		0.31	22.1		0.6
Succinate	0.23	4.0	2.65	3.3	20	8.9
Fumarate	0.65	0.85	0.22	13.0	4.1	0.7
Malate	3.33	7.3	0.94	19.0	30	2.8
Amino acids						
Alanine	21.7	32	1.34	269	3268	76.7
Asparagine	1.5	4.7	0.58	459	1142	81.7
Aspartate	16.3	21	2.57	717	577	35.0
Glutamate	53.0	170	74.69	658	1112	229.0
Glutamine	28.7	64	6.14	1243	2401	80.0
Glycine	2.1	2.9	1.51	244	247	31.0
Histidine	0.72	6.0	0.15	432	3141	53.8
Isoleucine	0.33	1.6	0.11	111	140	12.9
Leucine	0.73	1.0	0.36	131	125	27.1
Methionine	0.14	0.20	0.05	58.8	66	10.5
Phenylalanine	0.19	1.6	0.13	61.2	430	23.8
Proline	0.95	3.9	0.66	206	925	101.4
Serine	5.7		0.53	453		8
Threonine	5.9	4.0	0.47	758	220	29.3
Tryptophan	0.11	0.51	0.02	130	788	11.9
Tyrosine	0.26	1.6	0.18	145	832	44.3
Valine	2.1	10	0.51	243	490	40.9
Ornithine		4.1	0.49		502	49.1
Adenine nucleotides						
ATP	7.39	7.0	5.95		1.4	2.0
ADP	1.03	1.3	2.31		1.4	2.0
AMP	0.27	0.6	0.91		3.1	9.4

Table 3. Environmental heterogeneity in bioprocess development

Organism	Gradients of	Compartment system	Reference
Penicillium chrysogenum	Substrate	STR	(de Jonge et al., 2011; Zhao et al., 2012)
Corynebacterium glutamicum	Substrate/oxygen	STR-PFR	(Käß et al., 2014)
Saccharomyces cerevisiae	Substrate	STR/STR-PFR	(Aboka et al., 2009; van Heerden et al., 2014)
Pichia pastoris	Oxygen	STR-STR	(Lorantfy et al., 2013)
Escherichia coli	Oxygen/carbon dioxide	STR-STR	(Baez et al., 2011; Sandoval-Basurto et al., 2005)
Bacillus subtilis	Substrate/oxygen/pH	STR-PFR	(Amanullah et al., 2001; Junne et al., 2011)
CHO cells	Fluid dynamic stress	STR-PFR	(Nienow et al., 2013)

Generation of *in vivo* kinetic properties of cellular metabolism by pulse response experiments

There has been a surge of the application of models in industrial bioprocesses to improve the understanding of the cellular metabolism and to identify genetic engineering targets to reach a desired product yield. The establishment of kinetic models of a cell factory is the first step for the integration of fluid dynamics and microbial kinetics. The fundamental workflow for the establishment of kinetic models has been excellently reviewed by Almquist et al. (2014). However, the establishment of metabolic models is often hampered by the lack of information on the *in vivo* enzyme kinetic properties of the metabolic reactions. Toward this end, it has been well proposed that *in vivo* studies using pulse response experiments could contribute to parameter estimation of enzyme kinetics (Visser et al., 2004). In parallel, the number of kinetic parameters should be reduced as much as possible by using approximative kinetic formats such as the linear-logarithmic approach (Heijnen, 2005). A successful kinetic model should be simple but yet complete enough to describe sufficient aspects of the dynamic reaction kinetics and also convenient to be integrated into CFD models.

Pulse or stimulus-response experiments are an ideal tool to obtain understanding of the *in vivo* regulation mechanisms of organisms to cope with external perturbations. In general, an external stimulus is imposed on a steady state chemostat system whereupon the perturbed metabolome is quantitatively acquired via well-established fast sampling and quenching protocols (van Gulik et al., 2012; Mashego et al., 2004). The BioScope is a very helpful tool in pulse response experiments and is used as a special PFR in which an external perturbing agent such as the growth-limiting substrate, often glucose, can be pulsed. For example, in a glucose-limited chemostat cultivation of *Saccharomyces cerevisiae*, Visser et al. (2004) conducted pulse response experiments using the BioScope with both glucose and ethanol as the perturbing agents. By tracing the concentration change of intracellular and extracellular metabolites as well as energy and reducing equivalents, it was suggested that fructose-1,6-biphosphate stimulated pyruvate kinase and PEP did not function in the allosteric regulation of phosphofructokinase. Further, the ethanol pulse might perturb the metabolism in an indirect way, which may be caused by the ratio variation of nicotinamide adenine dinucleotide (NAD) to its reduced form (NADH) (Vemuri et al., 2007). Pulse experiments are very helpful as well in understanding the robustness of organisms. Robustness, the ability to maintain a balanced or functional state in coping with environmental

changes or perturbations is the fundamental property of life (Stelling et al., 2004). Organisms in general have a specific robustness to survive unfavorable environments. Upon an external perturbation, different organisms may have similar but not the same responses. For instance, in the same steady state chemostat cultivation systems, *Penicillium chrysogenum* and *Saccharomyces cerevisiae* were transiently exposed to the same glucose pulse, but a more pronounced change of the glycolytic flux was observed in *Saccharomyces cerevisiae* (Mashego et al., 2006; Nasution et al., 2006b).

Metabolic network modeling

Techniques for modeling the metabolism of microorganisms have evolved over many decades. In general, these models can be categorized into two main groups: stoichiometry-based models and kinetics-based models.

Genome-scale metabolic models (GEMs) are the most widely used stoichiometry-based models for studying cell physiology (Gu et al., 2019; Kim et al., 2017). By setting up an appropriate objective function (e.g. maximum yield of product on substrate) and flux constraints, the GEMs can be used to predict the cellular flux distribution, the pathway variability (Mahadevan and Schilling, 2003) and can further be used to seek for potentially beneficial genome modifications (Burgard et al., 2003; Thiele and Palsson, 2010). Various computational approaches, featured by flux balance analysis and its variations, have been developed to characterize the steady-state flux solution space. The community has also developed high-quality tools to help maximize the applications of GEMs (Ebrahim et al., 2013; Heirendt et al., 2019; Wang et al., 2018b). Nevertheless, the optimal solutions obtained from these GEMs can easily deviate from the real flux distribution. To address this issue, it is advisable to include extra information/constraints, such as protein concentration (Lloyd et al., 2018; Sánchez et al., 2017), transcriptional regulation (Lerman et al., 2012), metabolic regulation, thermodynamics (Canelas et al., 2011; Niebel et al., 2019; Saa and Nielsen, 2015) and enzyme kinetics, and so forth. For instance, Sánchez et al. (2017) presented a method which enhances a *GEM* with *Enzymatic Constraints* using *Kinetic* and *Omics* data (GECKO). It accounts for enzymes as a part of reactions in a GEM, allowing that the simulated flux of each reaction does not exceed its maximum capacity. This methodology has been tested in a GEM of *Saccharomyces cerevisiae* and showed better predictability of maximum growth rates on different carbon sources as well as a more accurate growth rate shift when knocking out *NDI1* (mitochondrial NADH dehydrogenase). Recently, Niebel et al. (2019) developed a GEM for

Saccharomyces cerevisiae with additional upper limits on the Gibbs energy dissipation rate. By introducing this upper limit, the extra-constrained GEM was able to predict the physiology and intracellular metabolic fluxes beyond the critical growth rate, above which the Crabtree effect of *Saccharomyces cerevisiae* kicks in. These results also suggested that the predictability of GEM's can be improved by incorporation of additional constraints.

In contrast, kinetics-based models are often used to describe dynamic features that are not captured by GEMs. The history of using kinetic models for enzyme reactions can be referred to as early as the 1960s (Monod et al., 1963; Monod et al., 1965) when efforts were focused on single enzyme mechanisms. Till now, the feasibility and importance of detailed kinetic models of specific processes (Douma et al., 2010), single metabolic pathways (Deshmukh et al., 2015; Fujimoto et al., 2012), and large-scale kinetic models (Andreozzi et al., 2016; Kesten et al., 2015; Smallbone et al., 2013) have been demonstrated. The general feature of different model categories is shown in Figure 1. The ultimate goal of kinetic modeling is to build a genome-scale kinetic model which includes the kinetic properties of all enzyme catalyzed reactions (Smallbone et al., 2010). Douma et al. (2010) have built a kinetic model that specifically described the gene regulation process of penicillin production, which is strongly controlled by the extracellular glucose level. Their dynamic gene regulation model predicts the expression of isopenicillin-N synthase based on glucose repression, fast decay of mRNA encoding for this enzyme as well as the decay of the enzyme itself. By combining this gene-regulation model with a simple stoichiometric kinetic model, the biomass and penicillin production during steady state chemostat and fed-batch cultivations could be accurately described. Meanwhile, detailed large-scale kinetic models have also been developed. A detailed glycolysis kinetic model was presented (Smallbone et al., 2013) and further extended and improved by Kesten et al. (2015). This detailed kinetic model consists of 49 kinetic expressions with 164 parameters. The model was fitted to published steady-state data (Wu et al., 2006) and tested with dynamic experimental data obtained from glucose and ethanol pulse experiments. Although it was not possible to quantitatively reproduce the dynamic behavior of each metabolite, for example, the adenine nucleotides and total pyruvate, the overall predictability of this complex kinetic model was comprehensive.

In addition, application of kinetic models enabled a much clearer and quantitative description of process development in scale-down simulators and knowledge under these downscaling scenarios was largely extended with a

deeper understanding of cell adaption to environmental changes. For example, Anane et al. (2019) studied the response of *E. coli* to glucose and dissolved oxygen oscillations via a model-based process prediction. They accelerated phenotype characterization for parallel scale down fed-batch cultivations in mini-bioreactors and most importantly, this model-based framework puts the scale-down/scale-up in mind at the beginning of process development.

Among the scopes of kinetic modeling, simulation and performance prediction of the complete fermentation process are gathering more and more attraction (Noorman, 2011). A study hotspot in this area is the prediction of cellular performance in large-scale industrial fermentations werein the organism has to cope with gradients, due to insufficient mixing (Noorman, 2011; Noorman and Heijnen, 2017; Straathof et al., 2019; Wang et al., 2018a). Unfortunately, these industrial fermentation conditions are extremely difficult to capture due to the complexity and difficulty of measuring these gradients throughout an industrial scale bioreactor. Therefore, it will be highly valuable for such an industrial process if a proper model can describe these large-scale dynamic conditions and provide explicit and quantitative suggestions for strain and process improvement. The integration of microbial kinetics and fluid dynamics offered great potential in describing details in a hundred-ton fermentor (Delvigne et al., 2017; Haringa et al., 2018). For instance, Lapin et al. (2004) and Lapin et al. (2006) first successfully combined a CFD framework with an *E. coli* kinetic model. In their workflow, the original *E. coli* model from Chassagnole et al. (2002) was first simplified based on the time hierarchy of the metabolism to enable the integration within a CFD framework and meet the computational ability. However, kinetic expressions with only glucose and/or oxygen uptake applied in the CFD models have limited value for the evaluation of a complete bioprocess.

Nevertheless, such integration comes with a cost. In simple words, it requires the metabolic model, designed with the end mindset for hydrodynamic coupling, to be largely simplified while keeping the critical cell response properties still in (Noorman and Heijnen, 2017). Hence, there remains an urgent need for metabolically structured models with a minimum number of kinetic parameters to be estimated, which are yet complete enough to describe sufficient aspects of the dynamic reaction kinetics and are convenient to be integrated into CFD models.

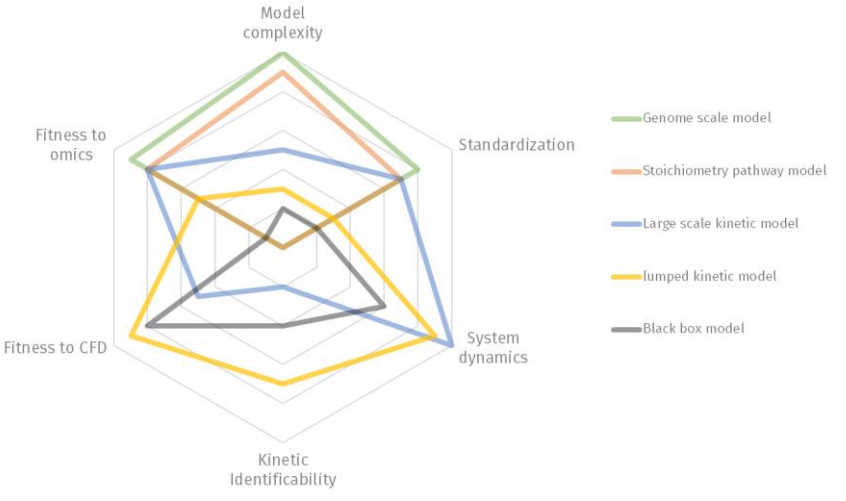

Figure 1 The general features of different model categories. Model complexity: the number of substances, conversion processes, kinetics involved in the model; fitness to omics: the potential of model performance improvement by further integration with omics data; fitness to CFD: the extendibility within a CFD framework for full-scale simulation; Kinetic Identifiability: the reliability of model defined reactions and metabolites as well as the accuracy and physical meaning of fitted parameters; System dynamics: the ability of predicting the dynamic responses of the metabolic system; Standardization: the acceptance of standard format throughout model development, preservation, easiness of access, tool development, and so forth. CFD, computational fluid dynamics

Integrating biotic and abiotic kinetic models

In the past few decades, unstructured metabolic models have been integrated with CFD models for simulation. In these approaches, the fluid phase was treated as continuum and the biophase was represented by a single chemical equation such as saturation-type kinetics of sugar uptake (Larsson et al., 1996). With the fast development of computer science and simulation algorithms, more detailed CFD models have already been established where Euler-Euler framing and Euler-Lagrange framing are the most frequently used approaches to describe the multiphase flow in bioreactors. However, both approaches suffered from high computational burden and required weeks for a full-process simulation. As an alternative, compartment based gradient estimation was explored with guidance from detailed CFD result. The integration of biokinetics with compartment model largely decreases the requirement of computational power and enables the full-tank, full-process simulation in real time.

Euler-Euler frame

In the Euler-Euler approach, the fluid phase and biophase are treated as continuum, and described in terms of their volume fractions (Buffo et al., 2012).

This approach has been applied in a number of bioreactors studies and also found wide applications in related fields (Kerdouss et al., 2006; Kurz et al., 2012; Sousa and Rangel, 2014; Zhang et al., 2009). For example, using segregated solutions of the Euler-Euler approach for the sake of less computational effort, the time course of production of gluconic acid was simulated and numerical results showed satisfactory agreement with the experimental data. It was found that as the biomass density incased, the mass transfer coefficient decreased and the gluconic acid production rate was reduced. This allowed to better understand the performance of this important bioprocess at different scale (Elqotbi et al., 2013). Moilanen et al.(2006) investigated the xanthan fermentation process in aerated fermentors by coupling gas-liquid mass transfer, xanthan bioreaction kinetics and non-Newtonian hydrodynamics with CFD. Gas-liquid hydrodynamics in xanthan fermentations including the bubble rise velocity and bubble size distribution was thus studied in detail. In a cellulase procdution case, an integrated model combining both the biomass kinetics and multiphase Euler-Euler formulation was successfully used to predict dynamic profiles such as the distributions of oxygen, cellulose and the shear stress within the fermentor (Bannari et al., 2012).

It is well known that population balance model (PBM), capable of illustrating the population heterogeneities, is often integrated with the Euler-Euler approach. For example, using the specific growth rate as a criterion, (Morchain et al., 2013) built a PBM to describe the heterogeneity of a cell population. It was shown that the model correctly represented the population growth rate dynamics. The model was used to predict the changes in the population growth rate as a response to the environmental change in the PFR. Scale-up effects have also been explained in detail by the model combining the Euler-Euler approach. PBM, and a kinetic model. Lab-scale and industrial-scale bioreactors were simulated by the model for comparison. It was confirmed that due to different time scales of mixing, mass transfer and chemical reactions, in bioreactors scale-up at a constant operating variable likely resulted in the formation of heterogeneities at a large scale (Morchain et al., 2014). The Euler-Euler approach is widely used, but here are several disadvantages associated with it. First, the biophase is treated as continuum, and it is therefore unable to distinguish between individual cells. As a consequence, it is impossible to describe the trajectory of a single cell in the flow field, which makes this approach less suitable for simulating the performance of the cell population upon an oscillating environment. Second, the cell state cannot be described accurately enough by just one or two parameters that can be used to distinguish

cell populations in PBM. Thus, a detailed description of the intracellular reaction network by PBM would lead to a high dimensional distribution function that is computationally intractable. As a consequence, such PBM has so far never been implemented (Fernandes et al., 2011; Henson et al., 2002).

Euler-Lagrange frame

In contrast, the Euler-Lagrange approach still treats the fluid phase as a continuum but the dispersed biophase is then tracked by the Lagrange approach (Lapin et al., 2010). As the biophase is present in the Lagrange frame, it becomes possible to track single cells and record the intracellular state as a function of time. As a result, the analysis of the lifelines of individual cells in space and time is possible. Recently, there are more studies applying the Euler-Lagrange approach for integration(Haringa et al., 2018; Lapin et al., 2004). By integrating a kinetic mode of glycolysis of yeast with the CFD model, Lapin et al. (2004) were the first who described temporal oscillations in glycolytic metabolites at the single-cell level in the presence of a spatially heterogeneous glucose concentration field (Wolf and Heinrich, 2000). To reduce the computational cost, the cell in this case represented a large collective of real cells. The simulation results showed that ideal mixing conditions contributed to synchronization of the individually autonomous oscillations at the population level, while in the presence of substate gradients, a dramatic loss of synchrony occurred. In order to further verify the integration method, a larger bioreactor was simulated with a more sophisticated cell kinetic model. The cell kinetic model contained a phosphotransferase system for the sugar uptake. The behavior of an *Escherichia coli* population in an oscillating environment was studied by integrating the CFD model with central carbon metabolism. Simulation results confirmed distinct difference in cell conditions that had been observed experimentally at different operation scales (Chassagnole et al., 2002; Lapin et al., 2006).

Compared to the Euler-Euler approach, the Euler-Lagrange approach is believed to be a more accurate method to describe the interaction between discrete cells and their environments. Cell population heterogeneity depends on the environments experienced at the single-cell level (Müller et al., 2010). The unique trajectory of a single cell can be depicted since the Lagrangian approach tracks single cells along their paths in bioreactors. Meanwhile, the effect of temporal and spatial environmental changes along the trajectory of single cells can be introduced as well.

Compartment-based integration

With either approach, the CFD considers detailed hydrodynamic features of a bioreactor. The computational burden constrains its application in real industrial settings. Even with considerable simplifications, weeks of simulation is still required for a complete fed-batch process. Compartment models form a middle ground between ideal reactor models and full CFD. Originally, those models were set up based on experimental data (Oosterhuis and Kossen, 1984; Vrabel et al., 1999; Vrábel et al., 2001). More recently, CFD simulation was used as a basis of constructing such a compartment model (Bezzo et al., 2003; Bezzo and Macchietto, 2004; Delafosse et al., 2014). With a significant simplification of the fluid field, the large-scale gradient can be estimated via integration of a black box model within seconds (Nadal-Rey et al., 2021b; Nadal-Rey et al., 2021a; Spann et al., 2019a; Spann et al., 2019b; Tajsoleiman et al., 2019). However, the assumption of instantaneous equilibrium of the cell's response to external conditions by the black box model is questionable. Similar to integration with CFD models, there are certain advantages of combining a structured kinetic model with a compartment model, via either population balances (Pigou et al., 2017; Pigou and Morchain, 2015) or parcel tracking (Delafosse et al., 2015).

The scope of this thesis

Bioprocesses use the power and versatility of nature via microorganisms that make bio-products from renewable feedstocks. Microorganisms can very well be engineered as efficient cell factories. However, the gap between the cell environment at lab and production scales is causing gross resource and asset utilization inefficiencies and is a barrier to fast and successful scale-up. Understanding the effect of environmental heterogeneities on cells in bioreactors is of great importance of response mechanisms upon an external perturbation and facilitate the establishment of pathway kinetics. A couple of scale-down systems have been devised to study the effect of environmental gradients and multiple, prolonged perturbations in large-scale practices. In parallel, cell kinetic models need to be designed and calibrated with a simple structure but yet should be efficient to describe the cell dynamics in selected scenarios. Still, kinetic models alone are not capable of interpreting the complex oscillating environment in the industrial scale bioreactor. Therefore, an extensive understanding of a real bioprocess needs the information of detailed flow field of the fermentor.

This thesis aims at developing a robust, model-based, simulation platform to speed up fermentation scale-up, driven by profound biological and physical understanding from series of continuous cultivations and representative scale-down cultivations. Further, we would like to demonstrate the application of computational approaches for better design of scale-down simulators. All of those will enable faster and more successful scale-up, improve the energy and resource efficiency of fermentations and accelerate bringing bio-innovations to the market.

The concept of using a simplified kinetic model to capture all essential dynamic features is demonstrated first, where a production process of penicillin with *Penicillium chrysogenum* is used as an example. The goal of **Chapter 2** is to assess the capability of a simplified structured model by using a smart lumping strategy for simplification purposes, with limited compromise on predictiability, relying on learnings from high-quality physiological and intracellular metabolomics data. Beginning with the end in mind, the scope of the developed model is limited to the impact of potential carbon gradients in large scale and get quantitatively correct predictions in a well-controlled feast-famine scale-down systems. We will combine the existing knowledge of penicillium kinetics (e.g. PAA transportation, penicillin pathway regulation, role of storage carbon, etc.) for a comprehensive cell model.

Then, further deploying the lumped model, in **Chapter 3**, the goal is to integrate it into a CFD framework using the Euler-Lagrange approach. In such a way, one can clearly see the cell's motion pattern coupled with its intracellular metabolic pattern. In this work package, we will compare multiple scenarios covering two types of kinetic models (black box and structured kinetic models) with both ideal-mixing scenarios and CFD at 54m3 industrial scale or 3L lab scale. For the computational aspect, we will explore different ways of coupling (1-way or 2-way). Finally, we will test both kinetic models in fed-batch fermentation mode to verify the model prediction against real experimental data.

In Chapter 4 and 5, we will present another important model strain, the yeast *Saccharomyces cerevisiae*. **Chapter 4** will focus on the regulation of storage carbon metabolism as these compounds have already shown to have important roles in multiple organisms, including *P.chrysogenum* and *S.cerevisiae*. Because of the focus on storage metabolites, we will largely simplify the remaining, lean part of the yeast (e.g. central metabolism, growth, overflow metabolism, etc.) as a black-box while describing the roles of different storage compounds, such as trehalose and glycogen explicitly. Nevertheless, the concept of this model is

also to be able to provide a smooth integration into CFD for large scale simulation, as the content of these storage compounds are also of industrial interest. Therefore, the developed model will also be tested by integration with previously obtained particle tracks via 1-way coupling.

In parallel with the kinetic model with a focus on storage, the goal of **Chapter 5** is to develop a lumped kinetic model with a similar structure as the model presented in Chapter 2, essentially a compact full cell model that is capable of describing the complete specific growth rate range (i.e., 0.005 1/h to 0.35 1/h). Instead of including a specific product pathway like in the *P.chrysogenum* model, we will focus more on describing *S.cerevisiae*'s ethanol formation properties under carbon overflow or oxygen limitation conditions, known as the Crabtree effect and Pasteur effects. This requires a proper description of the intracellular energy and redox levels by the model. The complete model will also be tested and validated against particle tracks obtained previously from a CFD simulation of a large scale *S. cerevisiae* fermentation process, using the Euler-Lagrange approach, proving its potential application in full-scale simulation. Further, this model can also serve as a scaffold for models wherein product formation is incorporated, to enable a faster and smoother modeling workflow for the future.

Finally, in **Chapter 6**, shortcomings of the model assessed in Chapter 5 are addressed and possible improvements are addressed in a qualitatively manner.

References

Aboka FO, Heijnen JJ, van Winden WA. 2009. Dynamic 13C-tracer study of storage carbohydrate pools in aerobic glucose-limited Saccharomyces cerevisiae confirms a rapid steady-state turnover and fast mobilization during a modest stepup in the glucose uptake rate. *FEMS Yeast Res* **9**:191–201.

Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J, Jirstrand M. 2014. Kinetic models in industrial biotechnology–improving cell factory performance. *Metab Eng* **24**:38–60.

Amanullah A, McFarlane CM, Emery AN, Nienow AW. 2001. Scale-down model to simulate spatial pH variations in large-scale bioreactors. *Biotechnol Bioeng* **73**:390–399.

Anane E, Sawatzki A, Neubauer P, Cruz-Bournazou MN. 2019. Modelling concentration gradients in fed-batch cultivations of E. coli–towards the flexible design of scale-down experiments. *Journal of Chemical Technology & Biotechnology* **94**:516–526.

Andreozzi S, Chakrabarti A, Soh KC, Burgard A, Yang TH, van Dien S, Miskovic L, Hatzimanikatis V. 2016. Identification of metabolic engineering targets for the enhancement of 1, 4-butanediol production in recombinant E. coli using large-scale kinetic models. *Metab Eng* **35**:148–159.

Baez A, Flores N, Bolivar F, Ramirez OT. 2011. Simulation of dissolved CO2 gradients in a scale-down system: A metabolic and transcriptional study of recombinant Escherichia coli. *Biotechnol J* **6**:959–967.

Bannari R, Bannari A, Vermette P, Proulx P. 2012. A model for cellulase production from Trichoderma reesei in an airlift reactor. *Biotechnol Bioeng* **109**:2025–2038.

Bannari R, Kerdouss F, Selma B, Bannari A, Proulx P. 2008. Three-dimensional mathematical modeling of dispersed two-phase flow using class method of population balance in bubble columns. *Comput Chem Eng* **32**:3224–3237.

Bezzo F, Macchietto S, Pantelides CC. 2003. General hybrid multizonal/CFD approach for bioreactor modeling. *AIChE Journal* **49**:2133–2148.

Bezzo F, Macchietto S. 2004. A general methodology for hybrid multizonal/CFD models: Part II. Automatic zoning. *Comput Chem Eng* **28**:513–525.

Bhargava S, Wenger KS, Marten MR. 2003. Pulsed addition of limiting-carbon during Aspergillus oryzae fermentation leads to improved productivity of a recombinant enzyme. *Biotechnol Bioeng* **82**:111–117.

Buffo A, Vanni M, Marchisio DL. 2012. Multidimensional population balance model for the simulation of turbulent gas–liquid systems in stirred tank reactors. *Chem Eng Sci* **70**:31–44.

Burgard AP, Pharkya P, Maranas CD. 2003. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* **84**:647–657.

Bylund F, Collet E, Enfors S-O, Larsson G. 1998. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. *Bioprocess Engineering* **18**:171–180.

Canelas AB, van Gulik WM, Heijnen JJ. 2008. Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions. *Biotechnol Bioeng* **100**:734–743.

Canelas AB, Ras C, ten Pierick A, van Gulik WM, Heijnen JJ. 2011. An in vivo data-driven framework for classification and quantification of enzyme kinetics and determination of apparent thermodynamic data. *Metab Eng* **13**:294–306. http://www.sciencedirect.com/science/article/pii/S1096717611000176.

Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. 2002. Dynamic modeling of the central carbon metabolism of Escherichia coli. *Biotechnol Bioeng* **79**:53–73.

Curtis WR. 2000. Encyclopedia of cell technology. *Hairy roots, bioreactor growth. New York: John Wiley & Sons*:827–841.

Delafosse A, Calvo S, Collignon M-L, Delvigne F, Crine M, Toye D. 2015. Euler--Lagrange approach to model heterogeneities in stirred tank bioreactors--Comparison to experimental flow characterization and particle tracking. *Chem Eng Sci* **134**:457–466.

Delafosse A, Collignon M-L, Calvo S, Delvigne F, Crine M, Thonart P, Toye D. 2014. CFD-based compartment model for description of mixing in bioreactors. *Chem Eng Sci* **106**:76–85.

Delvigne F, Takors R, Mudde R, van Gulik W, Noorman H. 2017. Bioprocess scale-up/down as integrative enabling technology: from fluid mechanics to systems biology and beyond. *Microb Biotechnol* **10**:1267–1274.

Deshmukh AT, Verheijen PJT, Seifar RM, Heijnen JJ, van Gulik WM. 2015. In vivo kinetic analysis of the penicillin biosynthesis pathway using PAA stimulus response experiments. *Metab Eng* **32**:155–173.

Doran PM. 1993. Design of reactors for plant cells and organs. *Bioprocess design and control*:115–168.

Douma RD, Verheijen PJT, de Laat WTAM, Heijnen JJ, van Gulik WM. 2010. Dynamic gene expression regulation model for growth and penicillin production in Penicillium chrysogenum. *Biotechnol Bioeng* **106**:608–618.

Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. 2013. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst Biol* **7**:74. https://doi.org/10.1186/1752-0509-7-74.

Elqotbi M, Vlaev SD, Montastruc L, Nikov I. 2013. CFD modelling of two-phase stirred bioreaction systems by segregated solution of the Euler–Euler model. *Comput Chem Eng* **48**:113–120.

Enfors S-O, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, Krüger E, Schweder T, Hamer G, O'beirne D, others. 2001. Physiological responses to mixing in large scale bioreactors. *J Biotechnol* **85**:175–185.

Fernandes RL, Nierychlo M, Lundin L, Pedersen AE, Tellez PEP, Dutta A, Carlquist M, Bolic A, Schäpper D, Brunetti AC, others. 2011. Experimental methods and modeling techniques for description of cell population heterogeneity. *Biotechnol Adv* **29**:575–599.

Festel G. 2010. Industrial biotechnology: Market size, company types, business models, and growth strategies. *Industrial Biotechnology* **6**:88–94.

Fujimoto N, Kosaka T, Yamada M. 2012. Menaquinone as well as ubiquinone as a crucial component in the Escherichia coli respiratory chain. *Chemical biology* **10**:187–208.

Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. 2019. Current status and applications of genome-scale metabolic models. *Genome Biol* **20**:121. https://doi.org/10.1186/s13059-019-1730-3.

van Gulik WM, de Laat W, Vinke JL, Heijnen JJ. 2000. Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. *Biotechnol Bioeng* **68**:602–618.

van Gulik WM, Canelas AB, Taymaz-Nikerel H, Douma RD, Jonge LP de, Heijnen JJ. 2012. Fast sampling of the cellular metabolome. In: . *Microbial Systems Biology*. Springer, pp. 279–306.

Haringa C, Mudde RF, Noorman HJ. 2018. From industrial fermentor to CFD-guided downscaling: what have we learned? *Biochem Eng J* **140**:57–71.

Haringa C, Tang W, Wang G, Deshmukh AT, van Winden WA, Chu J, van Gulik WM, Heijnen JJ, Mudde RF, Noorman HJ. 2018. Computational fluid dynamics simulation of an industrial P. chrysogenum fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization. *Chem Eng Sci* **175**:12–24. http://www.sciencedirect.com/science/article/pii/S0009250917305742.

van Heerden JH, Wortel MT, Bruggeman FJ, Heijnen JJ, Bollen YJM, Planqué R, Hulshof J, O'Toole TG, Wahl SA, Teusink B. 2014. Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science (1979)* **343**:1245114.

Heijnen JJ. 2010. Impact of thermodynamic principles in systems biology. In: . *Biosystems engineering II*. Springer, pp. 139–162.

Heijnen JJ. 2005. Approximative kinetic formats used in metabolic network modeling. *Biotechnol Bioeng* **91**:534–545.

Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V, Magnusdóttir S, Ng CY, Preciat G, Žagare A, Chan SHJ, Aurich MK, Clancy CM, Modamio J, Sauls JT, Noronha A, Bordbar A, Cousins B, el Assal DC, Valcarcel L v, Apaolaza I, Ghaderi S, Ahookhosh M, ben Guebila M, Kostromins A, Sompairac N, Le HM, Ma D, Sun Y, Wang L, Yurkovich JT, Oliveira MAP, Vuong PT, el Assal LP, Kuperstein I, Zinovyev A, Hinton HS, Bryant WA, Aragón Artacho FJ, Planes FJ, Stalidzans E, Maass A, Vempala S, Hucka M, Saunders MA, Maranas CD, Lewis NE, Sauter T, Palsson BØ, Thiele I, Fleming RMT. 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc* **14**:639–702. https://doi.org/10.1038/s41596-018-0098-2.

Henson MA, Müller D, Reuss M. 2002. Cell population modelling of yeast glycolytic oscillations. *Biochemical Journal* **368**:433–446.

Jazini M, Herwig C. 2014. Two-compartment processing as a tool to boost recombinant protein production. *Eng Life Sci* **14**:118–128.

Jem KJ, Fateen S, Michaels J. 1994. Mixing phenomena in industrial bioreactors with perfusion spin filters. In: . *Animal Cell Technology*. Elsevier, pp. 392–396.

Jolicoeur M, Chavarie C, Carreau PJ, Archambault J. 1992. Development of a helical-ribbon impeller bioreactor for high-density plant cell suspension culture. *Biotechnol Bioeng* **39**:511–521.

de Jonge LP, Buijs NAA, ten Pierick A, Deshmukh A, Zhao Z, Kiel JAKW, Heijnen JJ, van Gulik WM. 2011. Scale-down of penicillin production in Penicillium chrysogenum. *Biotechnol J* **6**:944–958.

Jørgensen H, Nielsen J, Villadsen J, Møllgaard H. 1995. Metabolic flux distributions in Penicillium chrysogenum during fed-batch cultivations. *Biotechnol Bioeng* **46**:117–131.

Junne S, Klingner A, Kabisch J, Schweder T, Neubauer P. 2011. A two-compartment bioreactor system made of commercial parts for bioprocess scale-down studies: impact of oscillations on Bacillus subtilis fed-batch cultivations.

Käß F, Junne S, Neubauer P, Wiechert W, Oldiges M. 2014. Process inhomogeneity leads to rapid side product turnover in cultivation of Corynebacterium glutamicum. *Microb Cell Fact* **13**:1–11.

Kenty B, Li Z, Vanden T, Lee S. 2005. Mixing in laboratory-scale bioreactors used for mammalian cell culture. In: . *the 229th American Chemical Society National Meeting*. San Diego, CA.

Kerdouss F, Bannari A, Proulx P. 2006. CFD modeling of gas dispersion and bubble size in a double turbine stirred tank. *Chem Eng Sci* **61**:3313–3322.

Kesten D, Kummer U, Sahle S, Hübner K. 2015. A new model for the aerobic metabolism of yeast allows the detailed analysis of the metabolic regulation during glucose pulse. *Biophys Chem* **206**:40–57.

Kim WJ, Kim HU, Lee SY. 2017. Current state and applications of microbial genome-scale metabolic models. *Curr Opin Syst Biol* **2**:10–18.

Kiss R, Croughan M, Trask J, Polastri G, Groenhout M, Banka A, Paul J, Koning-Bastiaan H. 1994. Mixing time characterization in large scale mammalian cell bioreactors. In: . *AIChE Annual Meeting*.

Kurz D, Schnell U, Scheffknecht G. 2012. CFD simulation of wood chip combustion on a grate using an Euler–Euler approach. *Combustion Theory and Modelling* **16**:251–273.

Langheinrich C, Nienow AW, Eddleston T, Stevenson NC, Emery AN, Clayton TM, Slater NKH. 1998. Liquid homogenization studies in animal cell bioreactors of up to 8 m3 in volume. *Food and bioproducts processing* **76**:107–116.

Lapin A, Klann M, Reuss M. 2010. Multi-scale spatio-temporal modeling: lifelines of microorganisms in bioreactors and tracking molecules in cells. In: . *Biosystems Engineering II*. Springer, pp. 23–43.

Lapin A, Müller D, Reuss M. 2004. Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler- Lagrange methods: Traveling along the lifelines of single cells. *Ind Eng Chem Res* **43**:4647–4656.

Lapin A, Schmid J, Reuss M. 2006. Modeling the dynamics of E. coli populations in the three-dimensional turbulent field of a stirred-tank bioreactor—A structured–segregated approach. *Chem Eng Sci* **61**:4783–4797.

Lara AR, Galindo E, Ram\'\irez OT, Palomares LA. 2006. Living with heterogeneities in bioreactors. *Mol Biotechnol* **34**:355–381.

Larsson G, Törnkvist M, Wernersson ES, Trägårdh C, Noorman H, Enfors S-O. 1996. Substrate gradients in bioreactors: origin and consequences. *Bioprocess Engineering* **14**:281–289.

Leckie F, Scragg AH, Cliffe KC. 1991. Effect of bioreactor design and agitator speed on the growth and alkaloid accumulation by cultures of Catharanthus roseus. *Enzyme Microb Technol* **13**:296–305.

Lerman JA, Hyduke DR, Latif H, Portnoy VA, Lewis NE, Orth JD, Schrimpe-Rutledge AC, Smith RD, Adkins JN, Zengler K, Palsson BO. 2012. In silico method for modelling metabolism and gene product expression at genome scale. *Nat Commun* **3**:929. http://dx.doi.org/10.1038/ncomms1928.

Lloyd CJ, Ebrahim A, Yang L, King ZA, Catoiu E, O'Brien EJ, Liu JK, Palsson BO. 2018. COBRAme: A computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput Biol* **14**:e1006302.

Lorantfy B, Jazini M, Herwig C. 2013. Investigation of the physiological response to oxygen limited process conditions of Pichia pastoris Mut+ strain using a two-compartment scale-down system. *J Biosci Bioeng* **116**:371–379.

Mahadevan R, Schilling CH. 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* **5**:264–276.

Mashego MR, Wu L, van Dam JC, Ras C, Vinke JL, van Winden WA, van Gulik WM, Heijnen JJ. 2004. MIRACLE: mass isotopomer ratio analysis of U-13C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol Bioeng* **85**:620–628.

Mashego MR, van Gulik WM, Vinke JL, Visser D, Heijnen JJ. 2006. In vivo kinetics with rapid perturbation experiments in Saccharomyces cerevisiae using a second-generation BioScope. *Metab Eng* **8**:370–383.

Moilanen P, Laakkonen M, Aittamaa J. 2006. Modeling aerated fermenters with computational fluid dynamics. *Ind Eng Chem Res* **45**:8656–8663.

Monod J, Wyman J, Changeux J-P. 1965. On the nature of allosteric transitions: a plausible model. *J Mol Biol* **12**:88–118.

Monod J, Changeux J-P, Jacob F. 1963. Allosteric proteins and cellular control systems. *J Mol Biol* **6**:306–329.

Morchain J, Gabelle J-C, Cockx A. 2013. Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. *AIChE Journal* **59**:369–379.

Morchain J, Gabelle J-C, Cockx A. 2014. A coupled population balance model and CFD approach for the simulation of mixing issues in lab-scale and industrial bioreactors. *AIChE Journal* **60**:27–40.

Müller S, Harms H, Bley T. 2010. Origin and analysis of microbial population heterogeneity in bioprocesses. *Curr Opin Biotechnol* **21**:100–113.

Nadal-Rey G, McClure DD, Kavanagh JM, Cassells B, Cornelissen S, Fletcher DF, Gernaey K v. 2021a. Development of dynamic compartment models for industrial aerobic fed-batch fermentation processes. *Chemical Engineering Journal* **420**:130402.

Nadal-Rey G, McClure DD, Kavanagh JM, Cornelissen S, Fletcher DF, Gernaey K v. 2021b. Understanding gradients in industrial bioreactors. *Biotechnol Adv* **46**:107660.

Nasution U, van Gulik WM, Kleijn RJ, van Winden WA, Proell A, Heijnen JJ. 2006a. Measurement of intracellular metabolites of primary metabolism and adenine nucleotides in chemostat cultivated Penicillium chrysogenum. *Biotechnol Bioeng* **94**:159–166.

Nasution U, van Gulik WM, Proell A, van Winden WA, Heijnen JJ. 2006b. Generating short-term kinetic responses of primary metabolism of Penicillium chrysogenum through glucose perturbation in the bioscope mini reactor. *Metab Eng* **8**:395–405.

Nasution UDM. 2007. A dynamic and steady state metabolome study of central metabolism and its relation with the penicillin biosynthesis pathway in Penicillium chrysogenum.

Neubauer P, Junne S. 2010. Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Curr Opin Biotechnol* **21**:114–121.

Niebel B, Leupold S, Heinemann M. 2019. An upper limit on Gibbs energy dissipation governs cellular metabolism. *Nat Metab* **1**:125–132.

Nienow AW, Scott WH, Hewitt CJ, Thomas CR, Lewis G, Amanullah A, Kiss R, Meier SJ. 2013. Scale-down studies for assessing the impact of different stress parameters on growth and product quality during animal cell culture. *Chemical Engineering Research and Design* **91**:2265–2274.

Noorman H. 2011. An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnol J* **6**:934–943.

Noorman HJ, Heijnen JJ. 2017. Biochemical engineering's grand adventure. *Chem Eng Sci* **170**:677–693.

Oosterhuis NMG, Kossen NWF. 1984. Dissolved oxygen concentration profiles in a production-scale bioreactor. *Biotechnol. Bioeng.;(United States)* **26**.

Otero JM, Nielsen J. 2010. Industrial systems biology. *Biotechnol Bioeng* **105**:439–460.

Pigou M, Morchain J. 2015. Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chem Eng Sci* **126**:267–282.

Pigou M, Morchain J, Fede P, Penet M-I, Laronze G. 2017. An assessment of methods of moments for the simulation of population dynamics in large-scale bioreactors. *Chem Eng Sci* **171**:218–232.

Rokem JS, Lantz AE, Nielsen J. 2007. Systems biology of antibiotic production by microorganisms. *Nat Prod Rep* **24**:1262–1287.

Saa P, Nielsen LK. 2015. A general framework for thermodynamically consistent parameterization and efficient sampling of enzymatic reactions. *PLoS Comput Biol* **11**:e1004195.

Sánchez BJ, Zhang C, Nilsson A, Lahtvee P, Kerkhoven EJ, Nielsen J. 2017. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol Syst Biol* **13**:935. http://msb.embopress.org/content/msb/13/8/935.full.pdf.

Sandoval-Basurto EA, Gosset G, Bol\'\ivar F, Ram\'\irez OT. 2005. Culture of Escherichia coli under dissolved oxygen gradients simulated in a two-compartment scale-down system: Metabolic response and production of recombinant protein. *Biotechnol Bioeng* **89**:453–463.

Schügerl K. 1993. Comparison of different bioreactor performances. *Bioprocess Engineering* **9**:215–223.

Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E, Swainston N, Dada JO, Khan F, others. 2013. A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. *FEBS Lett* **587**:2832–2841.

Smallbone K, Simeonidis E, Swainston N, Mendes P. 2010. Towards a genome-scale kinetic model of cellular metabolism. *BMC Syst Biol* **4**:1–9.

Sousa T, Rangel CM. 2014. A dynamic two phase flow model for a pilot scale sodium borohydride hydrogen generation reactor. *Int J Hydrogen Energy* **39**:5291–5300.

Spadiut O, Rittmann S, Dietzsch C, Herwig C. 2013. Dynamic process conditions in bioprocess development. *Eng Life Sci* **13**:88–101.

Spann R, Gernaey K v, Sin G. 2019a. A compartment model for risk-based monitoring of lactic acid bacteria cultivations. *Biochem Eng J* **151**:107293.

Spann R, Glibstrup J, Pellicer-Alborch K, Junne S, Neubauer P, Roca C, Kold D, Lantz AE, Sin G, Gernaey K v, others. 2019b. CFD predicted pH gradients in lactic acid bacteria cultivations. *Biotechnol Bioeng* **116**:769–780.

Stelling J, Sauer U, Szallasi Z, Doyle III FJ, Doyle J. 2004. Robustness of cellular functions. *Cell* **118**:675–685.

Straathof AJJ, Wahl SA, Benjamin KR, Takors R, Wierckx N, Noorman HJ. 2019. Grand research challenges for sustainable industrial biotechnology. *Trends Biotechnol* **37**:1042–1050.

Sweere APJ, Luyben KCAM, Kossen NWF. 1987. Regime analysis and scale-down: tools to investigate the performance of bioreactors. *Enzyme Microb Technol* **9**:386–398.

Tajsoleiman T, Spann R, Bach C, Gernaey K v, Huusom JK, Krühne U. 2019. A CFD based automatic method for compartment model development. *Comput Chem Eng* **123**:236–245.

Tang W, Pan A, Lu H, Xia J, Zhuang Y, Zhang S, Chu J, Noorman H. 2015. Improvement of glucoamylase production using axial impellers with low power consumption and homogeneous mass transfer. *Biochem Eng J* **99**:167–176. http://www.sciencedirect.com/science/article/pii/S1369703X15001254.

Taymaz-Nikerel H, de Mey M, Ras C, ten Pierick A, Seifar RM, van Dam JC, Heijnen JJ, van Gulik WM. 2009. Development and application of a differential method for reliable metabolome analysis in Escherichia coli. *Anal Biochem* **386**:9–19.

Teusink B, Smid EJ. 2006. Modelling strategies for the industrial exploitation of lactic acid bacteria. *Nat Rev Microbiol* **4**:46–56.

Thiele I, Palsson BØ. 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* **5**:93–121.

VanGulik WM, Antoniewicz MR, Delaat W, Vinke JL, Heijnen JJ. 2001. Energetics of growth and penicillin production in a high-producing strain of Penicillium chrysogenum. *Biotechnol Bioeng* **72**:185–193.

Vemuri GN, Eiteman MA, McEwen JE, Olsson L, Nielsen J. 2007. Increasing NADH oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences* **104**:2402–2407.

Visser D, van Zuylen GA, van Dam JC, Eman MR, Pröll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004. Analysis of in vivo kinetics of glycolysis in aerobic Saccharomyces cerevisiae by application of glucose and ethanol pulses. *Biotechnol Bioeng* **88**:157–167.

Vrabel P, der Lans R, Cui YQ, Luyben KCAM. 1999. Compartment model approach: Mixing in large scale aerated reactors with multiple impellers. *Chemical Engineering Research and Design* **77**:291–302.

Vrabel P, van der Lans RGJM, Luyben KCAM, Boon L, Nienow AW. 2000. Mixing in large-scale vessels stirred with multiple radial or radial and axial up-pumping impellers: modelling and measurements. *Chem Eng Sci* **55**:5881–5896.

Vrábel P, van der Lans RGJM, van der Schot FN, Luyben KCAM, Xu B, Enfors S-O. 2001. CMA: integration of fluid dynamics and microbial kinetics in modelling of large-scale fermentations. *Chemical engineering journal* **84**:463–474.

Wang G, Zhao J, Haringa C, Tang W, Xia J, Chu J, Zhuang Y, Zhang S, Deshmukh AT, van Gulik W, Heijnen JJ, Noorman HJ. 2018a. Comparative performance of different scale-down simulators of substrate gradients in Penicillium chrysogenum cultures: the need of a biological systems response analysis. *Microb Biotechnol* **11**:486–497. https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1751-7915.13046.

Wang H, Marcišauskas S, Sánchez BJ, Domenzain I, Hermansson D, Agren R, Nielsen J, Kerkhoven EJ. 2018b. RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor. *PLoS Comput Biol* **14**:e1006541. https://doi.org/10.1371/journal.pcbi.1006541.

Wolf J, Heinrich R. 2000. Effect of cellular interaction on glycolytic oscillations in yeast: a theoretical investigation. *Biochemical Journal* **345**:321–334.

Wu L, van Dam J, Schipper D, Kresnowati MTAP, Proell AM, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. 2006. Short-term metabolome dynamics and carbon, electron, and ATP balances in chemostat-grown Saccharomyces cerevisiae CEN. PK 113-7D following a glucose pulse. *Appl Environ Microbiol* **72**:3566–3577.

Yang Y, Xia J, Li J, Chu J, Li L, Wang Y, Zhuang Y, Zhang S. 2012. A novel impeller configuration to improve fungal physiology performance and energy conservation for cephalosporin C production. *J Biotechnol* **161**:250–256.

Zhang D, Deen NG, Kuipers JAM. 2009. Euler- Euler modeling of flow, mass transfer, and chemical reaction in a bubble column. *Ind Eng Chem Res* **48**:47–57.

Zhao Z, ten Pierick A, de Jonge L, Heijnen JJ, Wahl SA. 2012. Substrate cycles in Penicillium chrysogenum quantified by isotopic non-stationary flux analysis. *Microb Cell Fact* **11**:1–14.

Zou X, Xia J, Chu J, Zhuang Y, Zhang S. 2012. Real-time fluid dynamics investigation and physiological response for erythromycin fermentation scale-up from 50 L to 132 m3 fermenter. *Bioprocess Biosyst Eng* **35**:789–800.

Chapter 2

A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by *Penicillium chrysogenum*

Published as

Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J. J., Xia, J., Chu, J., & Noorman, H. J. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by *Penicillium chrysogenum*. Biotechnology and Bioengineering, 114(8), 1733–1743. https://doi.org/10.1002/bit.26294

Abstract

A powerful approach for the optimization of industrial bioprocesses is to perform detailed simulations integrating large scale computational fluid dynamics (CFD) and cellular reaction dynamics (CRD). However, complex metabolic kinetic models containing a large number of equations pose formidable challenges in CFD-CRD coupling and computation time afterward. This necessitates to formulate a relatively simple but yet representative model structure. Such a kinetic model should be able to reproduce metabolic responses for short-term (mixing time scale of tens of seconds) and long-term (fed-batch cultivation of hours/days) dynamics in industrial bioprocesses. In this paper, we used *Penicillium chrysogenum* as a model system and developed a metabolically structured kinetic model for growth and production. By lumping the most important intracellular metabolites in 5 pools and 4 intracellular enzyme pools, linked by 10 reactions, we succeeded in keeping a relatively simple model structure, which gives informative insight in the state of the organism. The performance of this 9-pool model was validated with a periodic glucose feast-famine cycle experiment at the minute time scale. Comparison of this model and a reported black box model of this strain showed the necessities of the structured model under feast-famine conditions. This proposed model provides deeper insight into the *in vivo* kinetics and, most importantly, and can be easily integrated into a computational fluid dynamic framework for simulating complete fermentation performance and cell population dynamics in large scale and small scale fermentors.

Keywords: structured model, black box model, feast-famine, kinetics, *Penicillium chrysogenum*

Abbrivations

$a_{cell,3.1}$	Specific area of cell (m^2/CmolX)
C_i	Extracellular concentration of compound I (moli/kg)
c_i	Kinetic constant (reaction dependent)
D	Dilution rate (1/h)
k_{di}	Degradation rate of enzyme i (1/h)
$k_{perm,3.1}$	membrane permeability constant for PAA (m/h)
K_i	Affinity constant (moli/gDW for intracellular compounds and moli/kg for extracellular compounds)
k_i	Kinetic constant (reaction dependent)
$m_{3.3}$	Fitted parameter for penicillin production kinetic (-)
$m_{ATP,2.2}$	Cell maintenance in terms of ATP (molATP/CmolX/h)
m_{PAA}	Cell maintenance on PAA futile cycle in terms of ATP (molATP/CmolX)
m_S	Cell maintenance in terms of glucose (molglc/CmolX/h)
pH_{ext}	Extracellular pH (-)
pH_{int}	Intracellular pH (-)
pK_{PAA}	pK value of phenylacetic acid (-)
q_i	Biomass specific uptake/consumption rate of compound I (moli/CmolX/h)
R	Sum of residual error of the cost function (-)
v_i	Rate of reaction i (moli/CmolX/h)
X_i	Intracellular concentration of compound I (moli/gDW)
$Y_{X/S,max}$	Maximum biomass yields on glucose (molglc/CmolX)
$Y_{P/S,max}$	Maximum production yields on glucose (molglc/molPenG)
α	Kinetic constant (reaction dependent)
β	Kinetic constant (reaction dependent)
γ	Redox level (-)
δ_0	Basis rate in glucose transport kinetic (1/h)
μ	Specific growth rate (1/h)

Introduction

The emerging fields of systems biology and synthetic biology aim at achieving a more fundamental understanding of biological complexity through system-level analysis (Kleijn et al. 2007), and enhancing fermentative productivity of desired compounds by pinpointing and relieving bottlenecks in the metabolic networks (van Gulik et al. 2000). However, a direct utilization of the observations made in lab-scale cultivations could be problematic due to 'scale-up effects' (Noorman 2011). In order to capture the complex dynamics of the microbial metabolism in large-scale bioreactors, there has been an increasing focus on the use of high-resolution mathematical models for a rational bioprocess design and optimization (Vasilakou et al. 2016; Wang et al. 2015). A proposed powerful approach to accomplish this is by integrating computational fluid dynamics (CFD) models and cellular reaction dynamics (CRD) models (Lapin et al. 2004; Tyo et al. 2010). In this way, lifelines for individual cells can be obtained for high-precision scale-up/down investigation and population heterogeneity studies.

There are several kinds of CRD models describing whole cell physiologies that could be considered for coupling with CFD. The most complex models are genome-scale metabolic flux models (GEMs) which contain a huge amount of intracellular details (Lu et al. 2016; Österlund et al. 2012). For genome/large-scale metabolic flux models, the step towards dynamic simulation poses a major challenge due to the limited mechanistic *in vivo* kinetic knowledge of each reaction. On the other side, the conventional black box kinetic model is able to respond to the local residual glucose concentration with minimum model complexity, by neglecting all intracellular metabolic details. Using a black box model, Larsson et al. (1996) simulated the glucose concentration gradient in a 30 m^3 cultivation of *Saccharomyces cerevisiae* by integrating CFD and a hyperbolic glucose uptake model. This simple method predicted a spatial glucose gradient that was favorably validated with experimental data. This methodology guided the design of scale-down simulators and further study of the influence of glucose perturbation on the performance of *Saccharomyces cerevisiae* cultivation (Bylund et al. 1999). Still, the black box model has limitations: the cell's individual 'experiences', or 'life-lines' are not taken into consideration (Haringa et al. 2016; Lapin et al. 2010; Lapin et al. 2006). Nowadays, structured metabolic models are being developed and Kerkhoven et al. (2015) reviewed several of such kinetic models for yeast. Typically, they feature only one particular pathway of interest, outside of the context of the whole network and the strongly regulated and coherent response of the key metabolic functions. These more detailed metabolic kinetic models have their

associated issues, especially in parameter estimation (Hynne et al. 2001; Rizzi et al. 1997; Smallbone et al. 2013; Theobald et al. 1997; van Eunen et al. 2012). Efforts have been made to describe these kinetics in a more reliable way rather than using the classical hyperbolic enzyme kinetics. Heijnen (2005) reviewed six approximative kinetic formats from the perspective of metabolic modeling efficiency. Later, the most favored lin-log approach was successfully applied in the kinetic modelling of *P.chrysogenum* (Nikerel et al. 2012). All in all, the selection criteria of preferred CRD models for CFD coupling is a trade-off between purpose, complexity, simulation time frame and the data availability for parameter identification (Craven et al. 2013).

In this article, we use *P.chrysogenum* producing penicillin G as a model organism, and propose a metabolically structured model based on a published stoichiometric model and the concept of connecting important reactions in five lumped metabolite pools that are most sensitive to highly dynamic extracellular environment inside large-scale bioreactors. We demonstrated the applicability of our model in predicting the most relevant metabolic dynamics from the second to day timescale. Also its advantage to a classic black box model is addressed. With a simple enough structure to be further integrated into a CFD framework, this 9-pool structured kinetic model make afterward complete simulation of large scale fermentation possible.

Materials and Methods

Strain

The high yield *Penicillium chrysogenum* strain DS17690 was kindly donated by DSM Biotechnology Center (Delft, The Netherlands). This strain has been previously characterized in numerous studies (Berg 2013; de Jonge et al. 2011; Deshmukh et al. 2015; Douma et al. 2010c; Harris et al. 2006; van Gulik et al. 2001; van Gulik et al. 2000).

Data source

Steady state data (van Gulik et al. 2000): Biomass specific rates under glucose-limited chemostat conditions, covering a wide range of dilution rates, is available from van Gulik et al. (2000).

Feed ramp experiment (this study): To acquire intracellular information under slow dynamic conditions, a ramp experiment was performed where the dilution rate was linearly decreased from 0.05 1/h to 0.005 1/h in 100h (Ramp phase). The ramp was started after the system initially reached a steady state at 0.05

1/h. During the ramp phase, the fermentation settings remained the same as during the chemostat phase, except the air flow rate was adjusted to from 2.0 to 0.2L/min to get a reliable offgas O_2/CO_2 information at low feed rate. The dissolved oxygen concentration remained well above 0.2 mol/m^3 (80% of air saturation) throughout the process.

Feast-famine experiment (de Jonge et al. 2011): The experiments were carried out with periodic glucose feeding, with cycles of 360 seconds. The glucose was feed for the first 36 seconds in each cycle resulting in glucose accumulation in the broth and consumption (feast phase) during the first 180 seconds and glucose starvation (famine phase) for the remaining 180 seconds. During the 6 minutes cycle, datasets of online and offline measurements and extra and intracellular metabolites was collected.

Both the steady-state and ramp data were used for parameter estimation and model validation was performed using data from feast-famine experiments.

Medium

The composition of the chemostat medium was designed to support a steady state biomass concentration of approximately 5.6g/L dry weight, containing 15.0g/L glucose. 5.0g/L $(NH_4)_2SO_4$, 1.0g/L KH_2PO_4, 0.5g/L $MgSO_4 \cdot 7H_2O$, 0.68g/L phenylacetic acid (PAA) and 2ml/L of a trace element solution (Douma et al. 2010b). The same medium but with 0.41g/L PAA was used for the batch phase. For the details in medium preparation procedure, please refer to (Douma et al. 2010b)

Bioreactor setup

Aerobic glucose-limited chemostat cultures (pH=6.5, 25°C) of 4L working volume were carried out in a 7L turbine stirred bioreactor (Applikon, Schiedam, The Netherlands) at a dilution rate of 0.05 1/h. The aeration rate was 2L/min with a headspace overpressure at 0.3 bars and the stirrer speed was 500 rpm. Offgas was passed through a gas analyzer (NGA2000, Rosemount, USA) for the O_2 and CO_2 concentration measurement.

Dry weight

Dry cell weight was measured in triplicate, using glass fiber filters (type A/E; Pall Corporation, East Hills, NY; 47 mm with 1 µm pore size). The filters were pre-dried at 70°C until a constant weight was achieved and loaded with 5 g broth and washed with 10 mL demineralized water. The filters were then dried for 24 h at 70°C and weighted afterwards.

Rapid Sampling

Rapid sampling for extracellular glucose, penicillin and PAA was conducted using the cold steel-bead method described previously (Mashego et al. 2003).

Rapid sampling for intracellular metabolites was carried out through a specially designed rapid sampling device (Lange et al. 2001), with the sample processing procedure of de Jonge et al. (2012) for quantitative metabolite analysis, using U-^{13}C-labeled cell extract as internal standards (Wu et al. 2005).

Metabolite analysis

Intermediates of the glycolysis, TCA cycles, PP pathway and amino acids were measured by a GC-MS platform (Cipollina et al. 2009). Adenine nucleotides and penicillin pathway intermediates were analyzed by LC-MS/MS (Douma et al. 2010a; Nasution et al. 2006; Seifar et al. 2012). Extracellular glucose is also measured by the GC-MS platform according to de Jonge et al. (2013).

Simulation, parameter optimization, model validation

Specific rates and intra-/extracellular metabolite concentrations obtained from previous chemostat cultivations (van Gulik et al. 2000) and the ramp feed experiment were used to estimate the 31 free parameters of the kinetic model. Parameter estimation was carried out in Matlab R2015a by minimizing the relative error of experimental value and predicted value (Eq.1).

$$R = \sum_i \left| \frac{y_{sim,i} - y_{exp,i}}{y_{exp,i}} \right| \qquad \text{Eq.1}$$

The standard error of each estimated parameter was calculated using an earlier reported method (Dolan et al. 2007). The parameterized model was validated by simulating the feast-famine experiments published by de Jonge et al. (2011).

Theoretical Aspect

The 9-pool structured model

The metabolically structured model was derived from a previously published stoichiometric model for the same strain (van Gulik et al. 2000). Metabolites were lumped into five metabolite pools: glycolytic intermediates (Glyc), amino acids (AA), ATP, PAA and stored carbohydrates (Sto). These five pools were chosen:

1) To account for different response time scales: glycolytic intermediates and ATP pools have a time scale of seconds, amino acids have a time scale of minutes and Sto pool has a time scale of hours or days.

2) To allow the deployment of internal cellular resources, such as amino acids and stored carbohydrates in the absence of extracellular glucose.

3) To allow lumped reactions. For example, penicillin production is derived from amino acids, precursor and ATP; biomass production is based on the glycolytic pool, amino acids and ATP.

Cells compartments were not included as it is not possible to differentiate between metabolites from different compartments in the cell. Relationships among these 5 intermediate pools were established by defining 10 lumped reactions, using hyperbolic equations for most rates. In addition, for 4 reactions (glucose uptake, PAA export, penicillin production and storage conversion), the maximum rate capacity was not constant. Therefore, 4 enzyme pools were defined to allow variable capacities. The 9-pool structured model contains five lumped intracellular metabolite pools, four enzyme (capacity) pools and ten extracellular components (See Table I), connected via ten intracellular reactions (See Figure 1).

Detailed information on the pools' elemental composition, model stoichiometry, the kinetics and the previously available metabolic parameter values can be found in Supplement A. To compare the model performance with a unstructured model, a classic black box model for this strain developed by (Douma et al. 2010c) is shown in Supplement B. Intracellular and extracellular mass balances, which are the basis of our simulation model, are included in Supplement C. Measured intracellular metabolite concentrations for the metabolites that formed the five model pools for fitting and validation purposes are listed in Supplement D.

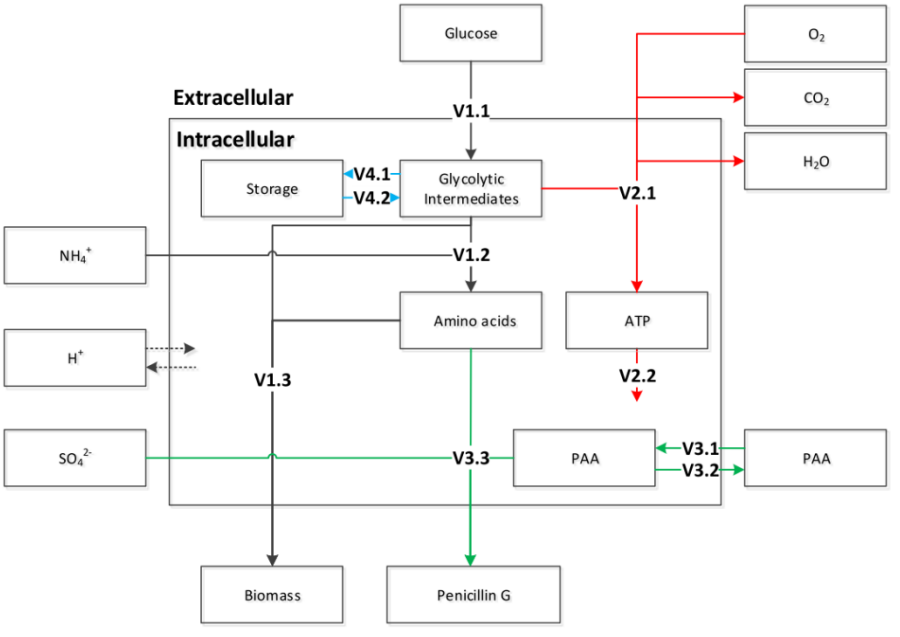

Figure 1 Overview of the 9-pool model for *P.chrysogenum*. Carbon metabolism is shown in black, energy metabolism in red, penicillin production in green, storage carbon cycling in blue.

Table I Intra- and extracellular pool definition, element composition, charge condition, redox level and approximate intracellular turnover time

	Pool name	Expression	Element composition[a]	γ-value	Approx. turnover time (s) (de Jonge 2016; Nasution et al. 2008)	Predicted turnover time at μ=0.05 1/h[b]
Intracellular	Glycolytic	Glyc	CH_2O	4	~5-10	7
	Amino acids	AA	$CH_{2.00}O_{0.70}N_{0.32}$	3.64	~500-3000	2800
	ATP	ATP	n.a.	n.a.[c]	~3-5	3
	PAA	PAA	$C_8H_7O_2^-$	36	~10-20	16
	Stored carbohydrates	Sto	$CH_{2.24}O_{0.98}$	4.28	~10000	12000
Enzyme[d]	Glucose uptake	$X_{E,1.1}$	n.a.	n.a.	n.a.	n.a.
	PAA export	$X_{E,3.2}$	n.a.	n.a.	n.a.	n.a.
	Penicillin conversion	$X_{E,3.3}$	n.a.	n.a.	n.a.	n.a.
	Storage conversion	$X_{E,4}$	n.a.	n.a.	n.a.	n.a.
Extracellular	Glucose	Glc	$C_6H_{12}O_6$	24	n.a.	n.a.
	Oxygen	O_2	O_2	0	n.a.	n.a.
	Carbon dioxide	CO_2	CO_2	0	n.a.	n.a.
	Ammonia	NH_4	NH_4^+	0	n.a.	n.a.
	Sulfuric acid	SO_4	SO_4^{2-}	0	n.a.	n.a.
	PAA	PAA	$C_8H_7O_2^-$	36	n.a.	n.a.
	Biomass	X	$CH_{1.79}O_{0.59}N_{0.16}$	4.13	n.a.	n.a.
	Penicillin G	P	$C_{16}H_{17}O_4N_2S^-$	74	n.a.	n.a.
	Water	H_2O	H_2O	0	n.a.	n.a.
	Proton	H	H^+	0	n.a.	n.a.

a: please reference Supplement A for more details.
b: Turnover time is calculated based on the pools and the influx(outflux) under steady state.
c: not available.
d: Capacities of reactions

Model extension with variable enzyme pools

Previous literature reported that enzyme levels are not constant but depend on growth rate (Heijnen and Romein 1995). Experimental results from the feed-ramp experiment also revealed an increase in both the extracellular glucose concentration and intracellular PAA concentration, suggesting a loss in glucose uptake capacity and PAA export capacity. As the 9-pool model is designed for describing a wide range of growth rates, it is necessary to have a variable maximum rate for kinetics of some reactions.

Based on experimental information, we have designed 4 variable enzyme pools: glucose uptake, PAA export, penicillin production and carbohydrate storage/release process. All rate capacities followed the intracellular enzyme balance (Stephanopoulos et al. 1998):

$$\frac{dX_E}{dt} = q_{E,syn} - \left(k_{dE} + \mu\right) X_E \qquad \text{Eq.2}$$

Where $q_{E,syn}$ is the specific enzyme synthesis rate in molE/CmolX/h, k_{dE} is the enzyme degradation rate in 1/h and μ is the cell growth rate in 1/h. The amount of transporter/protein is a result of production, degradation and growth dilution. The enzyme synthesis rate $q_{E,syn}$ kinetics differ between the enzyme pools, according to different experimental profiles.

Variable glucose uptake capacity

Experimental data from steady state and feed-ramp experiments suggested a loss of glucose uptake capacity at low growth rate (Figure S1), in line with previous reports (Heijnen and Romein 1995). This reduction leads to a higher extracellular glucose concentration.

Previous experimental results indicate a stable glucose affinity in either steady or oscillation conditions (de Jonge et al. 2011), attributing the uptake capacity loss to the protein synthesis. Following the same hypothesis, we propose that the synthesis of glucose transporter protein is a sigmoid function of the growth rate with a base level sufficient to provide the maintenance energy required at μ=0 1/h:

$$q_{E,syn} = q_{E,syn,\max} \frac{\left[\left(\mu + \delta_0\right)/k_{1.1}\right]^5}{\left[\left(\mu + \delta_0\right)/k_{1.1}\right]^5 + 1} \qquad \text{Eq.3}$$

Variable PAA export capacity

PAA export is reportedly governed by an ATP-binding cassette transporter (Douma et al. 2012). The enzyme synthesis rate was assumed to follow a linear relationship to the growth rate (Condon et al. 1995):

$$q_{E,syn} = \alpha + \beta\mu \qquad \text{Eq.4}$$

Combining Eq.2 and Eq.4 yields:

$$\frac{dX_E}{dt} = \alpha + \beta\mu - k_{dE}X_E - \mu X_E \qquad \text{Eq.5}$$

Here 'α' was set to zero, assuming that little to no transporter protein would be produced at zero growth.

Variable penicillin production pathway capacity

The kinetics for penicillin production are directly applied from previous research by Douma et al. (2010c). However, instead of the (extracellular) residual glucose concentration, we consider the concentration of intracellular glycolytic intermediates as the repression factor for synthesis of the enzyme in the penicillin pathway.

Variable carbohydrates storage/release capacity

Intracellular stored carbohydrates data from both chemostat and feed-ramp showed an increase in storage capacity at low growth rate, indicating that the cell tends to save more carbon to counter possible starvation. The Sto pool balance reads:

$$\frac{dX_{sto}}{dt} = \nu_{4.1} - \nu_{4.2} - \mu X_{sto} \qquad \text{Eq.6}$$

In steady state, the Sto pool size is only a function of the net storage flux and growth rate:

$$X_{sto} = \frac{\nu_{4.1} - \nu_{4.2}}{\mu} \qquad \text{Eq.7}$$

From the steady state data, one can easily calculate that the net storage flux ($\nu_{4.1}$-$\nu_{4.2}$) is around 0.0045 molC/CmolX/h at μ=0.05 1/h. This value is only 4% of the q_S (0.12 molC/CmolX/h (Nasution et al. 2008)) and only 18% of the storage influx (0.024 molC/CmolX/h (de Jonge et al. 2013)) under that condition. Therefore, for a correct storage pool size, one should carefully control the fluxes of storage and release so that their difference drops slower than the decline in growth rate. We propose to model the carbohydrate storage/release at two

levels (Eq.S11-Eq.S13, Table S5): (1) A variable 'enzyme' capacity similar to glucose uptake, PAA export and penicillin synthesis (See above). (2) Kinetic relations for storage and release on ATP, storage pools and the extracellular glucose concentration to determine the instantaneous rates. Unfortunately, performing a direct flux validation for these storage/release processes was difficult due to the high cycle rate and low net flux. Further validation on these rates could be realized through ^{13}C labelling experiments (Antoniewicz 2013; de Jonge et al. 2013; Tang et al. 2009).

Results and Discussion

Parameter estimation and model performance

Of the 40 model parameters, the value of 9 parameters could be obtained from previous reports (Table S6), the remaining 31 parameters were fitting based on the experimental result and the fitting algorithm mentioned above. The estimated error in the parameter is in estimation is typically in the range of 10-25%.

Figure 2 shows an overview of the model performance for steady states and feed-ramp experiments. Herbert-Pirt parameters for substrate utilization were calculated based on the stoichiometry and compared to previous reports (Table II). The glucose/biomass ratio is the same but the glucose/penicillin ratio is significantly lower in the 9 pool model. However, this is compensated by an additional energy cost due to a futile PAA cycle. In the black box model, these two features are combined, yielding to a higher glucose/penicillin ratio. Except for this, the new model showed good stoichiometric agreement with the black box model.

Table II Calculated yields and maintenance rates between 9-pool model and black box model

Model	$Y_{X/S,max}$ (molglc/CmolX)	$Y_{P/S,max}$ (molglc/CmolPenG)	m_S (molglc/CmolX/h)	m_{PAA}[a] (molglc/molPAA/h)
9-pool Model	0.249	2.484	0.00134	0.00142
Black box model (Douma et al. 2010c)	0.250	5.750	0.00150	n.a.[b]

a: PAA futile cycle rate was calculated based on PAA export rate at D=0.05 1/h (St.St.) with C_{PAA}=0.003molPAA/kg.
b: not available.

Parameter Sensitivity

The parameter sensitivity was expressed as the percentage variation of the predicted pool sizes and reaction rates when a single parameter was submitted to a change (Cloutier et al. 2008). Here the model sensitivity for the 37 estimated parameters (pH_{ext}, pH_{int} and pK values from Deshmukh et al. (2015) were excluded) was investigated for a glucose-limited steady state at a dilution rate of 0.05 1/h, via a 20% increase on all individual parameters. A symmetric response to a decrease of 20% was obtained for all parameters (data not shown).

The sensitivity analysis in Figure 3 illustrates clear, but limited variations of the intracellular pool sizes (X_i) and/or rates (v_i) with respect to parameter changes. It can be observed from Figure 3 that $v_{1.1}$, $v_{1.2}$, $v_{1.3}$, $v_{2.1}$ and $v_{2.2}$ show little sensitivity towards all parameters changes, so do the the dynamic parameters in $v_{1.1}$. The insensitivity of these rates and parameters, related to the carbon metabolism and energy production, is explained by their stoichiometric coupling when the cell growth rate ($v_{1.3}$) is fixed at 0.05 1/h. As expected, the pools of glycolytic intermediates, amino acids and storage carbohydrates show corresponding responses to the parameter variation. This suggests a certain stability of the 9-pool model. X_{ATP}, which is supposed to be a stiff node in this model, was confirmed to be relatively invariant in the sensitivity analysis. The 9-pool model was able to predict stable glucose-limited continuous cultivations with dilution rates no lower than 0.005 1/h as well as glucose-excess conditions (batch).

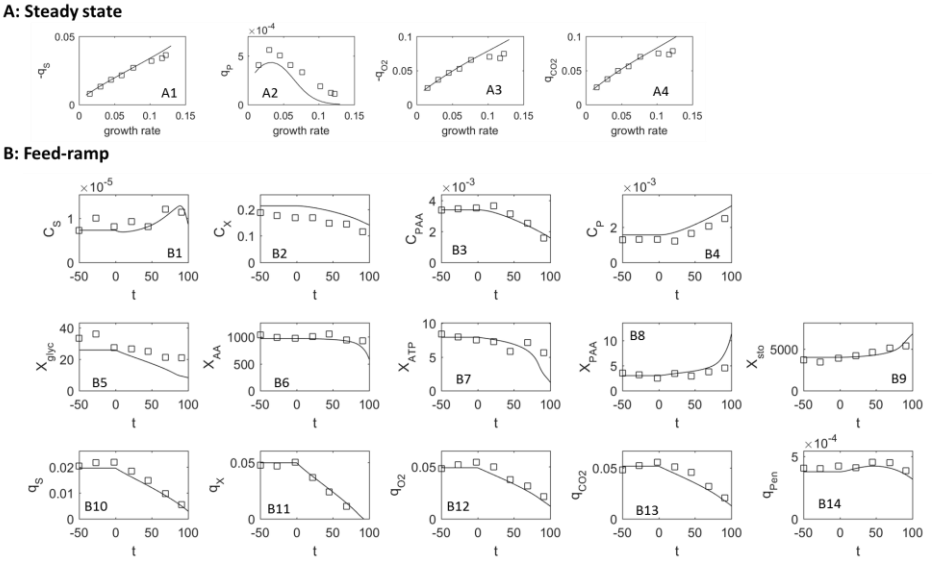

Figure 2 A: qi profiles as a function of the specific growth rate under chemostat conditions; B: extracellular Ci (moli/kg), intracellular Xi (µmol/gDW), specific rates qi (moli/CmolX/h) as a function of time during steady state (-50-0h) and ramp phases (0h-100h). Experimental data (□) and simulated result (line) predicted by the 9-pool model.

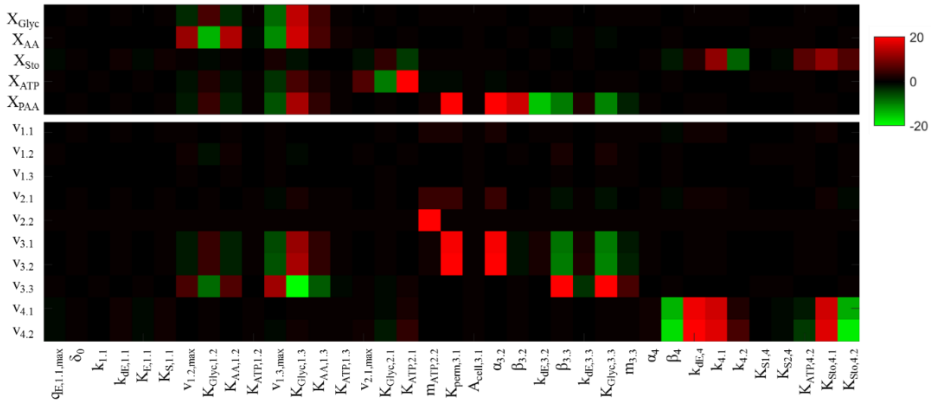

Figure 3 Sensitivity analysis results (+20%) on 31 model parameters for 5 intracellular pools and the 10 reaction rates at steady state (D=0.05 1/h). Relative errors in 5 intermediate pool sizes (upper panel) and 10 reaction rates (lower panel) are shown as percentage in the heat map (Red: +20%, Green: -20%).

Performance under feast-famine condition

Feast-famine cycles were selected for the 9-pool model validation. The application purpose of the model is within a CFD framework for simulation of an industrial scale bioreactor, where feast-famine regimes are expected. The average cycle time through the main regimes ranges from tens of seconds to several minutes depending on different bioreactor scales and configuration. The selected feast-famine experiment had a similar window of frequency and amplitude of the changes. The simulation started with a batch phase followed by a continuous phase under D=0.05 1/h. Once a steady state is reached, the feed strategy changed to an intermittent feed (de Jonge et al. 2011). In the simulation, the fluctuating feed regime lasted100 hours to reach reproducible cycles. In comparison, experimental stable cycle profile was obtained after at most 90 hours of intermittent feed (de Jonge et al. 2011).

Figure 4 shows the result of all intracellular pool profiles predicted by the 9-pool model compared to the available experimental data. It is clear that the two rapid pools, glycolytic intermediates and ATP, respond sufficiently quick (in several seconds) to the extracellular shift in glucose concentration. The Glyc pool accumulated a certain amount of carbon (around 40µmolC/gDW) during the feast phase, which was consumed almost immediately at the beginning of famine phase. Later in the famine phase, the carbon source was routed to the Sto pool. The model predicted that about 80µmolC/gDW carbon was released from the Sto pool so as to support the cell during starvation. These Glyc pool and Sto pool contribute a total 120µmolC/gDW of intracellular carbon supplement during the glucose depletion phase. This amount is in agreement with experimental results which suggested a transfer of intracellular carbon between feast and famine of about 113.7µmolC/gDW during the cycle (de Jonge et al. 2013).

The predicted glucose profile under feast famine conditions (Figure 5A) shows a good agreement with measured concentrations. De Jonge have reconstructed the q_S (de Jonge et al. 2011), growth rate(de Jonge 2016), q_P (de Jonge et al. 2011), respiratory intensity (de Jonge 2016; de Jonge et al. 2014) based on dynamic balance. The model output was compared to these estimated experimental rates (Figure 5B-G). The 9-pool model rates are close to experimental results in steady states (before t = 0) and showed some deviations during the feast famine cycles. Because of a long protein (glucose transporter) turnover time, the glucose uptake capacity in feast famine regime could be regarded at pseudo-steady state, at a value very close (0.0417 molglc/CmolX/h)

to reported $q_{S,max}$ (0.0449 molglc/CmolX/h). Consequently the uptake profile was in good agreement with the experimental rate calculated by de Jonge et al. (2011). For the growth rate and respiratory intensities, the model showed a more dynamic pattern than that in experimental estimates (Figure 5E, F). However, the estimated q_{O2}, q_{CO2} and μ are relatively uncertain due to the limitation of the used estimation method (de Jonge 2016). Most important is that, the 9-pool model managed to retain a certain growth rate and respiratory rates during the famine phase, due to the availability of the intracellular carbon source. It was reported that this intracellular carbon (e.g. trehalose, mannitol and glycogen, etc.) acted as emergency carbon supply during short stages of glucose depletion (de Jonge et al. 2013). By taking a closer look at the simulated storage rates (Figure 6), one notices that the storage process ($v_{4.1}$) was activated in the feast phase and inactivate during the famine phase, and vice versa for the carbohydrates release. This leads to a carbon flux at around 0.05 molC/CmolX/h towards the Glyc pool in the famine phase, which is comparable to the influx of extracellular glucose under at steady state conditions with D=0.016 1/h.

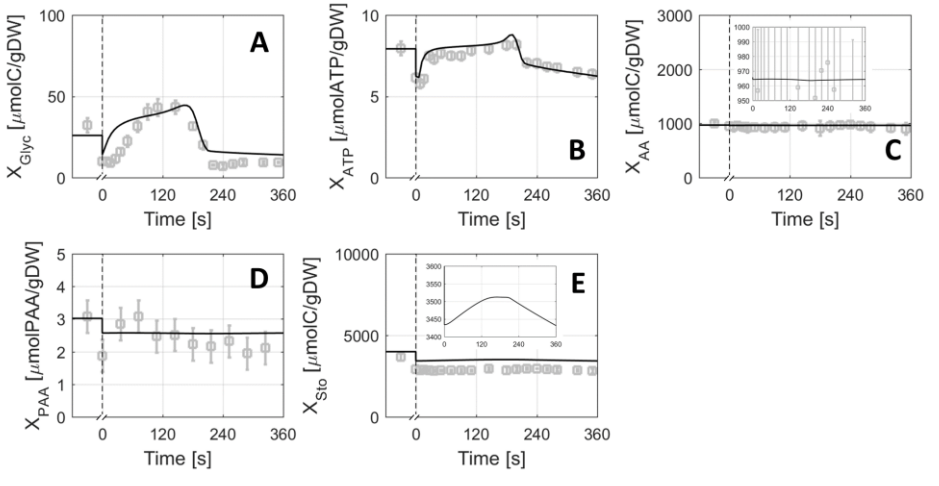

Figure 4. Comparison of intracellular profiles predicted by the 9-pool model (black solid line) and experimental results (grey open square) under steady state (before time 0) and feast-famine conditions (after time 0). Local scale magnification are shown for X_{AA} and X_{Sto}. Steady state: steady condition after 100 hours of continuous feed cultivation. Feast-famine: reproducible cycle after 100 hours of intermittent feed cultivation.

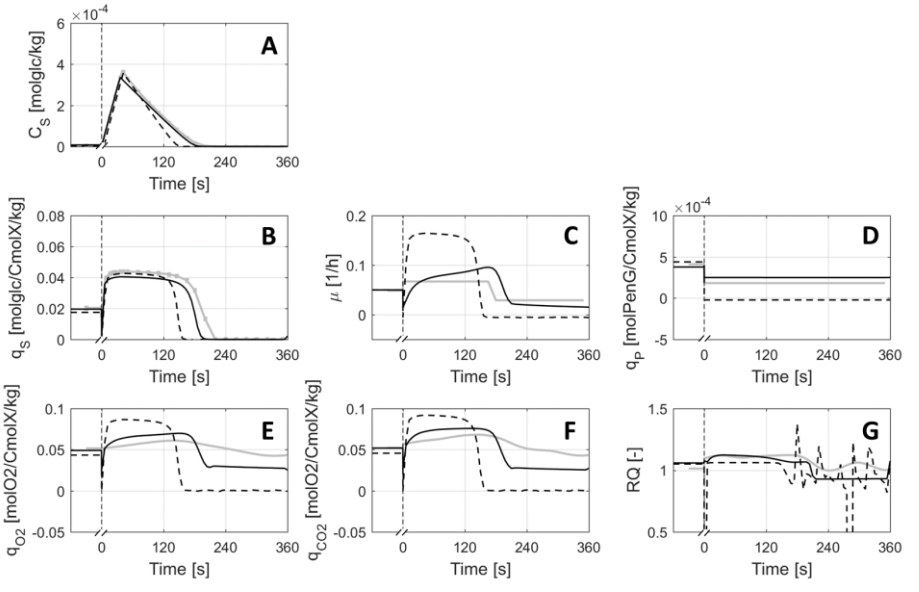

Figure 5 Comparison of extracellular and q-rate profiles predicted by 9-pool model (black solid Line), black box model (black dash line) and experimental results (grey solid line) under steady state (before time 0) and feast-famine condition (after time 0). Steady state: steady condition after 100 hours of continuous cultivation. Feast-famine: reproducible cycle after 100 hours of intermittent feed. Experimental q-rates were calculated/reconstructed based on de Jonge (2016)

By applying the intermittent feed strategy, it was observed that there was a significant drop in the penicillin production rate (de Jonge et al. 2011). The predictions of the 9-pool model were compared with experimental data obtained with continuous and intermittent deed strategies (**Figure 7**). The penicillin production ability in the 9-pool model is controlled by a limiting enzyme ($X_{E,3.3}$) whose synthesis is inhibited glycolytic intermediates. Under feast famine conditions, X_{Glyc} caries significantly (**Figure 4A**). This impacts the synthesis of this limiting enzyme under feast-famine conditions. Due to a long turnover time of the enzyme, the simulated q_P profile shows a steady (pseudo-steady) state, but at a lower value.

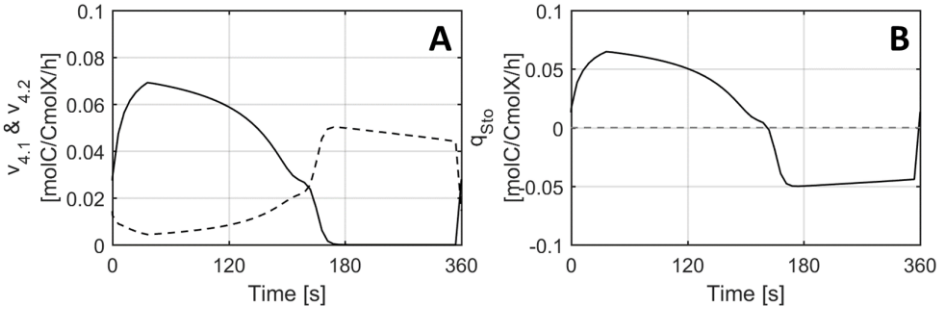

Figure 6 Predicted carbon storage and carbon release process in one reproducible feast-famine cycle. A: $v_{4.1}$ (storage): solid line and $v_{4.2}$ (release): dashed line. B: net storage rate (positive for storage, negative for release).

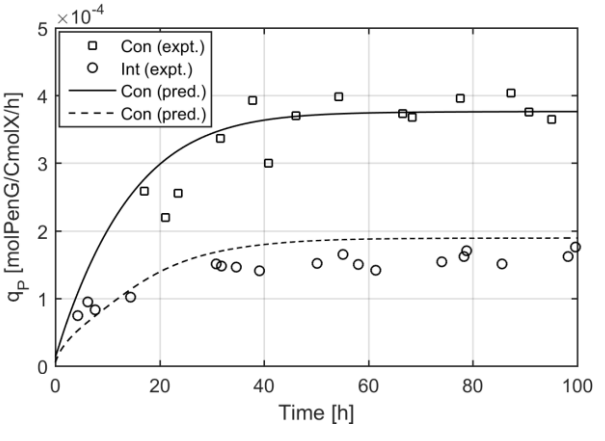

Figure 7. Predicted and experimental qP profile in the first 100 h of continuous feeding cultivation (circles and dashed line) and intermittent feeding condition (square and solid line). t=0 represents the end of the batch phase.

The simulation results showed the 9-pool model capacity in dealing with glucose oscillation conditions. Despite slight differences observed between simulation and reality, the model reflects the major metabolic characteristics of the cell dynamics, making it qualified for further integration.

Limitations of the Black box model

To compare the 9-pool model performance with a classic black box model, a published black box model of the applied strain (See Supplement B) was also tested. This black box model was previously designed on chemo-state results and has been tested under fed-batch conditions. Here, we compared its applicability to much more dynamic conditions, i.e. the feast-famine regime.

One deficiency of the black box model is the direct application of the Herbert-Pirt equation. In the black box model, the distribution of substrate uptake

always satisfies this equation (van Heerden et al. 2014) with no dynamics. This is a reasonable hypothesis in steady state (chemostat) or slow dynamic conditions (fed-batch or feed-ramp) where substrate is always available, but could be problematic under highly dynamic conditions where substrate is absent for short periods of time. In our feast-famine cases, the black box predicted an increase in biomass which was not observed in either the 9-pool model or experimental results. This 36% over-estimation in the biomass leads to a shorter feast phase and a longer famine phase (Figure 5).

In addition, there was no intracellular carbon source available in the black box model, yielding it incapable of handling feast-famine conditions. Obvious consequences of the absence of glucose were a negative growth (Figure 5C) and production rate (Figure 5D). These two rates have to be negative in the black box model when substrate is absent, so as to balance the constant m_S in Herbert-Pirt equation. Furthermore, respiration was predicted to be absent (Figure 5E, F). These issues of the black box model are structural deficiencies and cannot be overcome through parameter adjustment.

Conclusion

A metabolically structured 9-pool kinetic model was developed for penicillin fermentation. This model, characterized by five metabolite pools, four pathway capacity pools and ten kinetic reactions, showed sufficient accuracy in predicting intracellular pool sizes, extracellular concentrations, and reaction rates. By predicting feed-ramp and intermittent feed experiments, the kinetic model succeeded in reproducing dynamic processes both at minute and hour timescales, and the latter being representative for large scale fed-batch fermentations. Because of the lumped kinetics, a relatively limited computational power is needed and therefore this model can be further integrated with CFD simulations so as to realize a full, high-resolution industrial-scale bioprocess simulation. In addition, the pool structure allows the calculation of individual cell lifelines, which describe the performance and composition of cells (or clusters of cells) during their trajectories in large bioreactors. This enables to study of cell population dynamics under complex industrial conditions.

Further extensions can be made to this model. Depending on the model application, additional pools can be added. For instance, including oxygen and intracellular reducing agent pools (NAD^+-NADH pair), enables to study local variations in oxygen supply and the cell redox level distribution in large-scale fermentations, which are commonly regarded as a key feature in aerobic

fermentation. However, before further improving this model, the trade-off between model accuracy/application and computational complexity needs to be carefully considered for supporting scale-up and scale-down design and optimization.

References

Antoniewicz MR. 2013. ^{13}C metabolic flux analysis: optimal design of isotopic labeling experiments. Curr Opin Biotech 24(6):1116-1121.

Berg M. 2013. *Penicillium chrysogenum*: Genomics of an Antibiotics Producer. In: Horwitz BA, Mukherjee PK, Mukherjee M, Kubicek CP, editors. Genomics of Soil- and Plant-Associated Fungi: Springer Berlin Heidelberg. p 229-254.

Bylund F, Guillard F, Enfors SO, Trägårdh C, Larsson G. 1999. Scale down of recombinant protein production: a comparative study of scaling performance. Bioprocee Eng 20(5):377-389.

Cipollina C, ten Pierick A, Canelas AB, Seifar RM, van Maris AJA, van Dam JC, Heijnen JJ. 2009. A comprehensive method for the quantification of the non-oxidative pentose phosphate pathway intermediates in *Saccharomyces cerevisiae* by GC–IDMS. J Chromatogr B 877(27):3231-3236.

Cloutier M, Bouchard-Marchand É, Perrier M, Jolicoeur M. 2008. A predictive nutritional model for plant cells and hairy roots. Biotechnol Bioeng 99(1):189-200.

Condon C, Squires C, Squires CL. 1995. Control of rRNA transcription in *Escherichia coli*. Microbiological Reviews 59(4):623-45.

Craven S, Shirsat N, Whelan J, Glennon B. 2013. Process model comparison and transferability across bioreactor scales and modes of operation for a mammalian cell bioprocess. Biotechnol Progr 29(1):186-196.

de Jonge L, Buijs NAA, Heijnen JJ, van Gulik WM, Abate A, Wahl SA. 2013. Flux response of glycolysis and storage metabolism during rapid feast/famine conditions in *Penicillium chrysogenum* using dynamic ^{13}C labeling. Biotech J 9:372-385.

de Jonge LP. 2016. A scale-down study of the industrial penicillin fermentation using quantitative metabolomics [D]. The Netherlands: Delft University of Technology. 192 p.

de Jonge LP, Buijs NAA, ten Pierick A, Deshmukh A, Zhao Z, Kiel JAKW, Heijnen JJ, van Gulik WM. 2011. Scale-down of penicillin production in *Penicillium chrysogenum*. Biotech J 6(8):944-958.

de Jonge LP, Douma RD, Heijnen JJ, van Gulik WM. 2012. Optimization of cold methanol quenching for quantitative metabolomics of *Penicillium chrysogenum*. Metabolomics 8(4):727-735.

de Jonge LP, Heijnen JJ, van Gulik WM. 2014. Reconstruction of the oxygen uptake and carbon dioxide evolution rates of microbial cultures at near-neutral pH during highly dynamic conditions. Biochem Eng J 83(0):42-54.

Deshmukh AT, Verheijen PJT, Maleki Seifar R, Heijnen JJ, van Gulik WM. 2015. *In vivo* kinetic analysis of the penicillin biosynthesis pathway using PAA stimulus response experiments. Metab Eng 32:155-173.

Dolan KD, Yang L, Trampel CP. 2007. Nonlinear regression technique to estimate kinetic parameters and confidence intervals in unsteady-state conduction-heated foods. J Food Eng 80(2):581-593.

Douma RD, de Jonge LP, Jonker CT, Seifar RM, Heijnen JJ, van Gulik WM. 2010a. Intracellular metabolite determination in the presence of extracellular abundance: Application to the penicillin biosynthesis pathway in *Penicillium chrysogenum*. Biotechnol Bioeng 107(1):105-15.

Douma RD, de Jonge LP, Jonker CTH, Seifar RM, Heijnen JJ, van Gulik WM. 2010b. Intracellular Metabolite Determination in the Presence of Extracellular Abundance: Application to the Penicillin Biosynthesis Pathway in *Penicillium chrysogenum*. Biotechnol Bioeng 107(1):105-115.

Douma RD, Deshmukh AT, de Jonge LP, de Jong BW, Seifar RM, Heijnen JJ, van Gulik WM. 2012. Novel insights in transport mechanisms and kinetics of phenylacetic acid and penicillin-G in *Penicillium chrysogenum*. Biotechnol Prog 28(2):337-48.

Douma RD, Verheijen PJ, de Laat WT, Heijnen JJ, van Gulik WM. 2010c. Dynamic gene expression regulation model for growth and penicillin production in *Penicillium chrysogenum*. Biotechnol Bioeng 106(4):608-18.

Haringa C, Tang W, Deshmukh AT, Xia J, Reuss M, Heijnen JJ, Mudde RF, Noorman HJ. 2016. Euler-Lagrange computational fluid dynamics for (bio)reactor scale-down: an analysis of organism life-lines. Eng Life Sci 16(7):652-663..

Harris DM, Diderich JA, van der Krogt ZA, Luttik MAH, Raamsdonk LM, Bovenberg RAL, van Gulik WM, van Dijken JP, Pronk JT. 2006. Enzymic analysis of NADPH metabolism in β-lactam-producing *Penicillium chrysogenum*: Presence of a mitochondrial NADPH dehydrogenase. Metab Eng 8(2):91-101.

Heijnen JJ. 2005. Approximative kinetic formats used in metabolic network modeling. Biotechnol Bioeng 91(5):534-545.

Heijnen JJ, Romein B. 1995. Derivation of Kinetic Equations for Growth on Single Substrates Based on General Properties of a Simple Metabolic Network. Biotechnol Progr 11(6):712-716.

Hynne F, Danø S, Sørensen PG. 2001. Full-scale model of glycolysis in *Saccharomyces cerevisiae*. Biophys Chem 94(1–2):121-163.

Kerkhoven EJ, Lahtvee P-J, Nielsen J. 2015. Applications of computational modeling in metabolic engineering of yeast. FEMS Yeast Res 15(1):1-13.

Kleijn RJ, Liu F, van Winden WA, van Gulik WM, Ras C, Heijnen JJ. 2007. Cytosolic NADPH metabolism in penicillin-G producing and non-producing chemostat cultures of *Penicillium chrysogenum*. Metab Eng 9(1):112-23.

Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJT, Heijnen JJ. 2001. Improved rapid sampling for in vivo kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. Biotechnol Bioeng 75(4):406-415.

Lapin A, Klann M, Reuss M. 2010. Multi-scale spatio-temporal modeling: lifelines of microorganisms in bioreactors and tracking molecules in cells. Adv Biochem Eng Biotechnol 121:23-43.

Lapin A, Müller D, Reuss M. 2004. Dynamic Behavior of Microbial Populations in Stirred Bioreactors Simulated with Euler−Lagrange Methods: Traveling along the Lifelines of Single Cells. Ind Eng Chem Res 43(16):4647-4656.

Lapin A, Schmid J, Reuss M. 2006. Modeling the dynamics of E. coli populations in the three-dimensional turbulent field of a stirred-tank bioreactor—A structured–segregated approach. Chem Eng Sci 61(14):4783-4797.

Larsson G, Törnkvist M, Wernersson ES, Trägårdh C, Noorman H, Enfors SO. 1996. Substrate gradients in bioreactors: origin and consequences. Bioprocee Eng 14(6):281-289.

Lu H, Cao W, Ouyang L, Xia J, Huang M, Chu J, Zhuang Y, Zhang S, Noorman H. 2016. Comprehensive reconstruction and in silico analysis of *Aspergillus niger* genome-scale metabolic network model that accounts for 1210 ORFs. Biotechnol Bioeng: In press.

Mashego MR, van Gulik WM, Vinke JL, Heijnen JJ. 2003. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. Biotechnol Bioeng 83(4):395-399.

Nasution U, van Gulik WM, Kleijn RJ, van Winden WA, Proell A, Heijnen JJ. 2006. Measurement of intracellular metabolites of primary metabolism and adenine nucleotides in chemostat cultivated *Penicillium chrysogenum*. Biotechnol Bioeng 94(1):159-66.

Nasution U, van Gulik WM, Ras C, Proell A, Heijnen JJ. 2008. A metabolome study of the steady-state relation between central metabolism, amino acid biosynthesis and penicillin production in *Penicillium chrysogenum*. Metab Eng 10(1):10-23.

Nikerel IE, Verheijen PT, Gulik W, Heijnen J. 2012. Model-Based Design of Superior Cell Factory: An Illustrative Example of *Penicillium chrysogenum*. In: Wittmann C, Lee SY, editors. Systems Metabolic Engineering: Springer Netherlands. p 221-270.

Noorman H. 2011. An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. Biotech J 6(8):934-943.

Österlund T, Nookaew I, Nielsen J. 2012. Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. Biotechnol Adv 30(5):979-988.

Rizzi M, Baltes M, Theobald U, Reuss M. 1997. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. Biotechnol Bioeng 55(4):592-608.

Seifar RM, Deshmukh AT, Heijnen JJ, van Gulik WM. 2012. Determination of delta- L-alpha-aminoadipyl -L-cysteinyl-D-valine in cell extracts of *Penicillium chrysogenum* using ion pair-RP-UPLC-MS/MS. J Sep Sci 35(2):225-230.

Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E, Swainston N, Dada JO, Khan F and others. 2013. A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. FEBS Lett 587(17):2832-2841.

Stephanopoulos G, Aristidou AA, Nielsen J. 1998. Metabolic engineering: principles and methodologies: Academic press.

Tang YJ, Martin HG, Myers S, Rodriguez S, Baidoo EEK, Keasling JD. 2009. Advances in analysis of microbial metabolic fluxes via ^{13}C isotopic labeling. Mass Spectrom Rev 28(2):362-375.

Theobald U, Mailinger W, Baltes M, Rizzi M, Reuss M. 1997. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. Biotechnol Bioeng 55(2):305-316.

Tyo KEJ, Kocharin K, Nielsen J. 2010. Toward design-based engineering of industrial microbes. Curr Opin Microbiol 13(3):255-262.

van Eunen K, Kiewiet JAL, Westerhoff HV, Bakker BM. 2012. Testing biochemistry revisited: How *in vivo* metabolism can be understood from *in vitro* enzyme kinetics. PLoS Comput Biol 8(4):995-1007.

van Gulik WM, Antoniewicz MR, deLaat WTAM, Vinke JL, Heijnen JJ. 2001. Energetics of growth and penicillin production in a high-producing strain of *Penicillium chrysogenum*. Biotechnol Bioeng 72(2):185-193.

van Gulik WM, de Laat WTAM, Vinke JL, Heijnen JJ. 2000. Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. Biotechnol Bioeng 68(6):602-618.

van Heerden JH, Wortel MT, Bruggeman FJ, Heijnen JJ, Bollen YJM, Planqué R, Hulshof J, O'Toole TG, Wahl SA, Teusink B. 2014. Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. Science 343(6174).

Vasilakou E, Machado D, Theorell A, Rocha I, Nöh K, Oldiges M, Wahl SA. 2016. Current state and challenges for dynamic metabolic modeling. Curr Opin Microbiol 33:97-104.

Wang G, Tang W, Xia J, Chu J, Noorman H, van Gulik WM. 2015. Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. Eng Life Sci 15(1):20-29.

Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C-labeled cell extracts as internal standards. Anal Biochem 336(2):164-171.

Supplementary material

Additional supplementary information can be found in the online publication:

https://onlinelibrary.wiley.com/doi/10.1002/bit.26294

Supplement A. 9-pool model structure.

Supplement B. Black box model.

Supplement C. Compound balances used for both BB model and MS model.

Supplement D. Original experimental results.

Chapter 3

Computational fluid dynamics simulation of an industrial *P. chrysogenum* fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization

Published as

Haringa, C., **Tang, W.**, Wang, G., Deshmukh, A. T., van Winden, W. A., Chu, J., van Gulik, W. M., Heijnen, J. J., Mudde, R. F., & Noorman, H. J. (2017). Computational fluid dynamics simulation of an industrial P. chrysogenum fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization. Chemical Engineering Science.
https://doi.org/https://doi.org/10.1016/j.ces.2017.09.020

Abstract

We assess the effect of substrate heterogeneity on the metabolic response of *P.chrysogenum* in industrial bioreactors via coupling of a 9-pool metabolic model with Euler-Lagrange CFD simulations. In this work, we outline how this coupled hydrodynamic-metabolic modeling can be utilized in 5 steps. (1) A model response study with a fixed spatial extra-cellular glucose concentration gradient, which reveals a drop in penicillin production rate q_P of 18-50% for the simulated reactor, depending on model setup. (2) CFD-based scale-down design, where we design a 1-vessel scale down simulator based on the organism lifelines. (3) Scale-down verification, numerically comparing the model response in the proposed scale-down simulator with large-scale CFD response. (4) Reactor design optimization, reducing the drop in penicillin production by a change of feed location. (5) Long-term fed-batch simulation, where we verify model predictions against experimental data, and discuss population heterogeneity. Overall, these steps present a coupled hydrodynamic-metabolic approach towards bioreactor evaluation, scale-down and optimization

Keywords: CFD, Ruler-Lagrange, Metabolic model, Scale-down, Industrial

Introduction

Due to the presence of gradients in substrate concentration(Enfors et al., 2001), dissolved oxygen concentration (Oosterhuis & Kossen, 1984) and other process variables in industrial bioreactors, organisms are subject to temporal variations in their environment. Such variations impose stresses on these organisms (Lara et al., 2006; Neubauer & Junne, 2010; Wang et al., 2014), which may in turn affect the process yield (de Jonge et al., 2011). There are cases where extracellular variations appear to be advantageous (Enfors et al., 2001), but typically the impact is negative as the process is driven away from the conditions set for yield optimization (de Jonge et al., 2011; Wang et al., 2015). Being related to mixing behavior, these gradients may occur in any reactor type, and are expected to amplify upon scale-up, which may hence come with a yield loss that should be considered when judging scale-up economics. Furthermore, knowledge on the impact of bioreactor heterogeneity can be used to guide design changes to the reactor and, with genetic engineering, the micro-organism itself.

Previously, we used Euler-Lagrange computational fluid dynamics (CFD) to study the environmental fluctuations experienced by micro-organisms (called lifelines) (Haringa et al., 2016) and showed how fluctuation statistics can be acquired from such simulation to guide scale-down (SD) simulator design (Haringa, Deshmukh, et al., 2017). These works focused on simulation and fluctuation quantification using the substrate uptake (qS) lifeline and did not quantitatively consider the metabolic response. When a dynamic metabolic model is available for the studied organism, coupled metabolic-hydrodynamic simulations can be used to evaluate the expected metabolic impact (Lapin et al., 2004, 2006). Combined with experiments in representative scale-down simulators, such a coupled hydrodynamic-metabolic approach can be used for: (1) scale-down verification: does a scale-down simulator result in the same metabolic response as observed in the large-scale CFD simulation? And (2) design optimization: what is the expected impact of reactor design changes or metabolic modifications based on numerical assessment? The most promising changes can then be experimentally tested in representative scale-down simulations, offering a powerful approach to rational bioreactor design and scale-up (Wang et al., 2014, 2015).

We numerically study five topics, outlined in Figure 1, highlighting the different aspects of the CFD-based scale-down workflow. A penicillin production process is used as a cease-study. Part I considers the coupled hydrodynamic-metabolic

simulation of a 54m^3 industrial *P.chrysogenum* fermentation(Haringa et al., 2016), focusing on mixing dynamics and neglecting slow processes such a biomass growth. We study the impact of mixing on metabolic variations using a 9-pool metabolic model (Tang et al., 2017). Part II focuses on the design of a representative lab-scale SD-simulator for the 54m^3 reactor. In part III, we perform numerical verification of the proposed SD-simulator performance, first assuming ideal mixing, and second by a CFD simulation of a 3L reactor with dynamic feed. In part IV, we discuss process optimization and propose a simple reactor alteration to improve the penicillin yield. To conclude, in part V, we simulate 60 h of a fed-batch fermentation for comparison with industrial data. With this we explore various aspects of the use of coupled hydrodynamic-metabolic modeling for process evaluation and optimization.

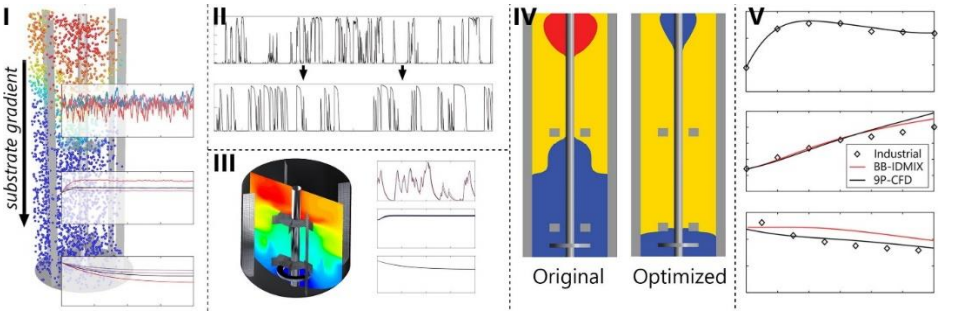

Figure 1. Graphic outline of the 5 subjects covered in this paper. I: Metabolic response simulation. II: CFD-guided scale-down simulator design. III: Numerical verification of scale-down reactor performance. IV: Numerical full-scale design optimization. V: industrial fed-batch simulation

Methodology

All CFD simulations were conducted in ANSYS FLUENT 15.7, MATLAB 8.6.0 was used for post-processing and ideal mixing simulations.

Metabolic model

The 9-pool metabolic model for P.chrysogenum developed by (Tang et al., 2017) contains 5 intra-cellular metabolite and 4 enzymatic pools, and couples to the extra-cellular substrate concentration C_S and phenylacetic acid (PAA) concentration C_{PAA}. The metabolite pools are: Glycolytic intermediates (X_{gly}), Amino acids (X_{AA}), Storage polymers (X_{sto}), ATP (X_{ATP}) and intra-cellular PAA (X_{PAA}), all reported in µmol/g_{dw} with g_{dw} being the dry biomass weight. Three dimensionless enzyme pools influence metabolic rates: $X_{E,11}$ (the substrate uptake capacity), $X_{E,32}$ (PAA export capacity), and $X_{E,4}$ (storage capacity). The 4th enzyme pool controls the biomass specific penicillin production rate q_P(Douma et al., 2010) and is reported in mol_P/$Cmol_X$/h (de Jonge et al., 2011; Douma et

al., 2010; van Gulik et al., 2000). For brevity, the mathematical model formulation is provided in Supplementary material A, together with additional information regarding FLUENT coupling. The effect of oxygen limitation has not been studied sufficiently to be included currently model (Haringa et al., 2016; Tang et al., 2017). Hence we currently assume sufficient oxygen supply in all cases; oxygen limitations will be considered in future extensions. For a model overview, we refer to (Tang et al., 2017).

Model simplifications. (Tang et al., 2017) developed and validated their model against a range of experimental data (de Jonge et al., 2011; van Gulik et al., 2000) including 360s feast-famine cycles (de Jonge et al., 2011). These results provide confidence that the model is able to capture the impact of circulation -timescale substrate variations. However, instabilities in X_{ATP} were encountered in our CFD simulations, which resulted from the sensitivity of the storage pool fluxes to turbulence-induced C_S fluctuations on the sub-second timescale, which were not accounted for in model development (for details see Supplementary material A). A structural solution of this issue requires deeper analysis of the signaling mechanism behind storage dynamics. As we currently lack the information to develop such improvements, we instead opted for a patch solution by assuming the ATP pool is in quasi-steady state, meaning the fluxes in- and out of the ATP pool balance, giving $dX_{ATP}/dt \approx 0$ (Nikerel et al., 2012). This converts the dynamic ATP-balance in an algebraic expression:

$$0 = \sum^{i} \left(v_i \left(X_{gly}, X_{AA}, X_{ATP}, \dots \right) \right) \qquad \text{Eq. 1}$$

For the current non-linear kinetics, Eq. 1 was evaluated for 100,000 randomly generated sets of intra-cellular pools. Subsequent correlation showed X_{ATP} can be modeled as $X_{ATP} = 8.25 \cdot X_{glyc}^3 / \left(10.5^3 + X_{glyc}^3 \right)$. The model response was deemed satisfactory under all tested conditions. Further details on the approach and verification against experimental data are reported in Supplementary material A.

CFD setup

54m³ reactor setup

We use the 54m³ reactor simulation (Haringa et al., 2016) with simplified single-phase hydrodynamics as the industrial base-case. We furthermore simulate the same case including aeration, with a superficial gas velocity of U_g = 0.05 m/s, measured under STP conditions. The headspace pressure in the reactor was

1.85 bar, which gives an air density of 2.4 kg/m^3 based on the log-mean pressure; the gas flowrate in the vessel was adjusted accordingly. The total domain height H_t = 11 m to account for broth expansion upon gassing, the gas-filled headspace is removed during parcel tracking (Haringa, Deshmukh, et al., 2017). A discrete population balance (8 bins, 0.5-12.7mm) with the kernels of (Luo & Svendsen, 1996) was employed to capture the bubble size distribution. A Sauter mean diameter d_b = 7mm was observed; we lack experimental data to verify this, unfortunately. Furthermore we used the standard k-ε model (dispersed turbulent formulation), multiple-reference frame impeller modeling, and the universal drag model for interphase momentum exchange. Other inter-phase forces were neglected (Gunyol et al., 2009; Haringa, Deshmukh, et al., 2017; Khopkar et al., 2003). Simulations using Casson rheology (Roels et al., 1974) diverged in volume fraction α. For simplicity, we hence set the broth rheology equal to water (Newtonian, μ_l = 0.001). We realize this is a strong deviation from reality; we defend this assumption by observing that measured air-broth circulation time lies in between the circulation times in pure water and air-water (Table 1), and capturing the range suffices for the current purpose. The air-water surface tension σ = 0.072 N/m, the turbulent Schmidt number was set to Sc_t = 0.2. Both single-phase and aerated simulations were conducted in a mesh with 180° periodicity. 235,000 hexahedral grid cells were used for single-phase cases, 923,000 hexahedral cells for aerated cases.

The gas-flow number $F_l = Q_g/ND^3$ = 0.1 implies the fermentor operates at the boundary of the 3-3 cavity regime and recirculation regime, where the mixing time τ_{95} is equal to or above that single phase-flow, respectively(Van't Riet & der Lans, 2011). Available industrial data on circulation time (Haringa et al., 2016) ($\tau_{circ} \approx \tau_{95}/4$, (Noorman, 2011)) suggests the later; the circulation time τ_{circ} is compared to simulation results in Table 1. The single-phase and two-phase simulation under- and over-estimate τ_{circ} for aerated broth with 30%, respectively. Note the experimental value is based on a single measurement and hence comes with a significant margin of error; furthermore, transient effect may lead to a natural variability in recorded mixing times (McClure et al., 2015), introducing additional uncertainty. With the present industrial data, it is unfortunately not possible to quantify this uncertainty. We regard the single-phase and aerated simulation as a lower and upper bound mixing time scenario, with the true mixing behavior in the range. This level of accuracy suffices for our current demonstration purposes, but we stress the need for further investigation into modeling true aerated, non-Newtonian fermentation broths, and associated with that, a wider range of large-scale validation data (gas hold-

up, local mixing curves and preferably local DO/substrate concentrations). Table 1 shows the gas-holdup is over-estimated compared to both air-water and air-broth experiments. This is likely an effect of model approximations, such as omitting inter-phase forces (except drag) and the empirical nature of inter-phase models/population balance kernels. For broth, the simplified rheology and effect of surfactants and anti-foam on the broth-water surface tension σ play an additional role. Currently we are not directly interested in gas-holdup, but in case oxygen dynamics are included, this aspect requires further study.

Table 1 Validation parameters of 54m^3 fermentor simulations. The holdup for air-water is averaged over 3 experiments, while a single experiment is reported for broth. The CFD circulation time is based on 5% saturation of the probe signal (mixing time τ_{95} is based on 95% saturation). For the experimental data, the half-circulation time was determined by recording the time lag between H_2SO_4 insertion at the top and probe response at the bottom.

Parameter	Exp. water	Exp. broth	Exp. air-water	Exp. air-broth	CFD water	CFD air-water
Gas hold-up	n/a	n/a	16.4±0.8	12.6	n/a	20.4
d_b [mm]	n/a	n/a	n/m	n/m	n/a	7.1
τ_{circ} [s]	19.3	77.0	41.6	25.7	18.2	32.9

n/a = not applicable; n/m = not measured

3-L laboratory reactor steup

A round-bottom vessel with a working volume of 3L (Tang et al., 2017) is simulated for scale-down verification (452,000 hexahedral grid cells. Geometric parameters are reported in Supplementary material B. The gas flowrate applied in prior scale-down experiment is 2L/min (0.66 VVM) (Tang et al., 2017; Wang et al., 2018), giving Fl = 0.009 with an agitation rate of N = 10 s^{-1} (600 RPM) (Tang et al., 2017). This value is outside of the range probed in mixing experiments (Van't Riet & der Lans, 2011), but implies τ_{95} is similar to or slightly higher than for effect of gas flow and model single-phase water. All walls were no-slip while the top surface had a no-shear free surface condition. Computational mixing simulations at 600 RPM yield a dimensionless mixing time N τ_{95} = 22, in excellent agreement with experiments (Supplementary material B); the dimensionless circulation time $\tau_{circ} \approx \tau_{95}/4$ (Noorman, 2011).

At high C_X, the high effective liquid viscosity μ_l may practically lead to transitional flow, possibly increasing θ_{95} significantly. Previous non-Newtonian simulation of aerated lab-scale reactors did not produce realistic mixing results due to stagnant zones (Moilanen et al., 2007), and preliminary work using a low-Re k-ε model with μ_l=0.15 Pa.s led to parcel tracking issues, with parcels sticking in the impingement point of the impeller discharge stream. We hence opted to decrease the agitation rate N to 1.67 s^{-1} to assess the effect of mixing time on the performance of a lab-scale scale-down simulator, and again assume a

Newtonian fluid with μ_l=0.001. This approach suffices for our current interest in the qualitative effect of a significant change in τ_{95}; we do stress that for predictive quantitative modeling a more realistic rheology model is required. For 600 and 100 RPM, τ_{circ} = 0.55 and 3.3s, respectively. Experimental evaluation of mixing behavior in real fermentation broths is required to comment on whether this range of τ_{circ} represents lab-scale practice.

Metabolic model coupling

The 9-pool metabolic model (Tang et al., 2017) is coupled to the Lagrangian (parcel) phase to study the response of microorganisms to environmental variations (Haringa, Noorman, et al., 2017; Lapin et al., 2004, 2006). In the 9-pool model the glucose uptake rate q_S is subject to transporter control, where the availability of transporter ($X_{E,11}$) is controlled by growth rate μ (h^{-1}). This means that strictly speaking 2-way coupling is required to resolve the substrate environment, which requires simulating long timespans due to the long transporter adaptation time, and is therefore computationally expensive (see Supplementary material C).

The long adaption time allows for the assumption that the average transport capacity $\overline{X_{E,11}}$ is homogeneous in the fermentor. As $\overline{X_{E,11}} = f(\mu)$, its value can be estimated based on growth rate under ideally mixing conditions, μ_{id}. For the applied model, the average growth rate under dynamic conditions $\bar{\mu}$ was typically close to μ_{id}, and the estimated $\overline{X_{E,11}}$ was similarly close. A-priori estimation of $\overline{X_{E,11}}$ allows to use 1-way coupling, as was done in earlier work(Haringa, Deshmukh, et al., 2017; Haringa et al., 2016; McClure et al., 2016), which means the number of tracked parcels N_p does not influence the substrate gradient and cab be freely chosen. This simplification does not hold when intra-cellular dynamics affect q_S at short timescales ($\approx\tau_{circ}$) (Haringa, Noorman, et al., 2017; Lapin et al., 2004, 2006), or when $\overline{X_{E,11}}$ under dynamic conditions differs strongly from the ideal-mixing assessment.

The above 1-way coupled approach was used to study mixing timescale dynamics, assuming constant C_X, $\overline{X_{E,11}}$, feed rate F and liquid-filled height H. This practically represents a chemostat cultivation, where the dilution rate D_r is equal to the mean growth rate $\bar{\mu}$. Parcel tracking for both 1- and 2-way coupling is conducted in FLUENT, but segregating the extra-cellular and intra-cellular reactions allows 1-way coupling to be executed after rather than during the FLUENT simulation, using MATLAB to perform the metabolic computations. Additional information regarding the practical implementation of the metabolic computations is provided in Supplementary material A. The (statistical) steady

state allows to simulate O(10) mixing times to acquire fluctuation statistics; lifelines of 80h are subsequently generated to study the adaptation of q_P to mixing-time dynamics (with constant C_X, $\overline{X_{E,11}}$ by construction) by joining together individual lifelines, exploiting the statistically-steady extra-cellular nature.

For the fed-batch simulation we use 2-way coupling to include temporal changes in C_X and $X_{E,11}$, meaning the metabolic computations are conducted in FLUENT as part of the simulation. The long variation time of both parameters allows the assumption that C_X and $\overline{X_{E,11}}$ are spatially homogeneous ($\overline{X_{E,11}}$ may be heterogeneous within the population, but to a same degree at every spatial location). This means that each time step$\overline{C_X}$ and $\overline{X_{E,11}}$ can be calculated as the parcel population ensemble average, and the local uptake rate can be computed from the Eulerian framework as

$$r_{S,C} = \overline{C_X} \cdot k_{11} \cdot \overline{X_{E,11}} \cdot \frac{C_S}{K_S + C_S} \qquad \text{Eq. 2}$$

This simplified 2-way coupling requires the parcel number to be sufficient to capture overall heterogeneity, for which N_P=O(10^3) typically suffices (Haringa et al., 2016; McClure et al., 2016); full 2-way coupling would require N_P=O(10^5)-O(10^6) (Haringa, Noorman, et al., 2017). 1- and 2-way coupling require similar computation time per hour flow-time, but 2-way coupling does require the full fermentation time to be simulated to account for changes in C_X and $X_{E,11}$. A comparison of assumptions between 1 and 2-way coupling is given in Table 2.

Table 2 Comparison of the assumptions between 1-way and 2-way coupling method used in this work. 1-way coupling is here used for chemostat cultivation, and 2-way coupling for fed-batch cultivation.

Method	1-way coupling	2-way coupling
Parcel tracking	FLUENT 15.7	FLUENT 15.7
Metabolic computation	MATLAB 8.6.0 (post-process)	FLUENT 15.7 (in-process)
C_X	Fixed	Variable
$X_{E,11}$	Fixed	Variable
μ	Variable, stat. steady	Variable
D_r	stat. steady $D_r = \bar{\mu}$	Dr ≈ 0 (fed-batch)

Overview of cases

We provide an overview of all simulations (Table 3), both conducted with CFD (FLUENT) and with the ideal or instantaneous mixing assumption (MATLAB), including the made assumptions and sections where these simulations are conducted. There is some variability in the applied timestep size Δt in FLUENT; in all cases it was ensured the particle trajectories were completed within the default accuracy settings. In all cases, glucose concentration C_S was variable, and the PAA concentration was fixed at C_{PAA} = 3mmol/kg.

As noted in previous section, the uptake capacity $q_{S,max} = k_{11} \cdot X_{E,11}$ in the 9-pool model depends on the growth rate μ. In the chemostat simulations, we aimed at μ≈0.03 h^{-1}. To maximize q_P: at this value of μ, the 9-pool model predicts $q_{S,max} \approx 1.13$ mmol/g_{dw}/h under well mixed conditions, which is marked lower than the $q_{S,max}$ = 1.6 mmol/g_{dw}/h reported by (de Jonge et al., 2011), measured for μ = 0.05 h^{-1}. The large scale simulation TU-A, TG-A, MU-A were conducted with $q_{S,max} = k_{11} \cdot X_{E,11} = 1.13\ mmol/g_{dw}/h$ and K_S = 9.8 μmol/kg.

The scale-down analysis and associated lab-scale CFD simulation (*part II* and *part III*) were conducted before the 9-pool model was available, which meant we had to rely on the kinetic parameters of (de Jonge et al., 2011), as in our previous work where we solely considered glucose uptake (Haringa et al., 2016). For consistency, we hence report a set of CFD simulations (TU-B, TG-B, MU-B) which use the 9-pool model, but with the uptake kinetics as published by (de Jonge et al., 2011), K_S = 7.8 μmol/kg and $q_{S,max}$ = 1.6 mmol/g_{dw}/h. We note that the fluctuations in q_S and the intra-cellular pools are too strong in these cases. The purpose of these simulations is to show that the intra-cellular response predicted between the industrial and lab-scale simulations matches; not to predict the metabolic response in the absolute sense.

Part I: Model response study. Part I focuses on TU-A (1-phase hydrodynamics, top feed) and TG-A (2-phase hydrodyn., top feed), to study the metabolic response to extra-cellular variations in an industrial-scale reactor with a statistically steady extra-cellular environment. As in our earlier work, a late fermentation stage was modeled, with C_X = 55 g/kg and substrate feed rate F = 1.23 g/m^3s (Haringa et al., 2016). The 1-way coupling approach means $X_{E,11}$ remains unchanged in time. All other pools were variable, and initialized based on ideal mixing results. For consistency with part II, III, TU-B and TG-B are also reported here. The results are compared with a CFD simulation coupled with the dynamic gene regulation model of (Douma et al., 2010) (1-phase, top feed, case TU-1P), and ideal-mixing simulations with both the dynamic gene regulation (ID-1P) and 9-pool (ID-9P) model.

Part II: Scale-down design. In part II we show a representative single-vessel SD-simulator with dynamic feed can be designed from the lifelines gathered in part I, using CFD-case TU-1B as a basis. Two designs are proposed, with biomass concentrations C_X = 55g/kg and C_X = 27.5 g/kg, respectively. As noted above, the uptake kinetics of (de Jonge et al., 2011) were used. AS in previous work (Haringa, Deshmukh, et al., 2017), the default SD protocol is based on matching q_S-lifelines between the scales.

Part III: Scale-down verification. First the performance of the scale-down protocols from *Part II* is assessed assuming ideal mixing (case ID-SD-27 and ID-SD-55). Next, CFD simulations of the 3L lab scale reactor were conducted with the C_X = 27.5 g.kg scale-down protocol, to study the effect of non-ideal mixing on SD performance. Instantaneous feed pulse injection was assumed in a small volume near the top surface. The hydrodynamics were frozen, but the substrate field was updated every timestep. The feed pulse scheme was supplied to FLUENT via a user defined function coupled to a lookup table. The fast mixing required time resolutions of Δt = 0.002 s for N = 600 RPM (case CFD-SD-600), Δt = 0.01 s for N = 100 RPM (case CFD-SD-100); this limited the resolved flow-time to 650s, in which 42 feed pulses were applied. This number is too small for a proper replication of the industrial-scale fluctuation statistics; therefore, scale-down performance was judged by comparing the model performance with the ideal-mixing response for the same 42 pulses.

Part IV: Design optimization. Industrial-scale CFD simulations were conducted with the substrate feed directly in the top impeller discharge stream (1-phase hydrodynamics), referred to as MU-A and MU-B.

Part V: Full-scale fed-batch verification. We simulated a 60 h timespan of a feed-batch fermentation (top feed, 1-phase hydrodyn.) which was conducted in the current 54m^3 geometry, named TU-FB, to verify model performance with industrial data which was kindly provided by the DSM biotechnology center. The simulation was started at t = 10 h after batch start C_X = 14 g/L. All model parameters are initialized based on the ideally-mixed 9-pool model outcome for the given starting conditions. In the industrial fermentation the total broth mass increased from 36,000 to 46,000 kg over the simulated timespan. However, explicitly modeling the volume change is computationally costly. As an approximation, kept the volume constant at 54 m^3, with the hydrodynamics of MU-1; as both impellers are submerged at all times, the change in τ_{circ} over the course of the fermentation is assumed to be minor. To compensate for the higher volume, the provided feed profile (reported in Figure 6) was adjusted to ensure an equal feed in g/kg/s between the simulation and industrial fermentation at all times. Experimental data for q_P and μ were used to evaluate model performance for TU-FB, as well as an ideal-mixed simulation with the model of (Douma et al., 2010), case ID-1P-FB.

Table 3 Overview of all the simulations, both CFD and ideal/instantaneous-mixing based (IDM), conducted in this work. All cases were conducted as chemostats, except for TU/ID-FB, which are a fed-batch simulations. Naming convention: T = top feed. M = mid feed (impeller discharge stream). U = ungassed. FB = fed-batch (2-way coupied). ID = instantaneously mixed. SD = scale-down. 9-P indicates the 9-pool model of (Tang et al., 2017) is used for metabolic coupling, 1-P indicates the Dynamic Gene Regulation model of (Douma et al., 2010) is used. A and B indicate which kinetic parameter values are used. SD-100 and SD-600 indicate agitation rates of 100 and 600 RPM, respectively

Name	CFD/IDM	Met. Model.	Part	Gassing	Coupling	Feed	$q_{S,max}$ [mmol/g_{dw}/h]	K_S [µmol/kg]	C_X [gdw/kg]	N_P	Δt [s]
TU-A	CFD	9-P	I	no	1-way	top	1.13	9.8	55	4000	0.1
TG-A	CFD	9-P	I	yes	1-way	top	1.13	9.8	55	4000	0.2
MU-A	CFD	9-P	IV	no	1-way	imp.	1.13	9.8	55	4000	0.3
TU-B	CFD	9-P	I	no	1-way	top	1.6	7.8	55	4000	0.03
TG-B	CFD	9-P	I	yes	1-way	top	1.6	7.8	55	4000	0.2
MU_B	CFD	9-P	IV	no	1-way	imp.	1.6	7.8	55	4000	0.3
TU-1P	CFD	1-P	I	no	1-way	top	1.6	7.8	55	n/a	st.st.
ID-1P	IDM	1-P	I	n/a	n/a	n/a	1.6	7.8	55	n/a	st.st.
ID-9P	IDM	9-P	I	n/a	n/a	n/a	1.6	7.8	55	n/a	st.st.
ID-SD-27	IDM	9-P	II/III	n/a	n/a	n/a	1.6	7.8	27	n/a	0.03
ID-SD-55	IDM	9-P	II/III	n/a	n/a	n/a	1.6	7.8	55	n/a	0.03
CFD-SD-100	CFD	9-P	III	no	1-way	top	1.6	7.8	27	5000	0.01
CFD-SD-600	CFD	9-P	III	no	1-way	top	1.6	7.8	27	5000	0.002
ID-1P-FB	IDM	9-P	V	n/a	n/a	n/a	var.	9.8	var.	n/a	1
TU-FB	CFD	9-P	V	no	2-way	top	var.	9.8	var.	2000	0.2

n/a = not applicable; var. variable; st. st. = simulation conducted in steady state.

Result and discussion

Part I: model response study

CFD simulations

We study the long-term adaptation of *P.chrysogenum* exposed to a strong substrate gradient. The most notable difference between TU-A/B and TG-A/B is the higher τ_{circ} for the latter, as discussed previously, yielding q_S fluctuations of longer duration. As q_S is locally saturated in all cases, the fluctuation amplitude hardly differs. Examples of single lifelines for TU-A and TG-A are shown in Figure 2, *top panel*.

Figure 2 shows the pool dynamics over an 80 h period for TU-A, TG-A, TU-B, TG-B. All cases show qualitatively similar behavior, but the higher $X_{E,11}$ for TU/TG-B has a clear negative impact on q_P. This illustrates the error introduced by taking kinetic parameters directly from literature, without accounting for the adaptation of $q_{S,max}$ to μ.

Practically, q_P is controlled by X_{gly}: high X_{gly} inhibits synthesis of penicillin producing enzyme, but it increases growth rate μ which enhances enzyme synthesis. The first effect scales with X_{gly}^6 (Tang et al., 2017), meaning that high values of X_{gly} are highly repressive, but below-average values of X_{gly} are hardly influential. This explains the large difference in q_P between the cases, eventhough all cases have a nearly equal average $\overline{X_{glyc}}$. The cases with the highest X_{gly} buildup show the biggest q_P loss. For aerated cases, the higher τ_{circ} translates to prolonged exposures to excess conditions, resulting in strong X_{gly} accumulation. Similarly, the higher transport capacity for TU/TG-B cases increased glycolytic accumulation. The effect of both kinetics and τ_{circ} is summarized in the Damköhler number *Da* = τ_{circ}/τ_{rxn}, where we take $\tau_{rxn} = K_S/(q_{S,max} \cdot C_X)$, the limit for $C_S \to 0$. This definition for τ_{rxn} does not require specifying a value of C_S, which makes it straightforward to evaluate for both experimental and CFD cases. Including the impeller-fed cases MU-A/B (*part IV*), a linear trend between the penicillin yield Y_{SP} (Table 4) and *Da* is observed: $Y_{SP} = 0.3417-0.0015Da$ ($R^2 = 0.97$), graphically shown in Supplementary material D.

Within the range of fluctuations, the effect of X_{gly} on μ, while non-linear in nature (Tang et al., 2017), can be reasonable linearized. Hence, the effect of high and low X_{gly} values on μ nearly averages out: $\overline{\mu(X_{gly})} \approx \overline{\mu(X_{gly})}$. Only the most extreme case (TG-B) deviates from this; the very lengthy exposures to starvation conditions leads to a lower $\bar{\mu}$. The data clearly shows that the

duration of exposures to excess- and starvation conditions strongly impacts the metabolic response. Since these time periods are highly distributed, there is considerable heterogeneity in X_{gly} at any given location. This feature is clearly visible in supplementary videos (available online with the original publication), and is inherently not captured by black-box models that instantaneous adaptions of the intra-cellular to the extra-cellular domain.

$X_{E,32}$ and $X_{E,4}$ are hardly affected in case TU/TG-B where $X_{E,11}$ was preconditioned for $\mu = 0.03\ h^{-1}$, whereas the higher uptake for TU/TG-B causes some changes in these pools. X_{AA} is hardly affected in all cases. The value of these pools is homogeneous within the population (Supplementary material E).

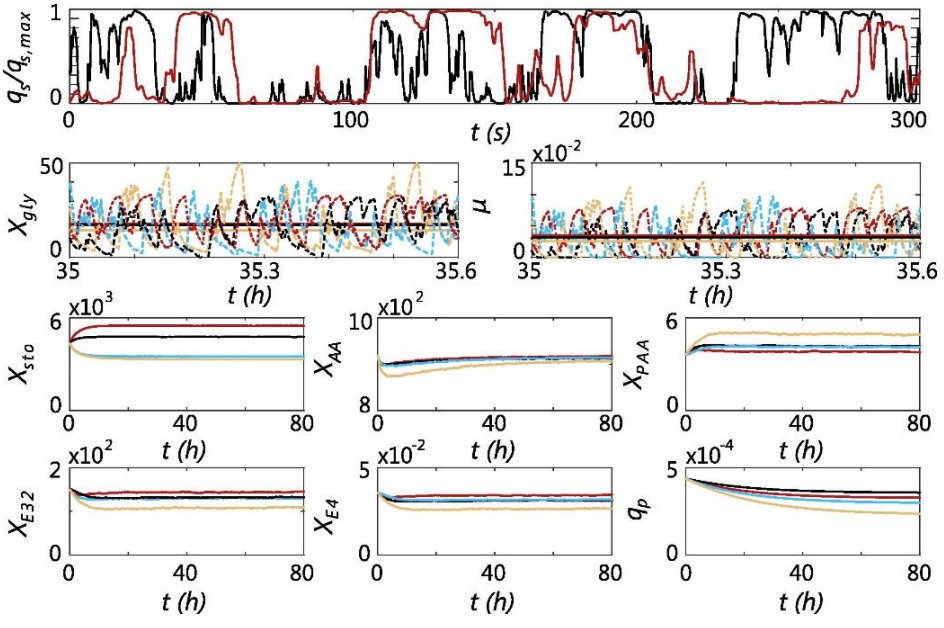

Figure 2 Long-term 9-pool model response for TU-A (black) , TG-A (red), TG-B (orange). The top panel shows examples of extra-cellular variations experienced by a single parcel in TU-A, TG-A. For X_{gly} and μ, the solid lines represent the mean, the dashed lines represent single parcel tracks to indicate the variations. All other lines represent are averages of 100 parcels. Intra-cellular pools X_{gly}, X_{AA}, X_{sto} and X_{PAA} have units $\mu mol/g_{dw}$. μ has units h^{-1}. q_P has units $mol/Cmol_X/h$. All other pools are dimensionless.

Table 4 Comparing yields and productivity between experimental data (van Gulik et al., 2000), the black box (BB) model (Douma et al., 2010) and the 9-pool model (Tang et al., 2017) with ideal mixing assumption and the 9-pool (9P) of (Tang et al., 2017)

Case	F [$mol_S/Cmol_X/h$]	μ [h^{-1}]	q_P [$mol_P/Cmol_X/h$]	Y_{SX} [$Cmol_X/Cmol_S$]	Y_{SP} [mol_P/mol_S]	Da
Exp.	0.0118	0.029	$5.33 \cdot 10^{-4}$	0.41	0.045	n/a
ID-1P	0.0125	0.032	$4.94 \cdot 10^{-4}$	0.43	0.040	0
ID-9P	0.0125	0.033	$4.40 \cdot 10^{-4}$	0.44	0.035	0
TU-1P	0.0125	0.043	$0.73 \cdot 10^{-4}$	0.57	0.006	n/a
TU-A	0.0125	0.032	$3.57 \cdot 10^{-4}$	0.44	0.029	32.1
TG-A	0.0125	0.035	$3.28 \cdot 10^{-4}$	0.47	0.026	58.0
TU-B	0.0125	0.033	$2.99 \cdot 10^{-4}$	0.44	0.024	57.0
TG-B	0.0125	0.027	$2.38 \cdot 10^{-4}$	0.36	0.019	103

n/a = not applicable.

Experimental data and yields

The CFD results are compared with experimental chemostat data (van Gulik et al., 2000) and ideal-mixing simulations using both model of (Douma et al., 2010) (ID-1P) and the model of (Tang et al., 2017) (ID-9P) in Table 4. Both models are known to under-predict q_P around $\mu = 0.03\ h^{-1}$ compared to steady-state experiments. As shown in Table 4, the CFD simulations show a yield loss between 18% (TU-A) and 46% (TG-B) compared to the 9-pool model with ideal mixing. The real circulation time for the 54 m^3 reactor lies in between the extremes simulated here; based on the *Da*-correlation a yield loss of 22% is expected for $\tau_{circ} = 25.7$ s, sing $X_{E,11}$ value for $\mu = 0.03\ h^{-1}$.

For demonstration, we have also coupled the model of Douma directly to FLUENT (TU-1P), which yields an extreme 85% decrease in Y_{SP} and strong increase in Y_{SX} (discussed in detail in (Haringa et al., 2016)). These results are deemed unrealistic; the model of Douma was not designed to cope with rapid substrate concentration fluctuations, and the results show that applying the model in a situation where such fluctuations are present leads to extreme results. Returning now to the 9-pool model of (Tang et al., 2017); although the chemostat assumption used here introduced some simplifications, we are confident the overall trends hold, making the outlined method suitable for a quick assessment of the impact of design changes on the fermentation process. The most promising cases can subsequently be studied in more detail with 2-way coupling and experimental scale-down assessment.

Part II: Scale-down design

A scale-down design analysis is conducted for TU-B. Feed protocols for a single-vessel, fluctuating feed scale-down simulation with variable pulse duration were designed based on the arc-analysis methodology proposed in (Haringa et al., 2016). In contrast to earlier work, we did not divide the lifelines in regimes first; the arc analysis method was directly applied to the full (smoothed) lifelines,

using a reference value q_{ref} = 0.05 $q_{S,max}$. This means the lifelines are divided in feast-arcs ($q_S/q_{S,max}$ > 0.05) and famine-arcs ($q_S/q_{S,max}$ < 0.05). The arc duration τ_{arc} is registered as the time between two consecutive crossings of q_{ref}, as graphically indicated in Figure 3A,B. The distribution in τ_{arc} is reported in Figure 3C. For famine arcs, we can assume negligible magnitude: $q_S \approx 0$, regardless of duration. For feast arcs, the maximum $q_S/q_{S,max}$ for each arc-trajectory, called $\Omega_{S,max}$, is recorded. This gives a correlation between magnitude $\Omega_{S,max}$ and duration τ_{arc} (Figure 3D). The rationale behind q_{ref} = 0.05 $q_{S,max}$ follows from the results: the famine arcs show a complex distribution in τ_{arc}, but with negligible amplitude. For the feast arcs, the τ_{arc} distribution is comparatively simple, and a clear correlation between $\Omega_{S,max}$ and τ_{arc} exists. Together, these statistics quantify q_S-lifeline fluctuations and form a basis for representative scale-down simulation.

Representative profiles of alternating feast-famine arcs are generated from the τ_{arc} distributions by inverse transform sampling; for each feast event, the maximum q_S is retrieved from the mean τ_{arc}-$\Omega_{S,max}$ correlation. Determining the feed rate F is straightforward from the mass balance, assuming an instantaneously mixed lab-scale reactor. During famine intervals, F = 0 and $q_S \approx 0$ by construction. The most truthful approach is to feed gradually over a period of 0.5 τ_{arc}, such that the arc-shape is symmetric (Figure 3E); this requires the lab-scale to operate at the industrial biomass concentration C_X = 55 g_{dw}/kg (case ID-SD-55). Applying instantaneous feed pulse administration (Figure 3F) relaxes this to C_X = 27.5 g_{dw}/kg (ID-SD-27); the rate-of-change in q_S is reduced as q_S decreases over the entire period τ_{arc}. The lower C_X leads to a reduced effective viscosity (Roels et al., 1974) which may facilitate practical operation. However, it must be ensured the change in rate-of-change does not result in a different metabolic response. Further decreasing C_X inherently compromises either the fluctuation duration or magnitude, and thereby the representation of q_S-lifelines.

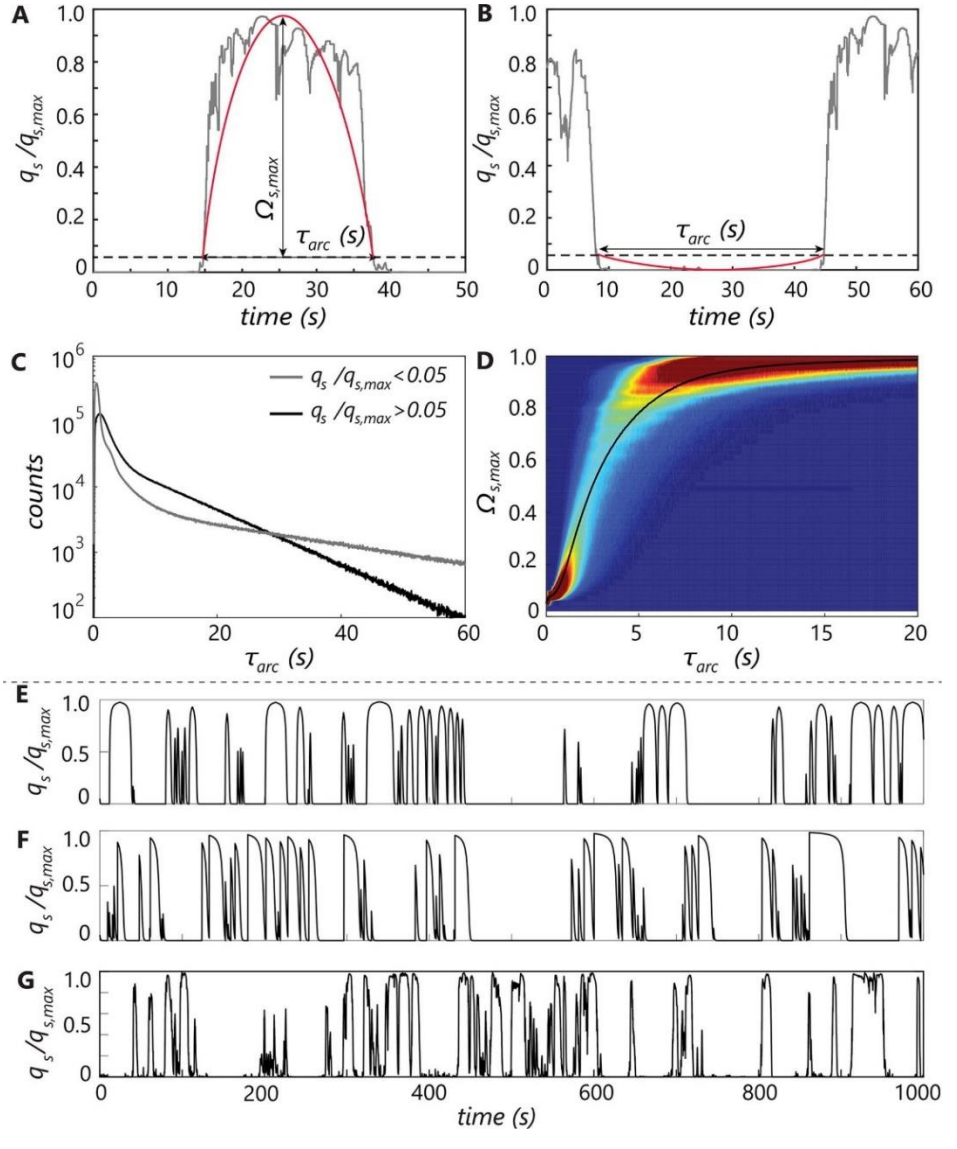

Figure 3 A and B: Graphical overview of the arc-analysis method for positive (feast arcs) and negative (famine arcs) fluctuations with respect to $q_{ref} = 0.05q_{S,max}$, respectively. C: Arc-time distributions famine (gray) and feast (black). D: Arc magnitude $\Omega_{S,max}$ under feast conditions as a function of arc time. Colors indicate bin fraction (normalized per timestep). Solid line: $\overline{\Omega_{S,max}}$ vs. τ_{arc} from CFD simulation. E: Generated lifeline, gradual feed pulses, $C_X = 55$ g/L. F: Generated lifeline, instantaneous feed pulses, $C_X = 27.5$ g/L. G: Example of CFD-lifeline for TU-B.

Table 5 Comparison of the instantaneous mixing cases ID-SD-27 and ID-SD-55 with CFD simulation TU-B. Inst. = instantaneous feed, Grad. = gradual feed. Δq_P is reported with respect to the ideal mixing benchmark ID-9P. The last three columns report the exposure (in %) to E (excess), L (limitation) and S (starvation) conditions, based on the definitions (Haringa et al., 2016)

Case	C_X [g/kg]	Feed	$\bar{\mu}$ [h^{-1}]	qP [molP/CmolX/h]	Δq_P	E	L	S
TU-B	55.0	-	0.033	$2.99 \cdot 10^{-4}$	-31.8%	6.8	36.2	57.0
ID-SD-55	55.0	grad.	0.0386	$2.90 \cdot 10^{-4}$	-33.7%	9.8	32.7	57.5
ID-SD-27	27.5	inst.	0.0340	$2.93 \cdot 10^{-4}$	-33.2%	8.0	34.2	57.8

Part III: Scale-down verification

In this section we assess the scale-down protocols of *part II*, first assuming instantaneous mixing and second using lab-scale CFD simulations. Note that instantaneous/ideal mixing in this context means the feed is immediately spatially distributed; due to the pulsed feed nature, there are temporal variations in q_S.

Instantaneous mixing

Both for ID-SD-55 and ID-SD-27, 5 statistically representative lifelines were generated and analyzed. Table 5 lists the metabolic response in q_P and μ compared to TU-B. Additionally, we conduct a regime analysis (using the definitions of (Haringa et al., 2016)) on the generated lifeline to determine the exposure to excess (E), limitation (L) and starvation (S) conditions. Case ID-SD-55 slightly over-estimates exposure to excess conditions. This results in a higher $\bar{\mu}$, mildly lower q_P and minor offsets in the intra-cellular pool sizes (reported in Supplementary material E), but overall we conclude that both cases excellently represent the large-scale simulation. The good performance of ID-SD-27 follows from the notion that the total uptake within a pulse of length t, $\int_0^t q_S dt$, is equal between the two pulse administration mothods, and the turnover time of X_{gly} is sufficiently slow to yield similar responses in X_{gly} (Supplementary material E). If the turnover time of X_{gly} was well below τ_{arc}, the metabolic response is expected to differ between the cases, and lowering C_X might not be allowed. We hence regard the possible reduction in C_X as a case-depended effect, and it should be evaluated as such. Furthermore, operating at industrial C_X whenever is possible may avoid unforeseen responses, not captured by the metabolic model. In case no predictions regarding the metabolic response are available, a scale-down simulator should in any case aim to produce the best possible replication of the extra-cellular environment (q_S-lifelines), and no compromises in C_X should be made.

CFD verification

In many cases, lab-scale fermentors can be assumed as ideally mixed. However, the combination of a very short τ_{rxn} (due to a low K_S) and mixing issues due to rheological issues (Moilanen et al., 2007) could lead to spatial heterogeneity in lab-scale fermentors. To assess whether this impacts scale-down performance, CFD simulation of a SD-simulator were conducted with C_X = 27.5 g/kg to probe the possible impact of non-instantaneous mixing. Spatial heterogeneity relates to the Damköhler number Da = τ_{cir}/τ_{rxn}. Here, τ_{circ} ≈ 0.5-3.3 s (for 600 and 100 RPM, respectively). As the C_S field is now dynamic, we employ a more general definition of the reaction time, $\overline{\tau_{rxn}} = \overline{C_S}/\overline{R_S} = (\overline{C_S} + K_S)/\left(q_{S,max} \cdot C_X\right)$ with $\overline{C_S}$ the volume average substrate concentration. Right after pulse administration, $\overline{C_S} \gg K_S$ and $Da \ll 1$: this implies the pulse will be mixed before $\overline{C_S}$ and thereby τ_{rxn} drop significantly, leading to a homogeneous broth and equal experiences by all micro-organisms in the domain.

This is reflected in the model response for both case CFD-SD-600 and CFD-SD-100. The q_S lifeline in Figure 4B (600 RPM) and C (100 RPM) show evidence of spatial heterogeneity directly following pulse administration, which for case CFD-SD-600 rapidly wears off, meaning the lifeline under the instantaneous mixing assuming is retrieved (Figure 4A). The heterogeneous period lasts longer for CFD-SD-100, but eventually the population synchronizes, and the metabolic response is hardly affected (Figure 4D). To comment on the role of non-ideal mixing in (aerated) SD-simulators with a high liquid viscosity μ_l, experimental measurements are required, but the results for CFD-SD-100 imply very poor mixing is required to yield significant heterogeneity in the population, and to yield a different metabolic response compared to the pulse-profile under the assumption of instantaneous mixing. This stems positive for practical application of fluctuating-feed SD-simulators.

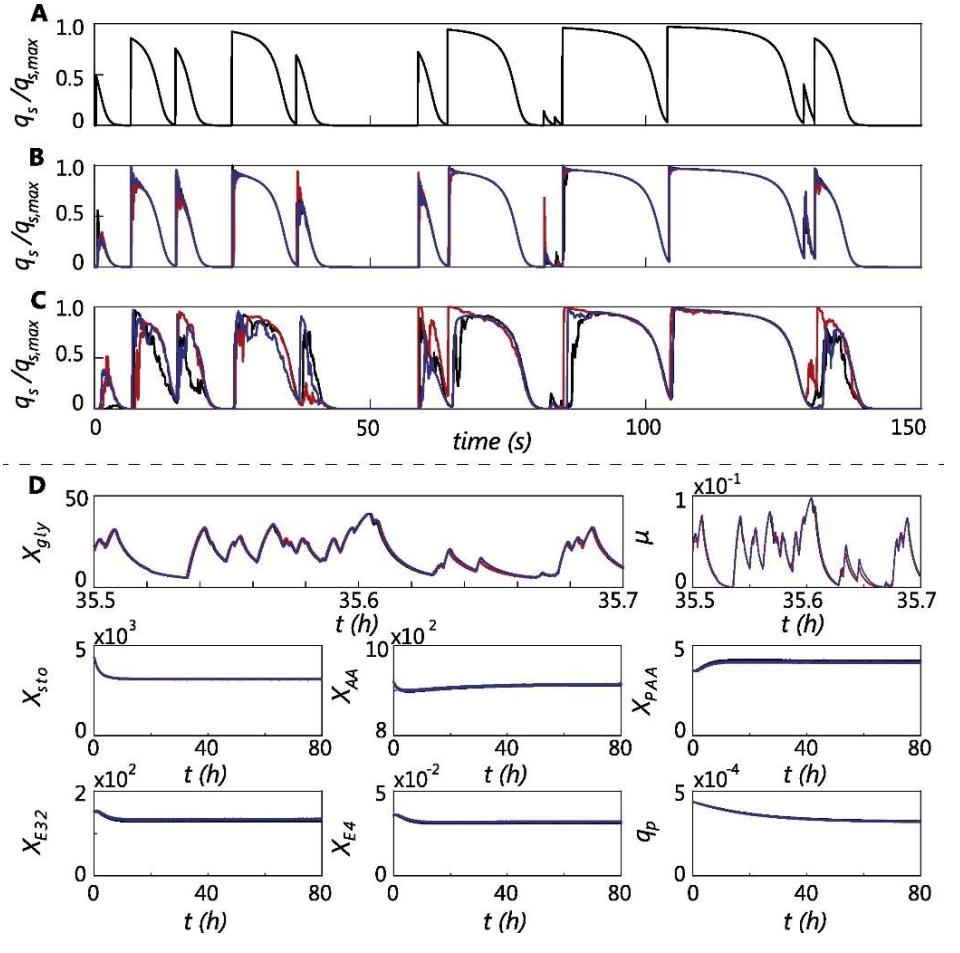

Figure 4 Particle tracks in scale-down simulation CFD. A: instantaneous mixing simulation. B: CFD-SD-600, 3 tracks. C: CFD-SD-100, 3 tracks. D: Response of intra-cellular pools in the 3L lab-scale reactor simulations. Black line: instantaneous mixing results. Red line: CFD-SD-600, average over 5000 tracks. Blue line: CFD-SD-100 simulation, average over 5000 tracks. Intra-cellular pools X_{gly}, X_{AA}, X_{sto} and X_{PAA} have units µmol/g_{dw}. µ has units h^{-1}. q_P has units mol/$Cmol_X$/h. All other pools are dimensionless.

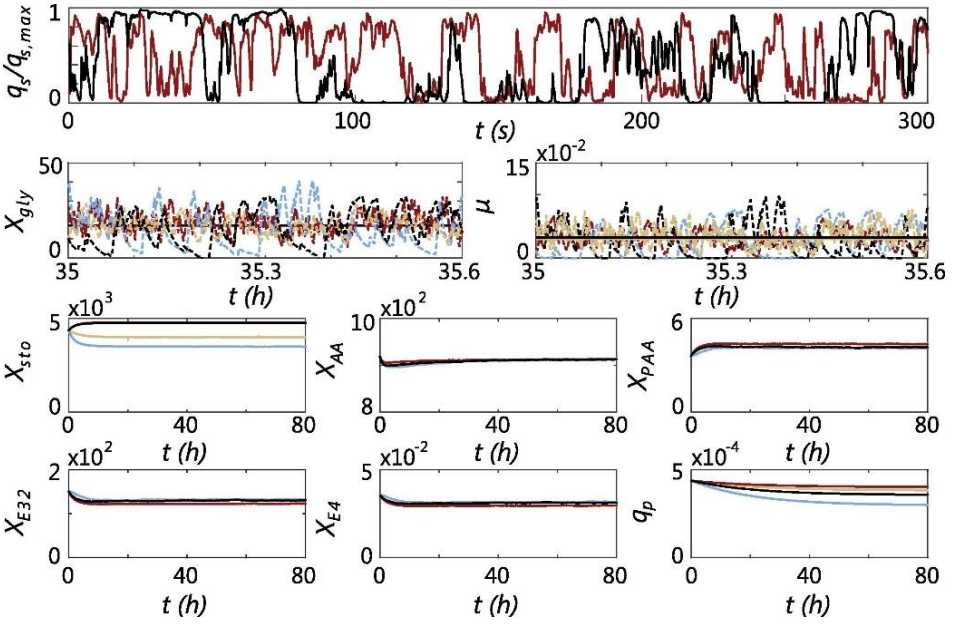

Figure 5 Long-term 9-pool model response large-scale simulations for TU-A (black), MU-A (red), TU-B (gray), MU-B(orange). The top panel shows examples of q_S-lifelines for single parcels in TU-A, MU-A. For X_{gly} and μ, the solid lines represent the average, the dashed lines represent single parcel tracks to indicate the variations. All other solid lines represent averages of 100 parcels. Intra-cellular pools X_{gly}, X_{AA}, X_{sto} and X_{PAA} have units $\mu mol/g_{dw}$. μ has units h^{-1}. q_P has units $mol/Cmol_X/h$. All other pools are dimensionless.

Part IV: Design optimization

Part I revealed that reducing the frequency of q_S variations the amplitude of X_{gly} fluctuations, which reduced inhibition of q_P. (Cronin, 1992) reduced τ_{95} by a factor 2-2.5 by placing the feed point just below the top impeller (Van't Riet & der Lans, 2011; Vrabel et al., 1999). We find τ_{95} = 23s (1-phase hydrodyn., MU-A/B) when the feed is placed in the top-impeller discharge stream, a 2.7-fold reduction in τ_{95} compared to the top feed. This exceeds expectations and may be excessively low for a true penicillin fermentation when rheology and aeration are account for, but we accept this result for the sake of demonstration. The pool response for simulations MU-A and MU-B is reported in Figure 5.

Compared to TU-A, the q_S-lifelines for MU-A show a lower fluctuation amplitude, and strong reduction in fluctuation duration (Figure 5**,** top). This translates to much milder X_{gly} variations that directly relate to a higher q_P for MU-A/B cases (Table 6). Again, $\overline{X_{gly}}$ and hence $\bar{\mu}$ remains virtually equal between the cases. The q_P loss is reduced to 8.6% (with respect to ID-9P), where the top-feed case with equal τ_{circ}, TU-A, showed a yield loss of 17%. The reduced exposure to starvation conditions furthermore is observed to yield a higher X_{sto} for MU-1 cases. An alternative process improvement may be to modify increase K_S by

modifying the glucose transporter, thereby reducing sensitivity to C_S fluctuations. Within the current metabolic model, this also requires altering the sensitivity of the storage/release process to C_S; we did not further pursue this option within the scope of this work.

Table 6 Comparing yields and productivity between experimental data (van Gulik et al., 2000), the black box (BB) model (Douma et al., 2010) and the 9-pool model (Tang et al., 2017) with ideal mixing assumption and the 9-pool (9P) of (Tang et al., 2017)

Model	Case	F [$mol_S/Cmol_X/h$]	μ [h^{-1}]	q_P [$mol_P/Cmol_X/h$]	Y_{SX} [$Cmol_X/Cmol_S$]	Y_{SP} [mol_P/mol_S]
9P	ID mix	0.0125	0.033	$4.40 \cdot 10^{-4}$	0.44	0.035
9P	TU-A	0.0125	0.033	$2.99 \cdot 10^{-4}$	0.44	0.024
9P	TU-B	0.0125	0.032	$3.57 \cdot 10^{-4}$	0.43	0.029
9P	MU-A	0.0125	0.030	$4.02 \cdot 10^{-4}$	0.40	0.032
9P	MU-B	0.0125	0.030	$3.83 \cdot 10^{-4}$	0.40	0.031

Part V: Industrial-scale fed-batch simulation

The long-term metabolic response in an industrial fed-batch reactor is simulated; the dynamic feed profile that was supplied to the simulation is reported in Figure 6A. C_X and μ are well captured (Figure 6B and C, resp.), although an ideal-mixed simulation with model of (Douma et al., 2010) (ID-FB) better captures the final 20h. The 9-pool CFD simulation, however, performs superior in predicting the gradual reduction in q_P (Figure 6D). The initial offset results from the lower peak q_P prediction by the 9-pool model around $\mu = 0.03 h^{-1}$.

The trends in intra-cellular pools (Figure 6E) reveal major temporal changes in the pool averages (solid lines), as well as the emergence of significant heterogeneity within the population; the dashed lines in Figure 6E represent the pool size standard deviation over 2500 tracks. The decreasing trend in all enzyme pools is a consequence of the reduction in $\bar{\mu}$ to 0.01 h^{-1}; the drop in $X_{E,32}$ reduces the PAA export capacity, giving rise to a strong PAA build-up. Similarly, a buildup in X_{sto} is observed. As before, the AA pool is least sensitive, although it undergoes some changes in later stages. The strong rise in population heterogeneity roughly coincides with the switch to a constant feed rate F ≈ 1.6 $kg/m^3/h$.

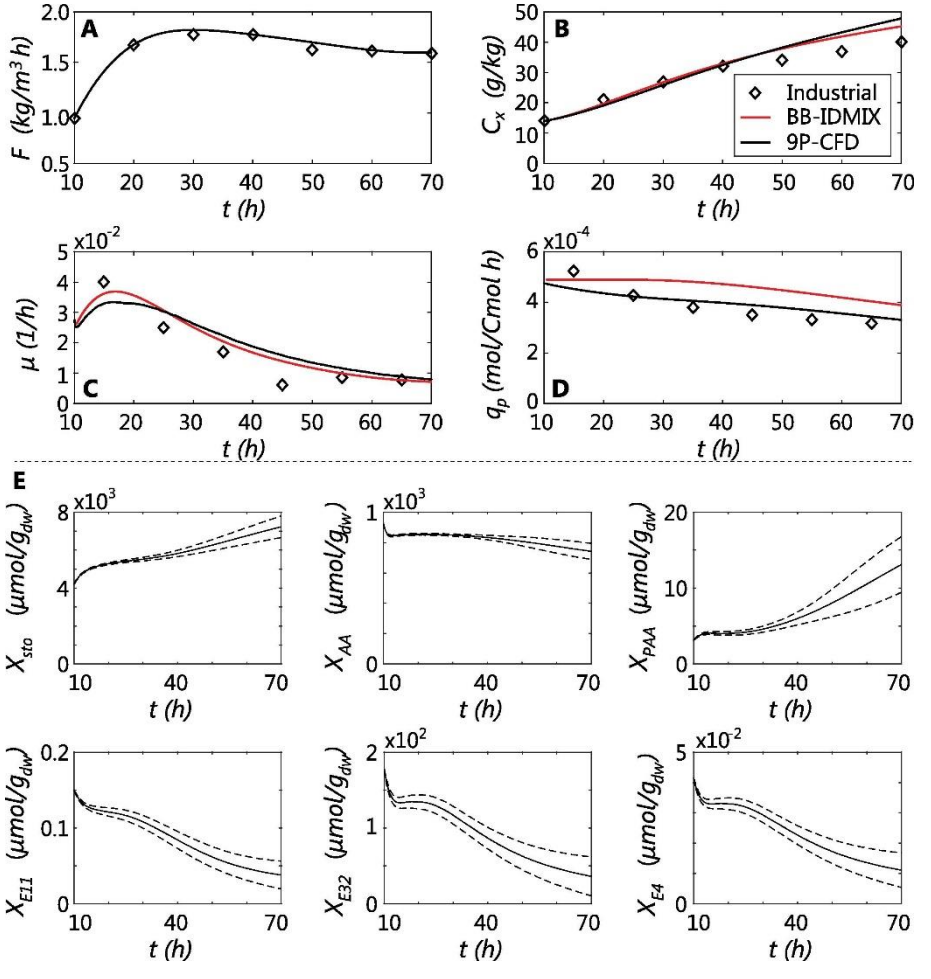

Figure 6 Response profiles of simulation TU-FB, comparing industrial data (diamonds), CFD response (black line) and an ideal-mixed black-box simulation (red line). A: Feed rate per unit reactor volume. B: Biomass concentration. C: Growth rate. D: Penicillin production. E: the response of 6 slow-responding intra-cellular pools, with the solid line the mean and dashed lines ±1 st. dev.

For brevity, figures further detailing the onset of and degree of population heterogeneity are reported in Supplementary material F. The high degree of heterogeneity in the enzyme pools may be surprising at first glance; their adaption timescale strongly exceeds τ_{circ}, and all parcels are expected to observe highly similar C_S fluctuations during the cultivation. The link between μ and $X_{E,11}$ plays a key role; a parcel residing in famine zone ($\mu \approx 0$) for a prolonged time undergoes a reduction in $X_{E,11}$. This reduces subsequent substrate uptake q_S with respect to the population average, further decreasing μ and hence $X_{E,11}$, thereby amplifying the original disturbance. A deeper analysis in

Supplementary material F shows that parcels with a below-average $X_{E,11}$ early on end in the bottom of the $X_{E,11}$ distribution. A prolonged exposure to excess conditions could reverse the disturbance, but the results show that this practically rarely occurs. The further the deviation from the population average, the more unlikely recovery becomes. We do recognize that the observation that starving organisms lower their uptake capacity appears counter-intuitive; we stress this is a model prediction, that should be verified experimentally, thereby showing how coupled simulation may generate new targets for experimental investigation. The variation in all other intra-cellular pools eventually stems from the variations in $X_{E,11}$; in the chemostat simulations of *part I*, where $X_{E,11}$ was necessarily fixed, no population heterogeneity was observed (Supplementary material E).

The parcels with high $X_{E,11}$ are the fastest growers; some acquire double the population average biomass over the cultivation time, whereas for the poorest growers $\mu \approx 0$ in the late process stage. As a low $\bar{\mu}$ has a negative effect on q_P, the fastest growers are also among the best penicillin producers, whereas the poor growers mostly accumulate storage material (Supplementary material F). Whether or not the predicted degree of heterogeneity is realistic requires an experimental scale-down study where population heterogeneity is probed on the single-cell level (Delvigne & Goffin, 2014; Zenobi, 2013). The simulations predict notable heterogeneity enzyme levels, which may provide suitable targets for fluorescent marking for experimental quantification. Besides bench-scale scale-down, the use of microfluidic tools (Dusny & Schmid, 2015; Grünberger et al., 2014) with highly controllable substrate feed rates may be a promising route towards studying the effects of substrate variations.

Concluding remarks

We reported on the use of coupled hydrodynamic -metabolic simulations to assess large-scale fermentation processes in five parts: (1) industrial-scale metabolic response analysis, (II) scale-down design, (III) scale-down verification, (IV) design optimization and (V) industrial-scale fed-batch analysis. Combined, these steps provide a methodology for the analysis, scale-down and optimization of large-scale fermentation processes. Combing the 9-pool metabolic model for *P.chrysogenum* of (Tang et al., 2017) with CFD simulations of a 54 m^3 fermentor (Haringa et al., 2016) (*part I*), we report a predicted penicillin yield loss of 18-45%, which correlated linearly with the Damköhler number, assuming a chemostat cultivation with 1-way metabolic coupling to simplify the simulation. The yield loss resulted from level of glycolytic

intermediates, relating the circulation time and substrate uptake capacity of the organism. These observations provide targets for reactor and metabolic optimization.

The arc analysis methodology of (Haringa et al., 2016) was used to design a representative single-vessel, dynamic-feed scale-down simulator, based on the q_S-lifeline (*part II*). Numerical evaluation to accurately reflect the metabolic response recorded in the industrial reactor. Capturing the rate-of-change experienced by micro-organisms on the industrial scale requires operating the lab-scale at the industrial biomass concentration C_X = 55 g/kg. The 9-pool model response shows that it is possible to compromise the rate of change without changing the metabolic response to some degree, allowing for a factor 2 reduction in C_X. This would facilitate operation, but may induce metabolic responses for which the model does not account. Hence, we do emphasize that operating the scale-down simulator at industrial C_X is preferred, especially when no metabolic response prediction is available, to ensure the best possible replication of q_S-lifelines. CFD simulations of the proposed scale-down simulator with pulsed feeding showed that non-instantaneous mixing at the lab scale (assessed for circulation times of 0.55 and 3.3 s) did not compromise the metabolic response, which gives confidence in the practical application of the proposed simulator. This operational window may depend on the organism and geometry, and should be evaluated per-case.

Changing the substrate feed location in the industrial-scale fermentor to improve substrate distribution reduced the ield loss from 18.4% to 8.6% (*part IV*). This showcases the prospects for in silico design optimization. To conclude, we present a 60 h fed-batch study (*part V*) with 2-way metabolic coupling, showing good agreement in μ and the glucose transport capacity $X_{E,11}$. The results illustrate the importance of simulating fed-batch dynamics including 2-way coupling to capture population heterogeneity. We do stress this does not imply that the 1-way coupled approach is futile; it is preferred for a rapid assessment of the metabolic response to design changes. We do, however, advise that the most promising chemostat cases are subsequently simulated with -way coupling (and/or experimentally assessed) to verify their performance when population heterogeneity is included.

Altogether, we outlined the different roles of coupled hydrodynamic-metabolic modeling in the assessment and improvement of large-scale fermentor designs. In future work, the proposed scale-down simulators are to be tested to verify model predictions; the predicted yield loss and population heterogeneity

provide clear targets for assessment and model verification. There is room for improvement in both the CFD models and dynamic metabolic models, which would greatly benefit from a broader availability of industrial-scale data for verification. Such improvements act towards increasing accuracy and reliability of the here-shown coupled CFD approach, but will not influence the methodology in itself. We believe the here-presented methodology, combined with practical scale-down simulation, opens up a new approach towards rational fermentor design and scale-dup, accounting for the effect of large-scale reactor heterogeneity.

References

Cronin, D. G. (1992). An experimental study of the mixing in a proto-fermenter agitated by dual Rushton turbines. *Trans. Inst. Chem. Eng., PartC*, *72*, 35–40.

de Jonge, L. P., Buijs, N. A. A., ten Pierick, A., Deshmukh, A., Zhao, Z., Kiel, J. A. K. W., Heijnen, J. J., & van Gulik, W. M. (2011). Scale-down of penicillin production in Penicillium chrysogenum. *Biotechnology Journal*, *6*(8), 944–958.

Delvigne, F., & Goffin, P. (2014). Microbial heterogeneity affects bioprocess robustness: Dynamic single-cell analysis contributes to understanding of microbial populations. *Biotechnology Journal*, *9*(1), 61–72.

Douma, R. D., Verheijen, P. J. T., de Laat, W. T. A. M., Heijnen, J. J., & van Gulik, W. M. (2010). Dynamic gene expression regulation model for growth and penicillin production in Penicillium chrysogenum. *Biotechnology and Bioengineering*, *106*(4), 608–618.

Dusny, C., & Schmid, A. (2015). Microfluidic single-cell analysis links boundary environments and individual microbial phenotypes. *Environmental Microbiology*, *17*(6), 1839–1856.

Enfors, S.-O., Jahic, M., Rozkov, A., Xu, B., Hecker, M., Jürgen, B., Krüger, E., Schweder, T., Hamer, G., O'beirne, D., & others. (2001). Physiological responses to mixing in large scale bioreactors. *Journal of Biotechnology*, *85*(2), 175–185.

Grünberger, A., Wiechert, W., & Kohlheyer, D. (2014). Single-cell microfluidics: opportunity for bioprocess development. *Current Opinion in Biotechnology*, *29*, 15–23.

Gunyol, O., Noorman, H. J., & Mudde, R. F. (2009). CFD simulations of a large-scale fermenter with muliple impellers. *8th World Congress of Chemical Engineering*, 1–4.

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017). Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669.

Haringa, C., Noorman, H. J., & Mudde, R. F. (2017). Lagrangian modeling of hydrodynamic–kinetic interactions in (bio) chemical reactors: practical implementation and setup guidelines. *Chemical Engineering Science*, *157*, 159–168.

Haringa, C., Tang, W., Deshmukh, A. T., Xia, J., Reuss, M., Heijnen, J. J., Mudde, R. F., & Noorman, H. J. (2016). Euler-Lagrange computational fluid dynamics for (bio) reactor scale down: an analysis of organism lifelines. *Engineering in Life Sciences*, *16*(7), 652–663.

Khopkar, A. R., Aubin, J., Xuereb, C., le Sauze, N., Bertrand, J., & Ranade, V. v. (2003). Gas- Liquid Flow Generated by a Pitched-Blade Turbine: Particle Image Velocimetry Measurements and Computational Fluid Dynamics Simulations. *Industrial & Engineering Chemistry Research*, *42*(21), 5318–5332.

Lapin, A., Müller, D., & Reuss, M. (2004). Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler- Lagrange methods: Traveling along the lifelines of single cells. *Industrial & Engineering Chemistry Research*, *43*(16), 4647–4656.

Lapin, A., Schmid, J., & Reuss, M. (2006). Modeling the dynamics of E. coli populations in the three-dimensional turbulent field of a stirred-tank bioreactor—A structured–segregated approach. *Chemical Engineering Science*, *61*(14), 4783–4797.

Lara, A. R., Galindo, E., Ram\'\irez, O. T., & Palomares, L. A. (2006). Living with heterogeneities in bioreactors. *Molecular Biotechnology*, *34*(3), 355–381.

Luo, H., & Svendsen, H. F. (1996). Theoretical model for drop and bubble breakup in turbulent dispersions. *AIChE Journal*, *42*(5), 1225–1233.

McClure, D. D., Aboudha, N., Kavanagh, J. M., Fletcher, D. F., & Barton, G. W. (2015). Mixing in bubble column reactors: experimental study and CFD modeling. *Chemical Engineering Journal*, *264*, 291–301.

McClure, D. D., Kavanagh, J. M., Fletcher, D. F., & Barton, G. W. (2016). Characterizing bubble column bioreactor performance using computational fluid dynamics. *Chemical Engineering Science*, *144*, 58–74.

Moilanen, P., Laakkonen, M., Visuri, O., & Aittamaa, J. (2007). Modeling local gas- liquid mass transfer in agitated viscous shear-thinning dispersions with CFD. *Industrial & Engineering Chemistry Research*, *46*(22), 7289–7299.

Neubauer, P., & Junne, S. (2010). Scale-down simulators for metabolic analysis of large-scale bioprocesses. *Current Opinion in Biotechnology*, *21*(1), 114–121.

Nikerel, I. E., Verheijen, P. J. T., Gulik, W. M. van, & Heijnen, J. J. (2012). Model-based design of superior cell factory: an illustrative example of Penicillium chrysogenum. In *Systems Metabolic Engineering* (pp. 221–270). Springer.

Noorman, H. (2011). An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnology Journal*, *6*(8), 934–943.

Oosterhuis, N. M. G., & Kossen, N. W. F. (1984). Dissolved oxygen concentration profiles in a production-scale bioreactor. *Biotechnol. Bioeng.;(United States)*, *26*(5).

Roels, J. A., den Berg, J., & Voncken, R. M. (1974). The rheology of mycelial broths. *Biotechnology and Bioengineering*, *16*(2), 181–208.

Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J. J., Xia, J., Chu, J., & others. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by Penicillium chrysogenum. *Biotechnology and Bioengineering*, *114*(8), 1733–1743.

van Gulik, W. M., de Laat, W., Vinke, J. L., & Heijnen, J. J. (2000). Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. *Biotechnology and Bioengineering*, *68*(6), 602–618.

Van't Riet, K., & der Lans, R. (2011). Mixing in bioreactor vessels. *Comprehensive Biotechnology*, *2*, 63–80.

Vrabel, P., der Lans, R., Cui, Y. Q., & Luyben, K. C. A. M. (1999). Compartment model approach: Mixing in large scale aerated reactors with multiple impellers. *Chemical Engineering Research and Design*, *77*(4), 291–302.

Wang, G., Chu, J., Noorman, H., Xia, J., Tang, W., Zhuang, Y., & Zhang, S. (2014). Prelude to rational scale-up of penicillin production: a scale-down study. *Applied Microbiology and Biotechnology*, *98*(6), 2359–2369.

Wang, G., Tang, W., Xia, J., Chu, J., Noorman, H., & van Gulik, W. M. (2015). Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. *Engineering in Life Sciences*, *15*(1), 20–29.

Wang, G., Wu, B., Zhao, J., Haringa, C., Xia, J., Chu, J., Zhuang, Y., Zhang, S., Heijnen, J. J., van Gulik, W., & others. (2018). Power input effects on degeneration in prolonged penicillin chemostat cultures: a systems analysis at flux, residual glucose, metabolite, and transcript levels. *Biotechnology and Bioengineering*, *115*(1), 114–125.

Zenobi, R. (2013). Single-cell metabolomics: analytical and biological perspectives. *Science*, *342*(6163), 1243259.

Acknowledgments

We want to thank our colleagues at ECUST Shanghai, DSM Sinochem pharmaceuticals and the DSM biotechnology center for our fruitful collaboration. We are grateful to the DSM biotechnology center for making fermentation data available to test our models. Special thanks to Prof. Matthias Reuss and Prof. Jianye Xia for their participation in many discussions and the insights they provided. This work has been conducted within a multi-party research project, between DSM-Sinochem Pharmaceuticals, TU Delft, East China University of Science and Technology and Guojia, subsidized by NWO and MoST (NWO-MoST Joint program, 2013DFG32630). All sponsors are gratefully acknowledged.

Supplementary material

Supplementary data associated with this Chapter can be found, in the online version, at https://doi.org/10.1016/j.ces.2017.09.020.

Chapter 4

Gene-regulation kinetic model capable of reproducing both long- and short-term storage carbohydrate dynamics in *S. cerevisiae*

Abstract

In industrial scale fermentation processes, microorganisms encounter a constantly changing environment, due to insufficient mixing in the large-scale bioreactors. Kinetic modelling is an ideal tool to quantitatively describe and predict microbial responses to these changes, in well-controlled scale-down systems and at production scale, via integration in a Computational Fluid Dynamics (CFD) platform. In this work, we constructed two different types of kinetic models (Computational Reaction Dynamics - CRD) to describe the dynamics of trehalose and glycogen, two important cellular storage carbohydrates in the yeast *Saccharomyces cerevisiae*, under multiple steady states and glucose oscillation conditions. Both storage kinetic models were connected to a simple black box model which describes the glucose uptake, cell growth and respiration. The first kinetic model is a complete black box model using conventional Michaelis-Menten type kinetics. Despite its capability of describing cell phenotypes under steady state conditions, this model cannot predict cellular behavior during short term dynamics, such as glucose feast-famine oscillations, due to structural limitations. The second model is a gene-regulation model where the dynamics of the synthesis and degradation of storage compounds is regulated at both enzyme and metabolite level at the same time. Even with simple linear enzyme kinetics, the gene-regulation model can already precisely reproduce all steady states and qualitatively the feast-famine scenario. This model also proves to be stable and maintains its dynamic features when integrated with cell lifelines in a 22 m^3 pilot tank. We foresee that the properties of the gene regulation model make it suitable for integration in a full-scale Euler-Lagrange CFD-CRD simulation or compartment based large scale bioreactor simulation. Moreover, unlike a black box model, the presented gene-regulation module can be incorporated into a fully structured growth model of *Saccharomyces cerevisiae* and thus gain in predictive capacity.

Keywords

kinetic model, gene regulation, steady states, feast-famine, *Saccharomyces cerevisiae*

Introduction

Microorganisms accumulate metabolites, polymers and energy carriers to cope with stress conditions, such as nutrient starvation, osmotic pressure, product inhibition, etc. Those stresses may not only be encountered in natural environments but could also be observed by cells in a well-controlled industrial scale fermentor. Among the various kinds of metabolites helping cells to mitigate those stresses, glycogen and trehalose have been studied most extensively (Choi et al., 2018; Gancedo & Flores, 2004; Nadal-Rey et al., 2021; Noorman, 2011; G. Wang, Zhao, et al., 2019; X. Wang et al., 2021; Wilson et al., 2010).

In *Penicillium chrysogenum*, trehalose acts as an emergency carbon source during short-term glucose starvation (L. de Jonge et al., 2014a). It was also found that the trehalose level in the cells was dynamically balanced by continuous fine-tuning of synthesis and degradation, which can be considered as an ATP consuming futile cycle. Nevertheless, the presence of such a trehalose synthesis and degradation cycle seems to be beneficial when we consider the cell's secondary metabolism. (X. Wang et al., 2021) knocked out trehalose synthesis by deleting the tps1 and tps2 genes of *P. chrysogenum* Wisconsin 54–1255 and observed not only morphology changes but also a significant drop in penicillin G production. Further, a high penicillin G producing strain featured a higher intracellular trehalose level compared to the reference wild type strain.

In *Saccharomyces cerevisiae*, trehalose and glycogen are deployed as carbon source when carbon supply is limited (C. a. Suarez-Mendez, 2015; C. A. Suarez-Mendez et al., 2014). Moreover, is has been observed that trehalose plays an important role in re-balancing the cell to a stable healthy status after experiencing a glucose pulse (C. A. Suarez-Mendez et al., 2016, 2017). In a trehalose storage-defective yeast transformant, a low glucose uptake capacity, lower respiration rate and lower energy charge status was observed (van Heerden et al., 2014).

Glycogen, the other reserve compound in yeast, has its own merits in being used as a carbon source. This long-chain polyglucan has little effect on the internal osmotic pressure of the cell and is crucial for yeast to survive long-term nutrient deprivation (Silljé et al., 1999). (Wilson et al., 2010) comprehensively reviewed the regulation of glycogen metabolism for both yeasts and bacteria. At fundamental level, they shared certain structure similarities while the process of regulation could be quite distinct.

In practical applications of modern biotechnology, the 'scale-up effect' (Delvigne et al., 2017; Noorman, 2011; G. Wang et al., 2014) is a serious challenge for industrial scale bioprocesses. Caused by non-ideal mixing and mass transfer limitations in combination with fast metabolic uptake, concentration gradients of substrate, dissolved oxygen and pH could be formed in the reactor space. When cells follow the liquid flow in the fermentor, they experience those gradients at a time scale of seconds to minutes and in many cases these gradients exert a negative impact on the key performance indicators (e.g. titer, rate and yield). To quantitatively evaluate the impact of these gradients on the cell physiology in industrial scale processes, modeling techniques such as computational fluid dynamics (CFD) coupled with computational reaction dynamics (CRD) were developed and successfully applied to both *Saccharomyces cerevisiae* (Haringa, Deshmukh et al., 2017b; Lapin et al., 2004) and *Penicillium chrysogenum* (Haringa, Tang et al., 2017) fermentations. However, trehalose and glycogen, two key compounds for yeast cells to deal with stress and substrate oscillations, were not highlighted before as key metabolic response factors. Due to the nature of their cyclic metabolism, the study of the storage compounds kinetics is less straightforward due to difficulties in obtaining the actual fluxes towards and away from the trehalose and glycogen nodes. (C. A. Suarez-Mendez et al., 2016) calculated the absolute in-/outflux of both storage compounds by providing the carbon supply in the form of ^{13}C labelled glucose during multiple steady states in chemostats. The absolute fluxes were calculated from the slope of increasing abundancy of ^{13}C labelled glucose. Further, a dynamic ^{13}13C-based metabolic flux analysis was applied in a dynamic oscillation scenario to identify the flux profiles during a glucose pulse condition (Aboka et al., 2009; L. de Jonge et al., 2014a; C. A. Suarez-Mendez et al., 2017).

In this study, we focused on modeling the cyclic metabolism of trehalose and glycogen, based on quantitative ^{13}C-labelled flux information. Two different kinetic models with different mechanisms of regulation were proposed, both capable of predicting a wide range of steady states reasonably well. However, when applied to a more industrially relevant scale-down system, the simple black-box based storage kinetics were not able to predict the dynamic storage responses within time scales of seconds to minutes, which appeared to be a structural problem. On the other hand, a gene-regulation based model, featuring decoupled regulation at both enzyme and metabolite levels, can reproduce the cell's periodic response to the feast-famine cycle and proved to be suitable for CFD integration by testing it with lifelines obtained in a 22 m^3

pilot tank. This piece of work clearly showed the necessity of incorporating gene regulation in the kinetic model, especially for modeling the dynamics at different timescales.

Material and Methods

Data sources

Three published sets of experimental data have been used in this Chapter for model construction and validation. The summary of the data sources is given in Table 1.

In short, we referred to (Canelas et al., 2011) for yeast growth in steady state chemostat cultures of which the dilution rates ranged from 0.02 h^{-1} to 0.30 h^{-1}. The measured glucose uptake rates, respiration rates and residual glucose concentrations for various growth rates were collected. Because the core black box model doesn't cover ethanol production and uptake as occurring at higher dilution rates, we only use data obtained below the critical dilution rate in this Chapter.

For the profile of storage carbon, we used data from (C. A. Suarez-Mendez et al., 2016) where ^{13}C-labelled MFA was carried out to obtain not only the intracellular levels of trehalose and glycogen but also the absolute in-/out-fluxes of those storage compounds.

Finally, we use glucose oscillation data from (C. A. Suarez-Mendez et al., 2017) for model validation under dynamic conditions. In this experiment, dynamic flux estimation based on a consensus model for yeast (Herrgård et al., 2008) with significant simplification, was executed using ^{13}C-tracing information for rapid, periodic 400s glucose oscillations. Both the levels of trehalose and glycogen, which were almost constant in one cycle due to the large turnover times of these pools and the absolute fluxes into and from these storage compounds were used for model validation.

Table 1. Summary of published data used in this Chapter

No.	Type	Strain	Experimental Setting	Reference
1	Steady states	CEN.PK 113-7D	Steady states reached via upward/downward step change in dilution rates	(Canelas et al., 2011)
2	Steady states		Steady states reached via downward step change in dilutions rates, ^{13}C labelling experiment at the end phase of each step	(C. A. Suarez-Mendez et al., 2016)
3	Glucose oscillation		Dilution rate at 0.01h^{-1}; feed profile repeated every 400s: fast feed (feast phase) in the first 20s and no feed (famine phase) in remaining 380s	(C. A. Suarez-Mendez et al., 2017)

Data processing and calculations

In (C. A. Suarez-Mendez et al., 2016), only the degradation rates and intracellular levels of trehalose and glycogen were reported. In this Chapter, the influx of trehalose and glycogen were also required and can be calculated from their respective molar balances:

$$\frac{dX_i}{dt} = v_{i,storage} - v_{i,release} - \mu \cdot X_i \qquad \text{Eq. 1}$$

where i stands for trehalose or glycogen, X is the intracellular level and μ is the specific growth rate. During steady state conditions, the left-hand side of Eq.1 equals zero. For these conditions the $v_{i,storage}$ can be obtained from Eq. 2 (C. A. Suarez-Mendez et al., 2016):

$$v_{i,storage} = v_{i,release} + \mu \cdot X_i \qquad \text{Eq. 2}$$

This leads to final absolute storage and release fluxes of both trehalose and glycogen as shown in Figure 1. The resulting fluxes, plotted against the specific growth rate, show completely different patterns for trehalose and glycogen. This provides strong evidence and necessity to apply different kinetics to describe the dynamics of the levels of trehalose and glycogen, respectively.

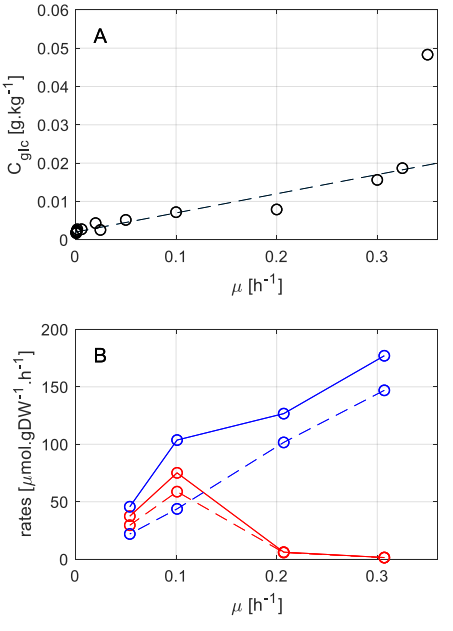

Figure 1. Residual glucose concentration (A) and specific synthesis and degradation rates of trehalose and glycogen (B) as a function of the growth rate in *S. cerevisiae* CEN.PK113-7D (C. A. Suarez-Mendez et al., 2016). For B, Red: Trehalose, Blue: glycogen, solid lines: synthesis rates, dashed lines: degradation rates.

Software and Script

All simulations were executed in MATLAB 2018a using the AMIGO toolbox (https://sites.google.com/site/amigo2toolbox). Ordinary Differential Equations (ODEs) were solved via CVODES (https://computing.llnl.gov/projects/sundials) by compiling the MATLAB model structure into C++ via Mex. Parameter estimations were executed via enhanced Scatter Search (eSS) with local solver fmincon (MATLAB).

Results and discussion

Model construction

A black box-based model for cell growth was constructed, coupled with kinetic equations for the synthesis and degradation rates of trehalose and glycogen. (Figure 2). In this model the specific glucose consumption rate is assumed to be a function of the residual glucose concentration according to saturation kinetics (Kesten et al., 2015; Lao-Martil et al., 2022; Smallbone et al., 2013). After consumption the glucose is assumed to be distributed between cell growth, maintenance and storage carbohydrate synthesis, in this case trehalose and glycogen. For the distribution of glucose between growth and maintenance the well-known Herbert-Pirt equation is used with as parameters the maximum biomass yield and a maintenance coefficient, which were fitted to published data (Canelas et al., 2011). Because we separated the storage and degradation process of trehalose and glycogen, we divided the maintenance parameter into two parts: a constant part and a variable part. The latter one represents energy costs, in terms of ATP, of the storage/degradation cycle, and is converted into glucose consumption via the calculated yield of ATP on glucose during fully aerobic catabolism. In this black box model, the biomass is defined as lean biomass, which implies that no storage compounds are included. Because the sum of glycogen and trehalose can vary between almost 0% to 15% of the cell weight (Canelas et al., 2011; Wilson et al., 2010), it is preferable to describe the trehalose and glycogen development separately for a better accuracy of the model. Therefore, we adapted the biomass formula from (Aboka et al., 2012) by excluding the storage compounds. The adapted lean biomass formula for our model was then recalculated as $C_1H_{1.761}N_{0.176}O_{0.541}$. The oxygen uptake and CO_2 production rates are calculated based on carbon and redox balancing.

The growth model is combined with storage compound synthesis and degradation to show the impact of trehalose and glycogen dynamics on both the carbon flow into central metabolism and the cell's macromolecular composition. In this research work, two different models were proposed to describe the observed kinetic features of both trehalose and glycogen synthesis and degradation.

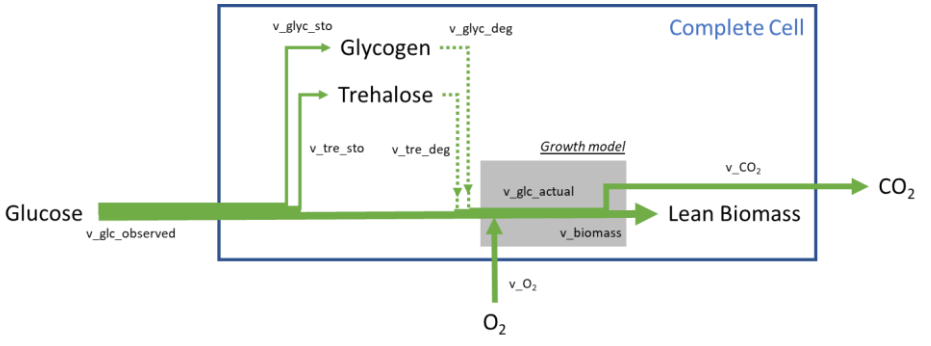

Figure 2. Schematic of the yeast model in this Chapter. The storage and degradation kinetics of trehalose and glycogen vary with different model scenarios. The growth model, which represents carbon distribution, cell growth and respiration remains the same.

Storage kinetics as a black box

Rate equations in this modeling effort follow classic Michaelis-Menten kinetics for all four reactions (i.e. v_tre_sto, v_tre_deg, v_glyc_sto, and v_glyc_deg). Trehalose accumulation was maximal at very low specific growth rates. Both trehalose storage and degradation rates showed their highest values at $\mu = 0.1$ h^{-1} (Figure 1), whereas trehalose was almost depleted with very low reaction rates at $\mu > 0.2$ h^{-1}. This suggests both rates could be activated (for $0h^{-1} < \mu < 0.1$ h^{-1}) and inhibited (for $\mu > 0.1h^{-1}$). (Canelas et al., 2011) showed that before reaching the critical growth rate ($\mu = 0.25$ h^{-1}) in glucose limited chemostats there is a linear relation between the specific growth rate and the specific glucose uptake rate (Canelas et al., 2011). The latter can further be well described by the extracellular glucose concentration(Figure 1A). Therefore, we proposed using extracellular glucose as both activator and inhibitor of trehalose storage and degradation. Further, we added trehalose inhibition (product inhibition) to the storage kinetics as well as trehalose activation (substrate activation) to the degradation kinetics.

$$v_{tre,sto} = v_{tre,sto,max} \cdot \frac{\left(C_{EC,glc}/K_{m,glc,tresto}\right)^{n_{glc,tresto}}}{\left(\left(C_{EC,glc}/K_{m,glc,tresto}\right)^{n_{glc,tresto}} + 1\right)} \cdot \frac{1}{\left(\left(K_{i,glc,tresto}/C_{EC,glc}\right)^{n_{i,glc,tresto}} + 1\right)} \cdot \frac{1}{\left(\left(K_{i,tre,tresto}/C_{IC,tre}\right)^{n_{i,tre,tresto}} + 1\right)} \quad \text{Eq. 3}$$

$$v_{tre,deg} = v_{tre,deg,max} \cdot \frac{\left(C_{EC,glc}/K_{m,deg,tredeg}\right)^{n_{glc,tredeg}}}{\left(\left(C_{EC,glc}/K_{m,glc,tredeg}\right)^{n_{glc,tredeg}} + 1\right)} \cdot \frac{1}{\left(\left(K_{i,glc,tredeg}/C_{EC,glc}\right)^{n_{i,glc,tredeg}} + 1\right)} \cdot \frac{\left(C_{IC,tre}/K_{tre,tredeg}\right)^{n_{tre,tredeg}}}{\left(\left(C_{IC,tre}/K_{tre,tredeg}\right)^{n_{tre,tredeg}} + 1\right)} \quad \text{Eq. 4}$$

Different from the pattern of trehalose metabolism, glycogen storage and degradation rates showed a nearly linear increase along with the specific growth rate (Figure 1B). Therefore we used the extracellular glucose concentration as the trigger, but did not apply glucose inhibition on the synthesis nor the degradation of glycogen. Glycogen synthesis is assumed to be inhibited at increasing intracellular glycogen content, while glycogen degradation is inhibited at decreasing intracellular glycogen content.

$$v_{glyc,sto} = v_{glyc,sto,max} \cdot \frac{\left(C_{EC,glc}/K_{m,glc,glycsto}\right)^{n_{glc,glycsto}}}{\left(\left(C_{EC,glc}/K_{m,glc,glycsto}\right)^{n_{glc,glycsto}} + 1\right)} \cdot \frac{1}{\left(\left(K_{i,glyc,glycsto}/C_{IC,glyc}\right)^{n_{i,glyc,glycsto}} + 1\right)} \quad \text{Eq. 5}$$

$$v_{glyc,deg} = v_{glyc,deg,max} \cdot \frac{\left(C_{EC,glc}/K_{m,glc,glycdeg}\right)^{n_{glc,glycdeg}}}{\left(\left(C_{EC,glc}/K_{m,glc,glycdeg}\right)^{n_{glc,glycdeg}} + 1\right)} \cdot \frac{\left(C_{IC,glyc}/K_{glyc,glycdeg}\right)^{n_{glyc,glycdeg}}}{\left(\left(C_{IC,glyc}/K_{glyc,glycdeg}\right)^{n_{glyc,glycdeg}} + 1\right)} \quad \text{Eq. 6}$$

Storage kinetics with gene regulation

In this scenario, the storage dynamics is not only controlled at the metabolite level but also depends on the capacities of the synthesis and degradation pathways. These capacities can be interpreted as variable enzyme levels which follow their intracellular mass balances. Based on our understanding of the storage kinetics during steady state and dynamic (feast-famine) conditions, we proposed eight different mechanisms, four at the enzyme level and four at the metabolite level to reproduce the experimental observations (Figure 3).

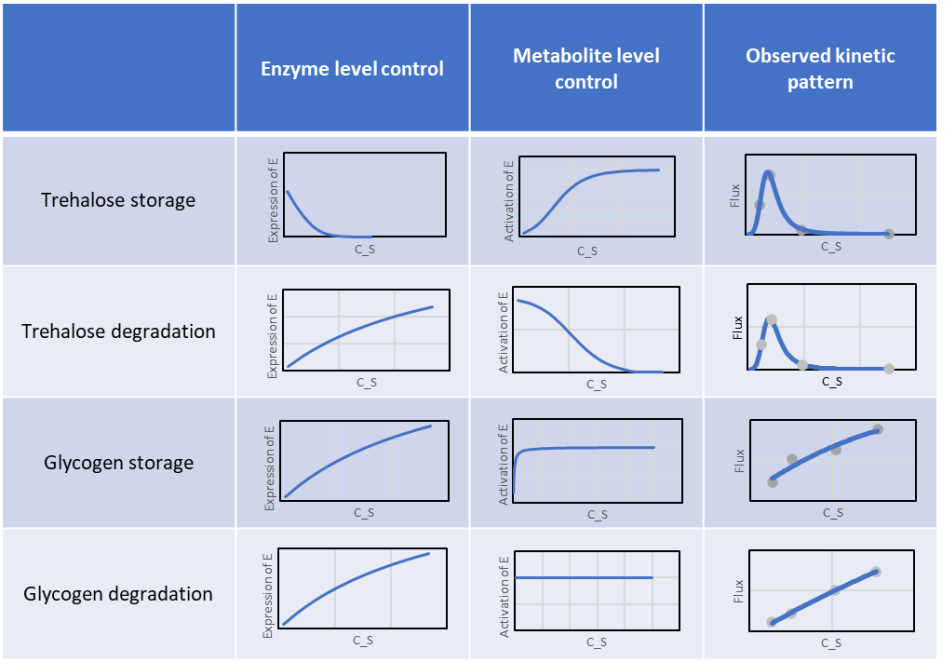

Figure 3. Proposed kinetic patterns at enzyme and metabolite level of storage and degradation processes for trehalose and glycogen, respectively.

For the synthesis of trehalose, glucose is, in general, being used as the substrate. The trehalose synthesis rate is proposed to depend on the residual glucose concentration according to saturation kinetics. Due to the observed decrease of trehalose synthesis for μ>0.1 h^{-1} (Figure 1), we infer that the flux capacity is declining as the growth rate increases. This leads to the first set of equations for the description of the trehalose synthesis rate:

$$v_{tre,sto} = v_{tps} \cdot \frac{\left(C_{EC,glc}/K_{m,glc,tresto}\right)^{n_{glc,tresto}}}{\left(C_{EC,glc}/K_{m,glc,tresto}\right)^{n_{glc,tresto}} + 1} \quad \text{Eq. 7}$$

$$\frac{d\left(v_{tps}\right)}{dt} = v_{tpssyn\_max} \cdot \frac{1}{\left(C_{EC,glc}/K_{i,glc,tpssyn}\right)^{n_{tps}} + 1} - \left(v_{growth} + k_{d,tps}\right) \cdot v_{tps} \quad \text{Eq. 8}$$

where in the enzyme formation rate is inhibited by high glucose concentrations and the enzyme will degrade at a constant rate.

For the trehalose degradation process, the overall reaction utilizes trehalose as substrate and produces two glucose molecules as product. The trehalose degradation rate is assumed to depend on the trehalose concentration inside

the cell according to saturation kinetics while it is inhibited by the product glucose. To reproduce the dynamic shape (Figure 1), a decreased enzyme capacity is assumed to be responsible for reducing the degradation rate at low growth rates. This can be interpreted as limited resources for the cell to synthesize those enzymes at low growth rates. The synthesis rate of the responsible enzyme is assumed to be linearly correlated with the growth rate.

$$v_{tre,deg} = v_{nth} \cdot \frac{1}{\left(K_{i,glc,tredeg}/C_{EC,glc}\right)^{n_{i,glc,tredeg}} + 1} \cdot \frac{C_{IC,tre}/K_{tre,tredeg}}{C_{IC,tre}/K_{tre,tredeg} + 1} \quad \text{Eq. 9}$$

$$\frac{d(v_{nth})}{dt} = v_{nthsyn\_max} \cdot v_{growth} - \left(v_{growth} + k_{d,nth}\right) \cdot v_{nth} \quad \text{Eq. 10}$$

The dynamics of glycogen storage is more straightforward: from the cell's response to a short-term glucose oscillation hardly any changes in the glycogen synthesis and degradation rates were observed during feast famine cycles (C. A. Suarez-Mendez et al., 2017), suggesting the storage rate is not largely affected by the availability of extracellular glucose. However, taking the glucose as the substrate of the storage reaction, we still use glucose as the activator of this storage process but begin with a very high affinity as initial value for the parameter estimation step. This also suggests that the relation between the glycogen storage rate and the specific growth rate is more likely to be governed by enzyme capacity changes than the metabolite level. Therefore, the enzyme synthesis rate is proposed to be linearly correlated to the growth rate.

$$v_{glyc,sto} = v_{gsy} \cdot \frac{C_{EC,glc}/K_{m,glc,glycsto}}{C_{EC,glc}/K_{m,glc,glycsto} + 1} \quad \text{Eq. 11}$$

$$\frac{d(v_{gsy})}{dt} = v_{gsysyn\_max} \cdot v_{growth} - \left(v_{growth} + k_{d,gsy}\right) \cdot v_{gsy} \quad \text{Eq. 12}$$

We propose a different metabolite level kinetic for the glycogen degradation rate compared to the other three, because we observed inconsistent cell responses in long-term steady state and short-term oscillation conditions. On one hand, under a series of steady states, the cells showed a positive correlation between the glycogen degradation rate and the growth rate and/or extracellular glucose (Canelas et al., 2011; C. A. Suarez-Mendez et al., 2016). This could be interpreted as either enzyme capacity-driven development (i.e. higher enzyme abundancy at higher growth rate, similar to the glycogen storage) or the activation by glucose (i.e. a higher glucose level stimulates the degradation process, similar as for trehalose degradation). However, the latter

hypothesis makes less sense because glucose is the product of glycogen degradation and less likely to be an activator of the degradation.

On the other hand, the glycogen degradation is reduced significantly during a short-term feast phase(C. A. Suarez-Mendez et al., 2014, 2017), that is, when the extracellular glucose concentration is high. This conflicts with the observed cellular response at high growth rates during steady state. (Wilson et al., 2010) found that UDP-glucose had clear regulatory effects on glycogen hydrolysis. From published *S.cerevisiae* metabolite data (Canelas et al., 2011; Nikerel et al., 2011; C. A. Suarez-Mendez et al., 2014, 2017), we found the profile of UDP-glucose stabilizes at around 2.5 µmol/gDW in multiple steady states, independent of the specific growth rate. Nevertheless, during the short-term glucose oscillations (C. A. Suarez-Mendez et al., 2017), the level of UDP-glucose was perturbed but quickly restored to the steady level (Figure 4). This kinetic feature fits with our hypothesis on the glycogen degradation process: the degradation rate follows the capacity increase at higher growth rate but is inhibited in a glucose perturbation. Therefore, for the kinetics of glycogen degradation, we proposed to use UDP-glucose rather than glucose as the regulatory metabolite for this process. Because glycogen is the substrate for the degradation, it also limits the degradation rate at low concentration. The enzyme formation and degradation have the same mechanisms as glycogen storage.

Because UDP-glucose is not part of the original model, and its absolute amount in the cell is negligible (<0.05% of the cell weight), we introduced a simple PID control system to mimic the UDP-glucose concentration but not include it as a part of the stoichiometry and mass balances. This semi-mechanistic control system can realize a constant level during steady state conditions and also a response to a short-term perturbation with only three parameters. The details of such PID control algorithm can be found in Supplementary information “PID based UDP-glucose modeling”.

$$v_{glyc,deg} = v_{gsh} \cdot \frac{1}{\left(C_{IC,udpg}/K_{i,udpg,glycdeg}\right)^{n_{udpg,glycdeg}} + 1} \cdot \frac{C_{IC,glyc}/K_{glyc,glycdeg}}{C_{IC,glyc}/K_{glyc,glycdeg} + 1} \quad \text{Eq. 13}$$

$$\frac{d\left(v_{gsh}\right)}{dt} = v_{gshsyn\_max} \cdot v_{growth} - \left(v_{growth} + k_{d,gsh}\right) \cdot v_{gsh} \quad \text{Eq. 14}$$

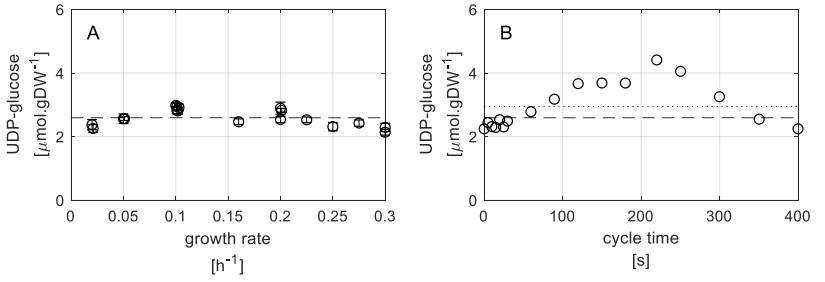

Figure 4. Reported UDP-glucose profile in steady states (A) and under glucose oscillation conditions (B). Dashed line for averaged UDP-glucose level in steady states and dotted line for average UDP-glucose level over a 400s oscillation cycle

Constant enzyme levels cannot reproduce both steady state and dynamic conditions

To obtain a complete model which is capable of simulating the full cell physiology, the growth-related parameters were first fitted to multiple steady states (Canelas et al., 2011) while all storage conversions were set to zero. This involves in total 6 parameters in the growth model. Once a reasonable fitting was obtained, we used them as the initial value for the second round of model fitting where parameters for storage conversions (4 kinetic expressions in total) were included. The steady state values of both storage compounds as well as the absolute fluxes of storage and degradation (C. A. Suarez-Mendez et al., 2016) were used for model evaluation. The fitting results of the growth-related parameters are shown in Table 2. Others are collected in the Supplementary information "Parameter estimation routine". The fitting performance is shown in Figure 5 with the experimental data.

As shown in Eq. 1, the absolute intracellular level of both storage compounds are the result of the dynamic balances between storage flux and degradation flux, which can be extremely small when the growth rate approaches zero. This makes the storage compound levels very sensitive to the predicted storage and degradation rates. Nevertheless, the simple black box model is well capable of reproducing the intracellular levels of both trehalose and glycogen, showing a satisfactory performance of this model under steady state conditions. The similar model kinetic structure for the storage and degradation steps also makes the small but distinct differences easier to mimic.

Table 2 Parameters estimated by AMIGO (eSS) for the growth of *Saccharomyces cerevisiae*

Parameter	Unit	Black box model	Gene-regulation model
V_glc_upt_max	molglc/CmolleanX/h	0.14	0.14
Km_glc_upt	mmolglc/kg	0.39	0.30
Yxs_max (alpha)	CmolleanX/molglc	3.4	3.4
Maintenance	molglc/CmolleanX/h	0.001	0.001
Yos	$molO_2$/molglc	6	6
P/O_ratio	molATP/molO	0.94	0.97

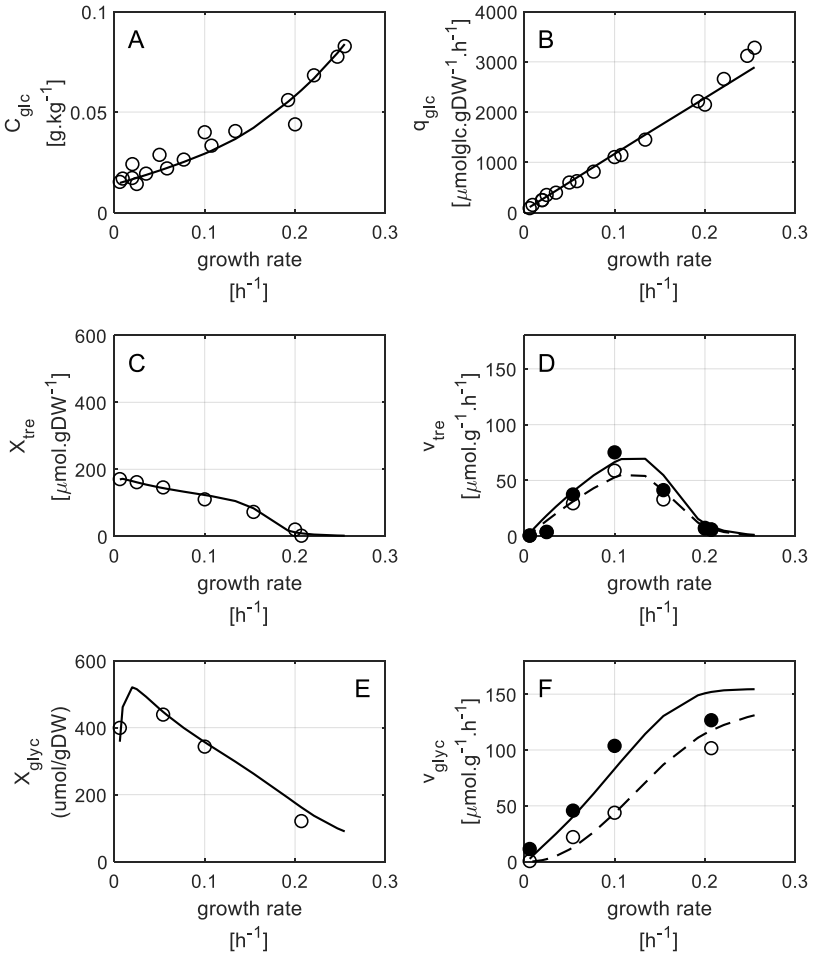

Figure 5. Performance of the black box storage model coupled with the black box core model. Circles represent experimental data and lines represent model predictions. For v_{tre} and v_{glyc}, there are solid lines, while filled circles represent storage rates; dashed lines and open circles represents degradation rates

Unfortunately, such a good agreement as obtained for the steady state experimental data is not possible for the feast-famine oscillation conditions (Figure 6). At a first glance, the absolute concentrations of trehalose and glycogen seem acceptable. Notably, although it is still overestimated, the reduction of the glycogen level compared to the level at steady state is still captured. The key problem of the prediction of this model is the description of

the storage kinetics under glucose oscillating conditions. It has been reported in both *S.cerevisiae* and *P.chrysogenum* [(L. de Jonge et al., 2014b, 2014a; C. a. Suarez-Mendez, 2015; C. A. Suarez-Mendez et al., 2016)] that the storage flux is expected to increase during the feast phase while decrease during the famine phase. Moreover, the degradation of storage carbon is observed to drop in the feast phase, whereas it is increased to a certain rate as emergency carbon source during the famine phase. The similar kinetic structures for storage and degradation (see Section "Storage kinetics as a black box") in this black box model now becomes a restriction of predicting completely different responses during feast-famine conditions.

The reason for this incapability when simulating fast dynamics is that this model structure only takes care of instant rate changes, i.e., regulation from metabolites/cofactors/signal agents. Although it was successfully applied to a wide range of steady states from 0.01 h^{-1} to 0.25 h^{-1}, it doesn't necessarily mean that this type of model is capable of correctly fitting those steady states. One key factor that was not included here is the enzyme abundancy change. This is also one of the key differences between long-term steady states and short-term oscillations: in many cases, enzyme levels will change with the growth rate (Lao-Martil et al., 2022; van den Brink et al., 2008; van Hoek et al., 1998) and will be induced or repressed as a result of changes in the cellular environment (Maier et al., 2002; Ozcan & Johnston, 1999). We conclude that changes in enzyme levels also play a key role in the cellular adaptation to glucose dynamics at different timescales. To clarify the difference of the cell's response, we compared the storage and degradation rates against different glucose uptake rates under long-term steady states and short-term oscillations. Figure 7 shows how different the correlation between storage/degradation rates and glucose uptake rate is. Taking the glycogen degradation rate as an example, neither the fold change nor the change of direction is the same under those two timescales. However, a similar comparison of model prediction showed very similar correlations under steady state and feast-famine conditions. Such manipulation of enzyme capacities is already beyond this black box model's design scope, which only uses hyperbolic curves to numerically fit the observations of multiple steady states.

In summary, it is concluded in this section that despite a good fitting performance of the proposed Michaelis-Menten based kinetic model on a wide range of steady states, it is not capable of reproducing cell responses in a short-term dynamic setting. Even more so, due to the incomplete mechanism of

storage carbohydrate regulation, the model application should be strictly limited to only steady state conditions and the absolute values of its parameters don't have a strong biological meaning, because those values implicitly cover changes in enzyme levels. It is still possible for such a type of model to reproduce short-term dynamic behavior as long as the assumption of constant enzyme levels is justified, e.g. as is the case for a single short term glucose perturbation. But the application of such model should be limited to the calibrated specific growth rate range. Extrapolation beyond this range will likely violate the assumptions during model construction and give unrealistic simulation results.

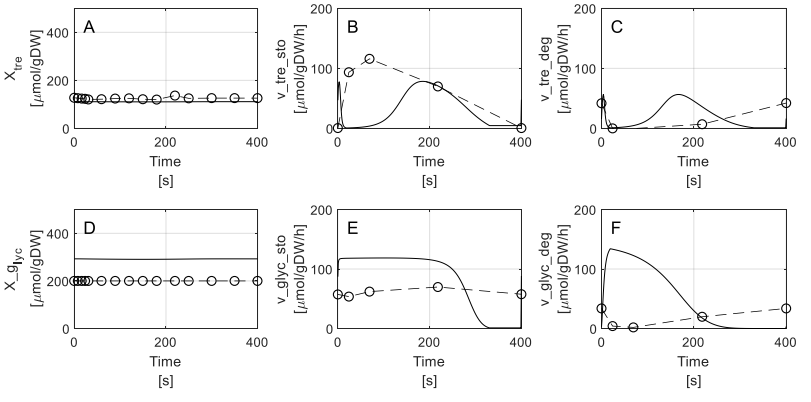

Figure 6.. Predictions of the black-box based storage kinetic model under feast-famine oscillation conditions at an average dilution rate 0.1 h^{-1}. Solid lines: the model prediction, circles with dashed lines: experimental data from (C. A. Suarez-Mendez et al., 2017)

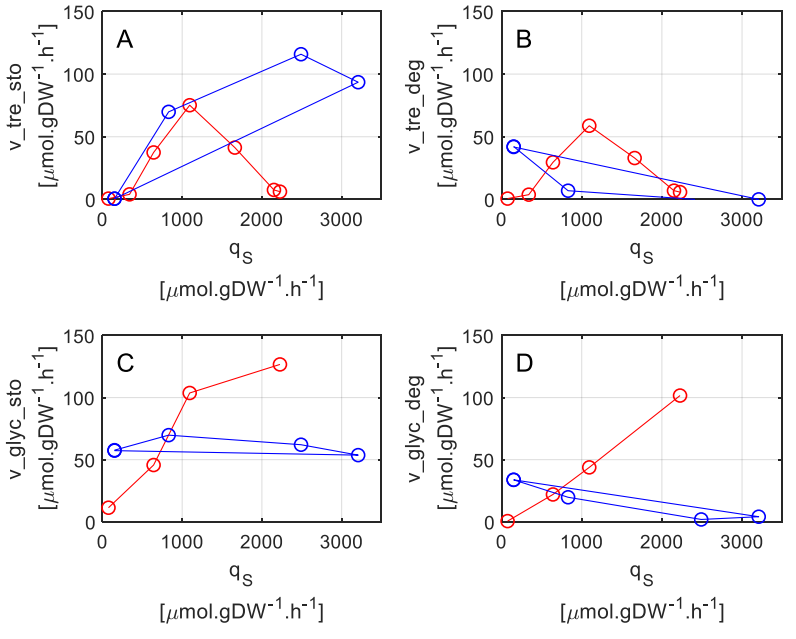

Figure 7. Storage compounds' rate correlation against glucose uptake rate Blue: oscillation (cycle); Red: steady states. This plot depicts a completely different dynamic pattern of those four steps under steady states and feast famine oscillations.

Gene-regulation model shows a good match of both steady states and oscillations

As shown in the previous section, one of the main differences between steady state and oscillation conditions is the enzyme dynamics. In steady states, the enzyme capacities of storage/degradation processes will change with the growth rate (Lao-Martil et al., 2022; van Hoek et al., 1998) and growth conditions while in short-term oscillations, slow enzyme dynamics are not likely to change within each cycle but will reach a dynamic balance after adaptation to the repeated oscillations. This is also the reason why the black box model is fundamentally not suitable for describing both long- and short-term dynamics (Douma et al., 2010; Tang et al., 2017; G. Wang, Haringa, et al., 2019).

Here the gene-regulation laws were applied separately for all four kinetics. In most of the cases, we set up a linear correlation between growth rate and the enzyme synthesis rate for simplicity purposes. The enzyme degradation rate was assumed to always follow first-order kinetics. Details on the kinetic

equations used are elaborated in the previous section “Storage kinetics with gene regulation”.

A similar parameter estimation procedure as for the black box model was followed, with as only exception the use of the initial values. The estimation of growth-related parameter starts with the final value determined for the black box model. This results in a similar but still not identical best-fit (Table 2) for the growth part. But still, a relatively satisfactory model performance on all steady states was obtained (Figure 8). It was mentioned that a good prediction of the absolute levels of both trehalose and glycogen relies on precisely predicted differences between their synthesis and degradation rates. These are significant challenges for adequate determination of the kinetics as we proposed because the rates are not only determined by different metabolites in different kinetic forms but also rather sensitive to the enzyme levels which are determined by their rates of formation and degradation. Consequently, the residual error of this gene-regulated model is larger than of the black box model. Especially for the trehalose dynamics, a slight underestimation of the maximum values is observed (Figure 8D).

The extra gain from the relatively complex gene regulation model is its capability of predicting short term dynamics under feast-famine conditions. As expected, the enzymes used in this model have long turnover times and their levels hardly show any dynamics in one oscillation cycle. The dynamics shown in Figure 9 contains only the rapid metabolite level responses caused by the external perturbation. For trehalose, the experimental data validate the predictions of both storage/degradation rates and intracellular trehalose concentrations. For the glycogen node, the storage and degradation rates were predicted with correct trends, but the predicted absolute level of glycogen was three times higher than observed experimentally (Figure 9D). As mentioned before, the absolute levels of those compounds are highly sensitive to the difference between the rates of synthesis and degradation. In the validation of the feast-famine cycles, we in fact see some overestimations of the glycogen synthesis rate (Figure 9E) which could partially explain the large overestimation of the glycogen level. By comparing the enzyme capacities in steady states and feast-famine dynamics, we noticed that the model predicts a similar flux capacity (average growth rate at 0.1 h^{-1}) for both scenarios (See supplementary information “Simulated storage enzyme capacity under steady states and oscillation condition”), which is the case for trehalose. For glycogen, the synthesis capacity decreased from 100 $\mu mol.gDW^{-1}.h^{-1}$. (Steady state, D = 0.1 h^{-}

[1]) to approximately 50 µmol.gDW^{-1}.h^{-1}(oscillation condition, average D = 0.1 h^{-1}). Due to the rapid changes in the extracellular environment, the cells may become more restricted in glycogen dynamics and have a higher overall maintenance energy demand due to possible futile cycles in storage metabolism or enzyme synthesis/degradation(L. de Jonge et al., 2014b; L. P. de Jonge et al., 2011; Lin & Neubauer, 2000; Reijenga et al., 2005). This type of cell adaptation, in some cases expressed as a higher flux distribution towards PP pathway for upregulated RNA synthesis for stress proteins(Li et al., 2018), is not covered by this gene regulation model, mainly because we used a simple black box model to describe the growth. To capture this kind of dynamics, a more detailed core growth model is required, which covers the adaption of other parts of metabolism. Nevertheless, even if we artificially reduce the capacity of storage (by reducing the synthesis rate of E_{gly}) in such a way that the predicted storage flux matches the experimental data (See supplementary information "Adapted gene regulated kinetic model simulating glucose oscillation conditions"), we still obtained an almost two times overestimated glycogen level from the model prediction. This reveals another issue which originates from either a relatively big estimation error on those rates, knowing that for glycogen no ^{13}C labelled information was available(C. A. Suarez-Mendez et al., 2014, 2017), or there are other exits from glycogen hydrolysation (Wilson et al., 2010) which were not included in the current models.

In summary, the gene-regulation model showed its capability in predicting both long-term steady states (Figure 8) and short-term oscillations (Figure 9), even though the enzyme synthesis is simply linearly correlated to the cell specific growth rate. The predicted absolute rate and compound concentrations at different steady states are the results of both enzyme balances and metabolite activation/limitation. For short term dynamics, the enzyme balance will gradually be adapted but for each oscillation cycle, only rapidly changing compounds and metabolites will trigger the enzyme kinetics, which may eventually consist of a 'contradictory' pattern against steady state observations (Figure 7).

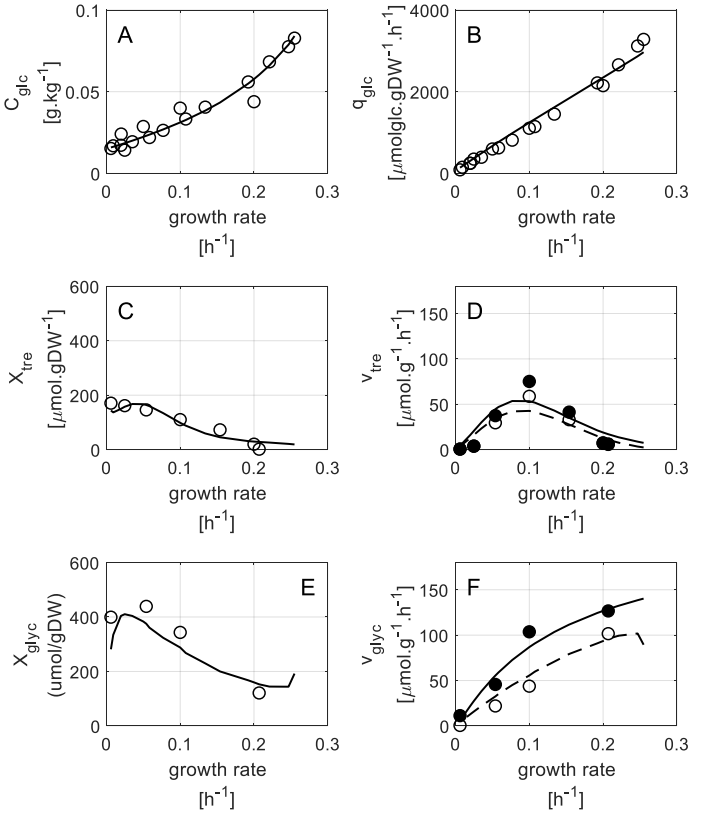

Figure 8. Performance of the gene regulation storage model coupled with the black box core model, applied to steady state data. Circles represent experimental data and lines represent model predictions. For vtre and vglyc, solid lines and filled circles represents storage rate; dashed lines and open circles represent degradation rates

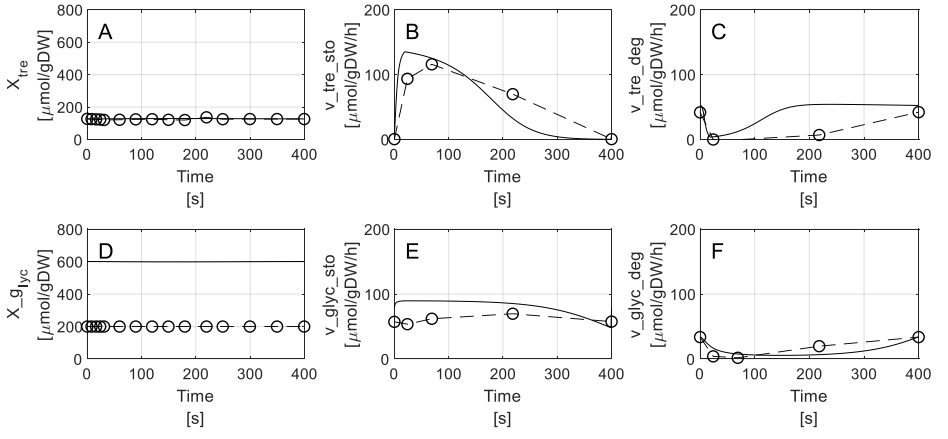

Figure 9. prediction of gene regulated storage kinetic model under feast-famine oscillation conditions at an average dilution rate 0.1 h^{-1}. Solid lines: the model prediction, circles with dashed lines: experimental data from (C. A. Suarez-Mendez et al., 2017)

Storage carbon acts as buffer under both feast and famine conditions

One of the pitfalls of the classic black box model is shown by (Tang et al., 2017) when simulating a glucose feast-famine scenario. If a constant maintenance energy parameter was assigned to the Herbert-Pirt equation, one must satisfy a glucose uptake rate that supports at least the maintenance demand which is an assigned constant value regardless the physiological conditions. When the glucose uptake rate, determined by the level of the extracellular glucose concentration, is smaller than the m_S value in the Herbert-Pirt equation, a negative growth rate will be calculated and cause issues in all other kinetics that rely on the growth rate as input. It may look like a cell-lysis rate or a short-term consumption of storage carbohydrate (which is part of the cell dry weight in such model assumption). But the interpretation of a 'lysis' rate from the Herbert-Pirt equation doesn't make sense. Because the rest of the model is not calibrated to this 'negative growth', this type of model prediction is clearly beyond the model's scope. In reality, cells do not start lysing immediately when faced with sudden starvation. (L. de Jonge et al., 2014b) has shown that trehalose, glycogen, and sometimes certain amino acids will be used as carbon source during a short period of starvation.

(Tang et al., 2017) demonstrated that including an extra storage pool and assigning proper kinetics to store/release carbon in feast-famine cycles can solve the carbon supply problem during starvation and prevent a negative growth rate being calculated. Because our gene-regulation model has a proper structure of storage allocation under feast-famine oscillations, we confirmed

that this model supplies enough carbon source into the central metabolism even during the famine phase (Figure 10). In addition, the storage carbon also helps buffering the cell from the initial 20s feast phase by re-distributing part of the intake carbon towards storage pools.

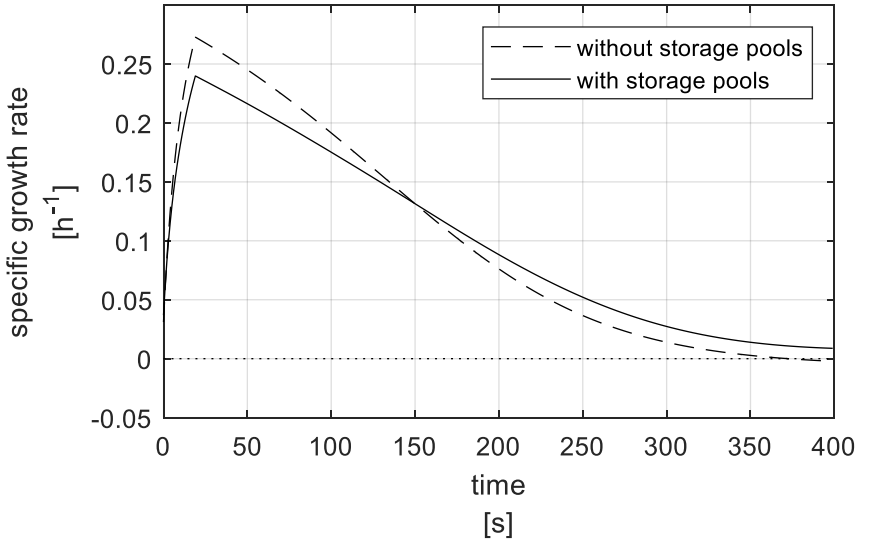

Figure 10. The predicted growth rate by the gene-regulation model with (solid line) and without (dashed line) storage kinetics in feast-famine conditions. The introduction of storage carbon first helps mitigating the pulse in the first 20 seconds by branching carbon flow from the central metabolism and supports minor "growth" in the famine phase. Negative growth predicted by the model without storage is a numerical result when simulated glucose uptake rate is smaller than the constant maintenance energy demand.

Successful integration of storage kinetic models with cell's lifelines from a 22m^3 tank

It has been suggested to execute a stability and sanity check before integrating the model to a computationally heavy and time-consuming full-scale CFD simulation. Therefore, we tested both models with multiple cell lifelines from our previous work (Haringa, Deshmukh, et al., 2017b). A lifeline consists of a series of glucose concentrations registered by Lagrangian particle tracking in the CFD simulation. The simulation scenario ensured a 0.08mM (30% of dissolved oxygen saturation at 1 bar) or higher DO level and therefore limitation of oxygen is not expected. The original lifeline lasts for 1190s and monitors the experienced glucose concentration every 0.03s. Here, we connected 120 separate original lifelines into a long, single lifeline which last for about 40 hours to enable the model to reach an unbiased dynamic state. Note that a slight adaptation of both models' glucose uptake kinetics is applied so that the

monitored glucose concentrations match the model prediction. This was done by overwriting the glucose affinity with the one used in (Haringa, Deshmukh, et al., 2017b). This small adaptation should still meet the objective to evaluate the model's stability and qualitative properties during glucose oscillations.

Figure 11 shows a snapshot of our models' responses in a short time lifeline window (720 s). Clear differences are visible between the two models, which agrees with validation of the models against feast-famine scale down conditions. More specifically, a proper storage profile of trehalose was observed with the gene-regulated model and the release of trehalose at low glucose concentrations (Figure 6B and Figure 9B). The black box model, however, showed an incorrect trend which tends to release carbon when glucose is abundant outside the cell. Similarly, we see a relatively dynamic glycogen response with this black box model, which is also aligned with what we observed previously (Figure 6 and Figure 9).

A key concept of our gene regulation model is the enzyme pool that responds to long-term dynamics while metabolite level regulations are responsible for short-term dynamics. Therefore, in the 20-hour lifelines, despite the highly dynamic rates of trehalose and glycogen storage/release, we indeed see more stable profiles for those enzymes (Figure 12) lasting for 20 hours. It is also important to notice that for each enzyme, the capacity of the 20 lifelines is likely to converge to a unique solution. This also perfectly matches the expectation as all lifelines represent the cell experiences in the same $22m^3$.

In summary, the stability and sanity checks were successfully passed by both kinetic models. The models remain mathematically stable during highly dynamic glucose concentrations and both models maintained their kinetic features under these conditions.

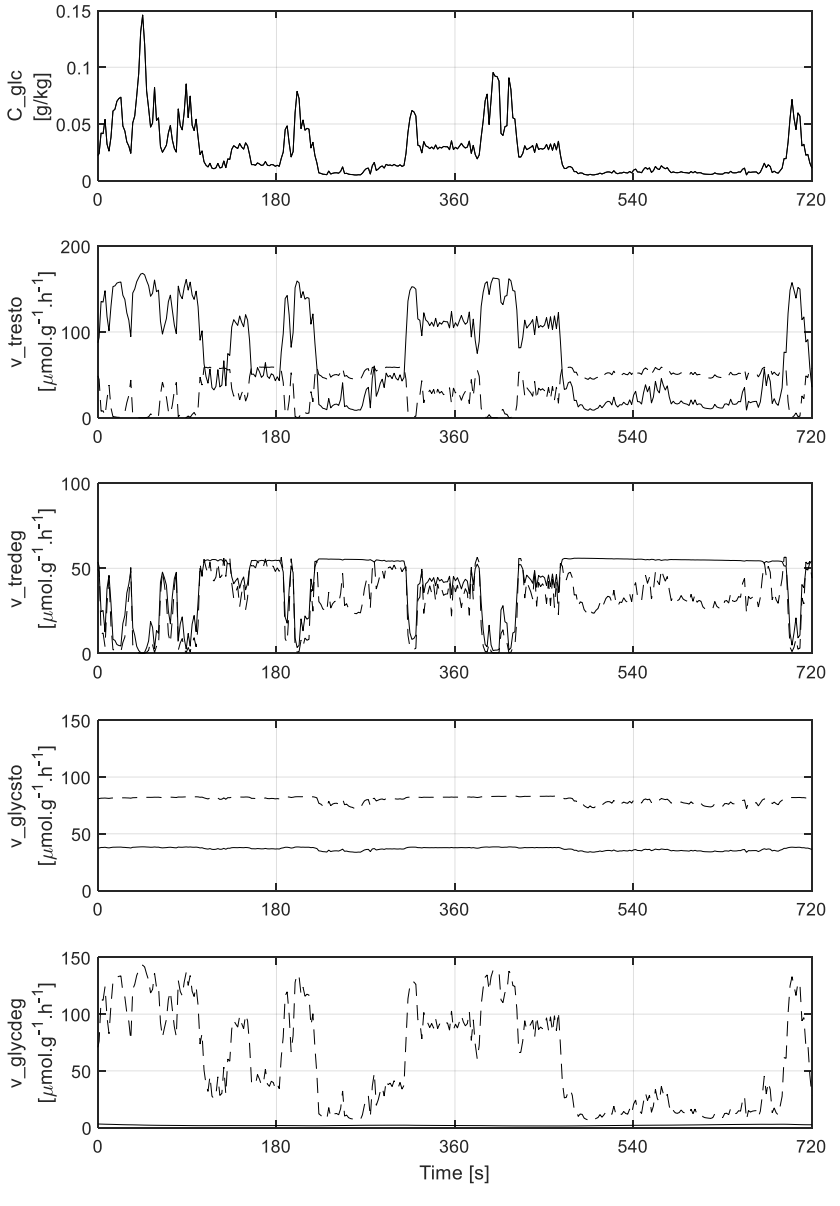

Figure 11 The responses of storage carbon kinetics in one cell lifeline extracted from a 22m³ pilot-scale fermentation. Only the extracellular glucose concentration (C_glc) was provided in the lifeline. Solid: prediction of gene regulated storage kinetic model; dashed: prediction of black box storage kinetic model

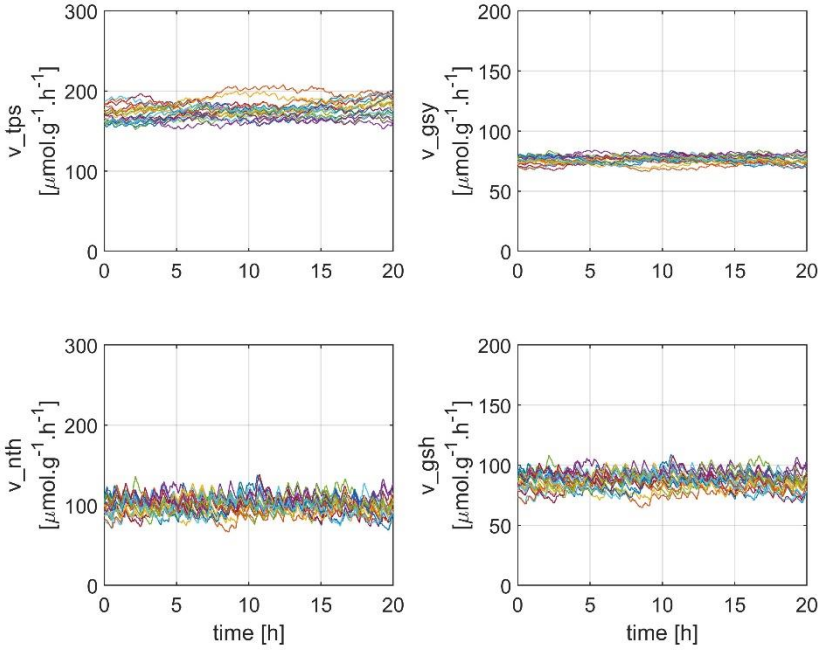

Figure 12 The predicted profiles of 4 enzyme pools with 20 lifelines. Each lifeline lasts for 20 hours, and an unbiased dynamic state is assumed to be reached for all 20 cases. The enzyme pool size is expressed as the capacity of conversion.

Conclusions and outlook

In this Chapter, we have presented two different types of storage carbohydrate kinetics that were coupled with the same black box growth model for *S. cerevisiae*. Trehalose and glycogen were studied separately as storage compounds due to their completely different dynamic features. One storage kinetic variant is a complete black box, using conventional Michaelis-Menten type kinetics. This model appeared well capable of reproducing cell growth, glucose uptake, storage content and storage fluxes under steady state conditions. However, this model was not capable of predicting cellular behavior during short term dynamic conditions such as feast-famine oscillations due to structural limitations. We therefore proposed a structured gene-regulation model to describe the same observations, where the dynamics of the synthesis and degradation of storage compounds was regulated at enzyme and metabolite level at the same time. By introducing simple linear enzyme kinetics, the gene-regulation model can already precisely reproduce all steady states and qualitatively also the feast-famine scenario. The properly modeled storage kinetics enables the black box core model to be able to simulate carbon supply oscillations without problems of negative growth. Finally, the gene-regulated model was tested against cell lifelines that were obtained previously in a 22 m^3 pilot scale reactor. This model is proved to be stable under such highly dynamic conditions and maintains its kinetic features in dealing with short- and long-term dynamics. These properties of the gene-regulation model make it suitable for use as particle ('parcel') kinetics in a complete full-scale Euler-Lagrange CRD-CFD simulation (Haringa, Deshmukh, et al., 2017a; Haringa, Noorman, et al., 2017; Haringa, Tang, et al., 2017; G. Wang, Haringa, et al., 2019) and/or compartment model for industrial fermentation (Haringa et al., 2022).

Nevertheless, the gene-regulation model can still be improved in several aspects:

- The use of 'ghost' species UDP-glucose is not favored, and real measurable UDP-glucose should be included as a part of the model in a future version.
- The remaining part of the model is still black box, which limits the usage of this model. With a more structured version, albeit with a relatively simple core model like the 9-pool kinetic model applied to *P. chrysogenum* (Tang et al., 2017), the model will gain in predictive capacity.

- We still see discrepancy in the prediction of the glycogen concentrations. The capacity of glycogen synthesis seems to be regulated by yet other compounds than glucose and the key synthesis/degradation enzymes, which is beyond the current model scope.

References

Aboka, F. O., Heijnen, J. J., & van Winden, W. A. (2009). Dynamic 13C-tracer study of storage carbohydrate pools in aerobic glucose-limited Saccharomyces cerevisiae confirms a rapid steady-state turnover and fast mobilization during a modest stepup in the glucose uptake rate. *FEMS Yeast Res*, *9*(2), 191–201. https://doi.org/10.1111/j.1567-1364.2008.00465.x

Aboka, F. O., van Winden, W. A., Reginald, M. M., van Gulik, W. M., van de Berg, M., Oudshoorn, A., & Heijnen, J. J. (2012). Identification of informative metabolic responses using a minibioreactor: a small step change in the glucose supply rate creates a large metabolic response in Saccharomyces cerevisiae. *Yeast*, *29*(3–4), 95–110. https://doi.org/doi:10.1002/yea.2892

Canelas, A. B., Ras, C., ten Pierick, A., van Gulik, W. M., & Heijnen, J. J. (2011). An in vivo data-driven framework for classification and quantification of enzyme kinetics and determination of apparent thermodynamic data. *Metabolic Engineering*, *13*(3), 294–306. https://doi.org/https://doi.org/10.1016/j.ymben.2011.02.005

Choi, K. R., Kim, W. J., & Lee, S. Y. (2018). Metabolomics for industrial fermentation. *Bioprocess and Biosystems Engineering*, *41*(7), 1073–1077. https://doi.org/10.1007/s00449-018-1967-3

de Jonge, L., Buijs, N. A. A., Heijnen, J. J., van Gulik, W. M., Abate, A., & Wahl, S. A. (2014a). Flux response of glycolysis and storage metabolism during rapid feast/famine conditions in Penicillium chrysogenum using dynamic 13C labeling. *Biotechnology Journal*, *9*(3), 372–385.

de Jonge, L., Buijs, N. A. A., Heijnen, J. J., van Gulik, W. M., Abate, A., & Wahl, S. A. (2014b). Flux response of glycolysis and storage metabolism during rapid feast/famine conditions in Penicillium chrysogenum using dynamic 13C labeling. *Biotechnology Journal*, *9*(3), 372–385. https://doi.org/10.1002/biot.201200260

de Jonge, L. P., Buijs, N. A. A., ten Pierick, A., Deshmukh, A., Zhao, Z., Kiel, J. A. K. W., Heijnen, J. J., & van Gulik, W. M. (2011). Scale-down of penicillin production in Penicillium chrysogenum. *Biotechnology Journal*, *6*(8), 944–958.

Delvigne, F., Takors, R., Mudde, R., van Gulik, W., & Noorman, H. (2017). Bioprocess scale-up/down as integrative enabling technology: from fluid mechanics to systems biology and beyond. *Microbial Biotechnology*, *10*(5), 1267–1274. https://doi.org/10.1111/1751-7915.12803

Douma, R. D., Verheijen, P. J. T., de Laat, W. T. A. M., Heijnen, J. J., & van Gulik, W. M. (2010). Dynamic gene expression regulation model for growth and penicillin production in Penicillium chrysogenum. *Biotechnology and Bioengineering*, *106*(4), 608–618.

Gancedo, C., & Flores, C.-L. (2004). The importance of a functional trehalose biosynthetic pathway for the life of yeasts and fungi. *FEMS Yeast Research*, *4*(4–5), 351–359. https://doi.org/10.1016/S1567-1356(03)00222-8

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017a). Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669.

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017b). Euler-Lagrange analysis towards representative down-scaling of a 22m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669. https://doi.org/https://doi.org/10.1016/j.ces.2017.01.014

Haringa, C., Noorman, H. J., & Mudde, R. F. (2017). Lagrangian modeling of hydrodynamic–kinetic interactions in (bio) chemical reactors: practical implementation and setup guidelines. *Chemical Engineering Science*, *157*, 159–168.

Haringa, C., Tang, W., & Noorman, H. J. (2022). Stochastic parcel tracking in an Euler-Lagrange compartment model for fast simulation of fermentation processes. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.28094

Haringa, C., Tang, W., Wang, G., Deshmukh, A. T., van Winden, W. A., Chu, J., van Gulik, W. M., Heijnen, J. J., Mudde, R. F., & Noorman, H. J. (2017). Computational fluid dynamics simulation of an industrial P. chrysogenum fermentation with a coupled 9-pool metabolic model: towards rational scale-down and

design optimization. *Chemical Engineering Science*. https://doi.org/https://doi.org/10.1016/j.ces.2017.09.020

Herrgård, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., Büthgen, N., Borger, S., Costenoble, R., Heinemann, M., Hucka, M., le Novère, N., Li, P., Liebermeister, W., Mo, M. L., Oliveira, A. P., Petranovic, D., Pettifer, S., Simeonidis, E., … Kell, D. B. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnology*, *26*(10), 1155–1160. https://doi.org/10.1038/nbt1492

Kesten, D., Kummer, U., Sahle, S., & Hübner, K. (2015). A new model for the aerobic metabolism of yeast allows the detailed analysis of the metabolic regulation during glucose pulse. *Biophysical Chemistry*, *206*, 40–57. https://doi.org/https://doi.org/10.1016/j.bpc.2015.06.010

Lao-Martil, D., Verhagen, K., Schmitz, J., Teusink, B., Wahl, S., & van Riel, N. (2022). Kinetic Modeling of Saccharomyces cerevisiae Central Carbon Metabolism: Achievements, Limitations, and Opportunities. *Metabolites*, *12*(1), 74. https://doi.org/10.3390/metabo12010074

Lapin, A., Müller, D., & Reuss, M. (2004). Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler- Lagrange methods: Traveling along the lifelines of single cells. *Industrial & Engineering Chemistry Research*, *43*(16), 4647–4656.

Li, C., Shu, W., Wang, S., Liu, P., Zhuang, Y., Zhang, S., & Xia, J. (2018). Dynamic metabolic response of Aspergillus niger to glucose perturbation: evidence of regulatory mechanism for reduced glucoamylase production. *Journal of Biotechnology*, *287*, 28–40. https://doi.org/https://doi.org/10.1016/j.jbiotec.2018.08.005

Lin, H. Y., & Neubauer, P. (2000). Influence of controlled glucose oscillations on a fed-batch process of recombinant Escherichia coli. In *Journal of Biotechnology* (Vol. 79). www.elsevier.com/locate/jbiotec

Maier, A., Völker, B., Boles, E., & Fuhrmann, G. F. (2002). Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Res*, *2*(4), 539–550. https://doi.org/10.1111/j.1567-1364.2002.tb00121.x

Nadal-Rey, G., McClure, D. D., Kavanagh, J. M., Cornelissen, S., Fletcher, D. F., & Gernaey, K. v. (2021). Understanding gradients in industrial bioreactors. *Biotechnology Advances*, *46*, 107660.

Nikerel, I. E., Canelas, A. B., Jol, S. J., Verheijen, P. J. T., & Heijnen, J. J. (2011). Construction of kinetic models for metabolic reaction networks: Lessons learned in analysing short-term stimulus response data. *Mathematical and Computer Modelling of Dynamical Systems*, *17*(3), 243–260. https://doi.org/10.1080/13873954.2010.548167

Noorman, H. (2011). An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnology Journal*, *6*(8), 934–943. https://doi.org/doi:10.1002/biot.201000406

Ozcan, S., & Johnston, M. (1999). Function and regulation of yeast hexose transporters. *Microbiology and Molecular Biology Reviews : MMBR*, *63*(3), 554–569. https://pubmed.ncbi.nlm.nih.gov/10477308

Reijenga, K. A., Bakker, B. M., van der Weijden, C. C., & Westerhoff, H. v. (2005). Training of yeast cell dynamics. *FEBS Journal*, *272*(7), 1616–1624. https://doi.org/10.1111/j.1742-4658.2005.04582.x

Silljé, H. H. W., Paalman, † J W G, Schure, E. G. ter, Olsthoorn, S. Q. B., Verkleij, A. J., Boonstra, J., & Verrips, C. T. (1999). Function of Trehalose and Glycogen in Cell Cycle Progression and Cell Viability in Saccharomyces cerevisiae. In *JOURNAL OF BACTERIOLOGY* (Vol. 181, Issue 2). https://journals.asm.org/journal/jb

Smallbone, K., Messiha, H. L., Carroll, K. M., Winder, C. L., Malys, N., Dunn, W. B., Murabito, E., Swainston, N., Dada, J. O., Khan, F., Pir, P., Simeonidis, E., Spasić, I., Wishart, J., Weichart, D., Hayes, N. W., Jameson, D., Broomhead, D. S., Oliver, S. G., … Mendes, P. (2013). A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. *FEBS Letters*, *587*(17), 2832–2841. https://doi.org/10.1016/j.febslet.2013.06.043

Suarez-Mendez, C. a. (2015). *Dynamics of Storage Carbohydrates Metabolism in Saccharomyces cerevisiae: A Quantitative* (Issue december). Faculty of Applied Sciences, Delft University of Technology.

Suarez-Mendez, C. A., Hanemaaijer, M., ten Pierick, A., Wolters, J. C., Heijnen, J. J., & Wahl, S. A. (2016). Interaction of storage carbohydrates and other cyclic fluxes with central metabolism: A quantitative approach by non-stationary 13C metabolic flux analysis. *Metabolic Engineering Communications*, *3*, 52–63. https://doi.org/https://doi.org/10.1016/j.meteno.2016.01.001

Suarez-Mendez, C. A., Ras, C., & Wahl, S. A. (2017). Metabolic adjustment upon repetitive substrate perturbations using dynamic 13C-tracing in yeast. *Microbial Cell Factories*, *16*(1), 161. https://doi.org/10.1186/s12934-017-0778-6

Suarez-Mendez, C. A., Sousa, A., Heijnen, J. J., & Wahl, A. (2014). Fast "Feast/Famine" Cycles for Studying Microbial Physiology Under Dynamic Conditions: A Case Study with Saccharomyces cerevisiae. *Metabolites*, *4*(2), 347. http://www.mdpi.com/2218-1989/4/2/347

Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J. J., Xia, J., Chu, J., & others. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by Penicillium chrysogenum. *Biotechnology and Bioengineering*, *114*(8), 1733–1743.

van den Brink, J., Canelas, A. B., van Gulik, W. M., Pronk, J. T., Heijnen, J. J., de Winde, J. H., & Daran-Lapujade, P. (2008). Dynamics of Glycolytic Regulation during Adaptation of Saccharomyces cerevisiae to Fermentative Metabolism. *Applied and Environmental Microbiology*, *74*(18), 5710–5723. https://doi.org/10.1128/aem.01121-08

van Heerden, J. H., Wortel, M. T., Bruggeman, F. J., Heijnen, J. J., Bollen, Y. J. M., Planqué, R., Hulshof, J., O'Toole, T. G., Wahl, S. A., & Teusink, B. (2014). Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science*, *343*(6174), 1245114.

van Hoek, P., van Dijken, J. P., & Pronk, J. T. (1998). Effect of Specific Growth Rate on Fermentative Capacity of Baker's Yeast. *Applied and Environmental Microbiology*, *64*(11), 4226–4233. https://doi.org/10.1128/aem.64.11.4226-4233.1998

Wang, G., Chu, J., Noorman, H., Xia, J., Tang, W., Zhuang, Y., & Zhang, S. (2014). Prelude to rational scale-up of penicillin production: a scale-down study. *Applied Microbiology and Biotechnology*, *98*(6), 2359–2369.

Wang, G., Haringa, C., Tang, W., Noorman, H., Chu, J., Zhuang, Y., & Zhang, S. (2019). Coupled Metabolic-hydrodynamic Modeling Enabling Rational Scale-up of Industrial Bioprocesses. *Biotechnology and Bioengineering*, *n/a*(n/a). https://doi.org/10.1002/bit.27243

Wang, G., Zhao, J., Wang, X., Wang, T., Zhuang, Y., Chu, J., Zhang, S., & Noorman, H. J. (2019). Quantitative metabolomics and metabolic flux analysis reveal impact of altered trehalose metabolism on metabolic phenotypes of Penicillium chrysogenum in aerobic glucose-limited chemostats. *Biochemical Engineering Journal*, *146*, 41–51. https://doi.org/https://doi.org/10.1016/j.bej.2019.03.006

Wang, X., Zhao, J., Xia, J., Wang, G., Chu, J., & Zhuang, Y. (2021). Impact of altered trehalose metabolism on physiological response of penicillium chrysogenum chemostat cultures during industrially relevant rapid feast/famine conditions. *Processes*, *9*(1), 1–16. https://doi.org/10.3390/pr9010118

Wilson, W. A., Roach, P. J., Montero, M., Baroja-Fernández, E., Muñoz, F. J., Eydallin, G., Viale, A. M., & Pozueta-Romero, J. (2010). Regulation of glycogen metabolism in yeast and bacteria. *FEMS Microbiology Reviews*, *34*(6), 952–985. https://doi.org/10.1111/j.1574-6976.2010.00220.x

Supplementary information

PID based UDP-glucose modeling

The UDP-glucose concentration follows a dynamic balance via a production rate (v_{in}) and a release rate (v_{out}):

$$v_{in} = v_{in,max} \cdot \frac{C_S}{C_s + K_S} \quad \text{Eq S1}$$

$$v_{out} = K_P \cdot e + K_I \cdot \int e + K_D \cdot \frac{de}{dt} \quad \text{Eq S2}$$

where $e = X_{UDP-glc} - X_{UDP-glc,stst}$, K_P, K_I and K_D are three PID parameters that determine the response of the release rate to a pulse in v_{in}.

We propose a simple a hyperbolic kinetic function for the formation rate of UDP-glucose, based on the availability of glucose. The utilization rate depends on the difference (e) of the current UDP-glucose level and its steady state value (before the critical growth rate, 2.5 µmol/gDW).

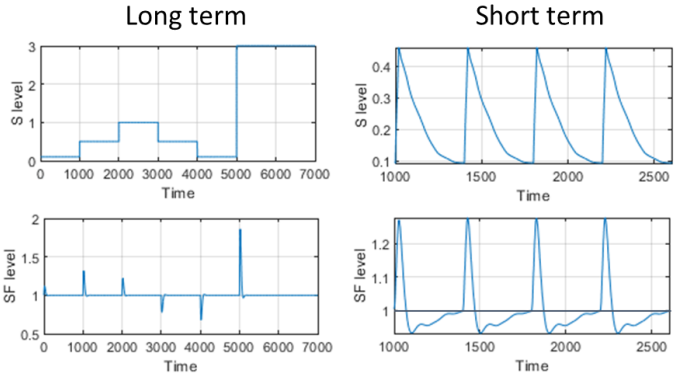

Figure S1. Demonstration of the PID based system response during long- and short-term perturbations. The SF(target) level is kept at the steady level (1) in different long-term shifts in S(substrate) and shows a dynamic response to periodic short-term perturbations. In either case, the target level converges to the given steady value.

In the application of this PID system, the v_{out} needs to be adapted as obtaining the differential term of the error, de/dt, is not easy. Fortunately, in this specific case, the PID output can be algebraically solved based on given system rules: in flux rate law and mass balance. We first worked out the original UDP-glucose balancing system:

$$v_{udpg\_in} = v_{udpg\_max} \cdot \frac{C_{EC\_glc}}{C_{EC\_glc} + K_{glc\_udpgin}} \quad \text{Eq S3}$$

$$u = K_P \cdot \left(C_{IC,udpg} - C_{IC,udpg,set}\right) + K_I \cdot Int_{err} + K_D \cdot \frac{d\left(C_{IC,udpg} - C_{IC,udpg,set}\right)}{dt} \quad \text{Eq S4}$$

$$v_{udpg\_out} = u \cdot \frac{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}}}{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}} + 1} \quad \text{Eq S5}$$

Note that Eq S5 prevents negative $C_{IC,udpg}$ in the system. Similar to other metabolic species, the UDP-glucose also follows the mass balance:

$$\frac{d\left(C_{IC,udpg}\right)}{dt} = v_{updgin} - v_{updgout} \quad \text{Eq S6}$$

And the integration of the error can be calculated in this ODE system via:

$$\frac{d\left(Int_{err}\right)}{dt} = C_{IC,udpg} - C_{IC,udpg,set} \quad \text{Eq S7}$$

From Eq S4 and Eq S6, one deduces:

$$u = K_P \cdot \left(C_{IC,udpg} - C_{IC,udpg,set}\right) + K_I \cdot Int_{err} + K_D \cdot \left(v_{updgin} - v_{updgout}\right)$$

Combining Eq S5 gives:

$$u = K_P \cdot \left(C_{IC,udpg} - C_{IC,udpg,set}\right) + K_I \cdot Int_{err} + K_D \cdot \left(v_{updgin} - u \cdot \frac{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}}}{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}} + 1}\right)$$

After reorganizing, one will get

$$u = \frac{K_P \cdot \left(C_{IC,udpg} - C_{IC,udpg,set}\right) + K_I \cdot Int_{err} + K_D \cdot v_{updgin}}{1 + K_D \cdot \frac{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}}}{\left(C_{IC,udpg}/K_{udpg,udpgout}\right)^{n_{udpgout}} + 1}} \quad \text{Eq S8}$$

For Eq S8, all right-hand terms can be obtained in the ODE system therefore the Eq S4 was replaced by Eq S8 in our gene regulated model.

The parameters used for this part of model are shown in Table S1. The parameters of this part of the model are mainly manually assigned due to lack of experimental data. The value of K_P, K_I and K_D were obtained via parameter scanning together with the complete gene-regulated model, with a range of 0-500.

Table S1 Parameters determined for UDP-glucose PID system

Parameter	Unit	Value
v_{udpg_max}	moludpg/CmolleanX/h	1
K_{glc_udpgin}	mmolglc/kg	0.1
K_p	-	100
K_i	-	15
K_d	-	55
$K_{udpg,udpgout}$	molupdg/CmolX	0.0065
$n_{udpgout}$		4
$C_{IC,udpg,set}$	molupdg/CmolX	0.0065

Parameter estimation routine

Table S2 Parameter list for the black box based storage model

Parameter name	Unit	Value	Remark
$v_{tre,sto,max}$	mol/CmolleanX/h	0.0036	
$K_{m,glc,tresto}$	mmol_glc/kg(broth)	0.021	
$n_{glc,tresto}$	-	1	
$K_{i,glc,tresto}$	mmol_glc/kg(broth)	0.015	Trehalose storage
$n_{i,glc,tresto}$	-	5	
$K_{i,tre,tresto}$	µmol_tre/gDW	138	
$n_{i,tre,tresto}$	-	5	
$v_{tre,deg,max}$	mol/CmolleanX/h	0.0049	
$K_{m,glc,tredeg}$	mmol_glc/kg(broth)	0.19	
$n_{glc,tredeg}$	-	1	
$K_{i,glc,tredeg}$	mmol_glc/kg(broth)	0.16	Trehalose degradation
$n_{i,glc,tredeg}$	-	5	
$K_{tre,tredeg}$	µmol_tre/gDW	0.98	
$n_{tre,tredeg}$	-	1	
$v_{glyc,sto,max}$	mol/CmolleanX/h	0.0041	
$K_{m,glc,glycsto}$	mmol_glc/kg(broth)	0.0107	
$n_{glc,glycsto}$	-	2	Glycogen storage
$K_{i,glyc,glycsto}$	µmol_glyc/gDW	370	
$n_{i,glyc,glycsto}$	-	5	
$v_{glyc,deg,max}$	mol/CmolleanX/h	0.0039	
$K_{m,glc,glycdeg}$	mmol_glc/kg(broth)	0.128	
$n_{glc,glycdeg}$	-	2	Glycogen degradation
$K_{glyc,glycdeg}$	µmol_glyc/gDW	1.04	
$n_{glyc,glycdeg}$	-	1	

Table S3 Parameter list for the gene regulated storage kinetic model

Parameter name	Unit	Value	Remark
$K_{m,glc,tresto}$	mmolglc/kg	0.12	
$n_{glc,tresto}$	-	2	
v_{tpssyn_max}	mol_tre/CmolleanX/h	0.00080	Trehalose storage
$K_{i,glc,tpssyn}$	mmolglc/kg	0.11	
n_{tps}	-	2	
$k_{d,tps}$	1/h	0.022	
$K_{i,glc,tredeg}$	mmolglc/kg	0.23	
$n_{i,glc,tredeg}$	-	4	
$K_{tre,tredeg}$	mmol_tre/CmolX	0.0054	Trehalose degradation
v_{nthsyn_max}	mol_tre/CmolleanX	0.03	
$k_{d,nth}$	1/h	0.78	
$K_{m,glc,glycsto}$	mmolglc/kg	0.0042	
v_{gsysyn_max}	mol_glyc/CmolleanX	0.0059	Glycogen storage
$k_{d,gsy}$	1/h	0.145	
$K_{i,udpg,glycdeg}$	mol_udpg/CmolX	0.0065	
$n_{udpg,glycdeg}$	-	7	
$K_{glyc,glycdeg}$	mmol_glyc/CmolX	10^{-7}*	Glycogen degradation
v_{gshsyn_max}	mol_glyc/CmolleanX	0.016	
$k_{d,gsh}$	1/h	0.45	

* Arbitrary number to prevent glycogen degradation when glycogen was depleted

Simulated storage enzyme capacity under steady states and oscillation condition

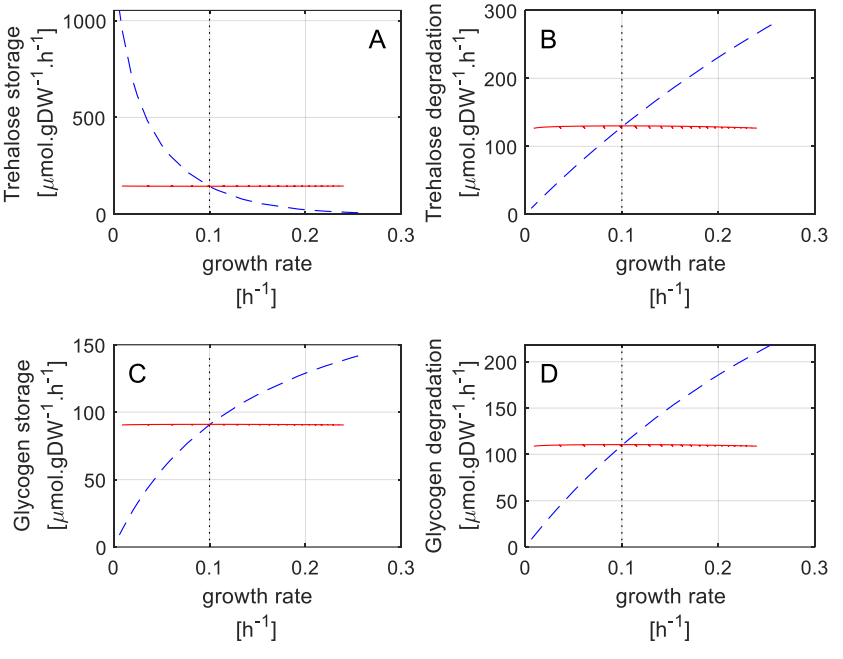

Figure S2 Four storage enzyme's capacity change during steady states (dashed blue line) and glucose oscillation condition (solid red line). The oscillation condition was executed under an average dilution rate of 0.1 h^{-1}. It is observed that enzyme capacity under oscillation condition doesn't deviate from the capacity of enzymes in steady states under the same dilution rate. This may not be true in real because of cell's global regulation under oscillating conditions. In our current gene-regulated model, the enzyme formation is linearly correlated to growth rate and the central growth model is a simple black box model. To properly calibrate the enzyme capacity of different condition but the same average dilution rate, further experiment and in-vivo enzyme capacity need to be determined.

Adapted gene regulated kinetic model simulating glucose oscillation conditions

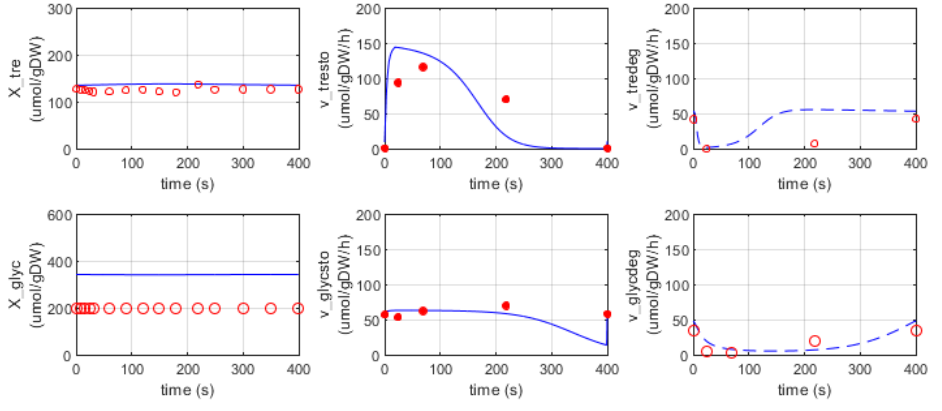

Figure S3 Adapted gene-regulation model (reduced v_gsy_syn) predicting oscillation conditions. Despite the proper estimation on glycogen storage and degradation rates, the glycogen content is still being overestimated by almost one-fold. This suggests either the estimation of previous glycogen fluxes have a large uncertainty (note that glycogen does not have label information but purely interfered by highly lumped stoichiometry model) or both models missed other outlets of glycogen metabolism.

Chapter 5

Towards a complete yet compact cell model for *Saccharomyces cerevisiae*

Abstract

Saccharomyces cerevisiae is a widely used model fungus in the biotechnology industry, due to its well-studied strain background, robustness to harsh environmental conditions, and ease of strain engineering. Transferring the strain from lab to industrial scale is always challenging due to the so-called scale-up effect, that is, sub-optimal performance, mainly related to rate-limiting mixing, mass and heat transfer rates. The integration of computational fluid dynamics (CFD) based reactor models with kinetic metabolic models for cellular behavior is one of the most comprehensive approaches to tackle this issue at early phases of strain and process development. For this, the kinetic models should be built with the 'end in mind', i.e., fit for use under dynamic conditions prevalent in the industrial bioreactor. Here, we first present a 7-pool kinetic model for yeast metabolism, validated against a wide range of published steady state data. The NAD^+/NADH ratio is described explicitly in this model and used as the trigger for ethanol production under glucose excess conditions (Crabtree effect) and oxygen limitation (Pasteur effect). With the proper built-in mechanism of ethanol formation, this model, with modifications on glucose uptake kinetics, can quantitatively capture the cellular responses to a single glucose pulse. When testing the model against data obtained from feast-famine cycles, the model showed discrepancies in some absolute intracellular compounds but is still able to capture the cellular responses qualitatively. These validations indicate that the current model is already able to display key physiological features of this yeast in industrial applications, but further improvements in model structure and additional complexity is needed for a more accurate description. In the last section, we further tested the model with simulated yeast lifelines from a 22 m^3 large pilot bioreactor. The lifelines were featured by highly dynamic glucose concentrations observed by the cells. Unlike the previously published 9-pool *Penicillium chrysogenum* model, we did not experience any computation stability issues on stiff nodes and the simulation was successfully executed for 40-hour long lifelines. In conclusion, 1) we demonstrated the capability of a simple structured kinetic model predicting *Saccharomyces cerevisiae* physiology at various dilution rates, including those beyond the critical growth rate where overflow mechanism was triggered; 2) the model is capable of predicting the ethanol production in a single glucose pulse event; 3) the model can qualitatively reproduce cellular responses in a long-term glucose oscillation scenario; 4) this model is stable enough to be integrated with highly dynamic lifelines from the large pilot bioreactor. To

further improve the accuracy of this model's prediction, we proposed an extension of the model in the Outlook Chapter of this thesis.

Keywords: lumped kinetic model, metabolic network, Crabtree effect, Pasteur effect, *Saccharomyces cerevisiae*

Introduction

Saccharomyces cerevisiae is the best-known yeast species and has been applied to produce fermented products for more than 10,000 years. More recently it also has become a very popular eukaryotic species for scientific research and commercial deployment in the biotechnology industry (Otero & Nielsen, 2010; Parapouli et al., 2020). The ease of introducing novel pathways and optimization of native cellular processes by metabolic engineering are rapidly expanding its range of applications as a cell factory. For example, *S. cerevisiae* has been successfully engineered for the production of bioethanol, succinic acid, farnesene, insulin and aroma compounds (Buijs et al., 2013; Nielsen et al., 2013; Parapouli et al., 2020). To successfully implement such bioprocesses in industry, a proper scale-up approach is required.

In addition to its versatility as production host, *S. cerevisiae* is well capable of growing on various carbon sources and is able to grow both aerobically and anaerobically (Verduyn et al., 1990). Further it has the characteristic feature of fermenting glucose to ethanol under aerobic conditions at elevated glucose concentrations (Crabtree effect) or oxygen-limiting conditions (Pasteur effect). These effects have been explained by various mechanisms, e.g. via an upper limit of ATP production through oxidative phosphorylation possibly due to spatial constraints for membrane proteins involved in oxidative phosphorylation (Pfeiffer & Morley, 2014; Schumacher, 2018), key intermediates' cross membrane shuttling (Xie et al., 2022), limited capacity of NADH recycling due to enzyme constraints (Hackett et al., 2016) or an upper limit in Gibbs free energy dissipation (Niebel et al., 2019). From an evolutionary perspective aerobic ethanol production can be an advantage as it provides a way to compete with other microorganisms in glucose-rich environments. Furthermore, successful efforts have been made to create a Crabtree-negative *Saccharomyces cerevisiae* via rational strain engineering combined with adaptive laboratory evolution (ALE) (Dai et al., 2018). No matter what the actual mechanisms are, these changes in fermentation mode bring challenges to precise industrial scale process control, especially considering the existence of carbon source and oxygen concentration gradients due to insufficient mixing and mass transfer (Haringa et al., 2018; Jem et al., 1994; Noorman, 2011; Wang et al., 2015).

One of the approaches to tackle these challenges is taking the microorganism's view via integration of a cellular metabolic kinetic model and computational fluid dynamics (CFD) (Haringa et al., 2018, 2022; Haringa, Tang, et al., 2017;

Nadal-Rey et al., 2021; Wang et al., 2019). The kinetic metabolic model describes the microbial response when facing dynamic environments, taking the history of individual cells into account, while the CFD model describes the details of the local cellular environment within the large-scale bioreactor. A successful application of the integration of biokinetics with CFD was first shown by Lapin and coworkers (Lapin et al., 2004) where the glucose uptake in *E. coli* (via the phosphotransferase system, PTS) was integrated into CFD with the Euler-Lagrange approach. Later, (Haringa, Tang, et al., 2017)a successful integration of simple yeast glucose uptake kinetics in an industrial scale stirred tank reactor was applied and based on the simulation results a representative scale-down simulator set-up in the lab was proposed (Haringa, Tang, et al., 2017). Shortly thereafter a structured kinetic model of *Penicillin chrysogenum*(Tang et al., 2017), validated against both long- and short- term dynamics, was integrated into a CFD framework (Tang et al., 2017). It was concluded from these works that an appropriate cellular kinetic model is very often the bottleneck for obtaining meaningful simulation results. This is not only caused by our limited knowledge of the mechanisms of long-term cellular adaptation and short-term perturbation responses, but also by the fact that such cellular kinetic models, when combined with a CFD model, are not mathematically stable and simple enough to avoid computational challenges (Haluk Resat, Linda Petzold, 2010; Miskovic et al., 2019; Wang et al., 2015, 2020).

Kinetic models describing yeast performance in various applications have already been subject of study for decades. As early as 1986,a black box based kinetic model that was capable of reproducing ethanol formation at high specific growth rates was developed, where an upper limit of the non-fermentative pathways was hypothesized to mimic overflow metabolism (Sonnleitner & Käppeli, 1986). Since then, various kinds of kinetic models for *S. cerevisiae* have been developed which were comprehensively reviewed (Lao-Martil et al., 2022). Among these models, many aimed for a better understanding of the physiology and dynamic behavior of the cells at a basic mechanistic level, where most efforts were put on central carbon metabolism (Hynne et al., 2001; Rizzi et al., 1997; Smallbone et al., 2013; van Heerden et al., 2014). These models were usually equipped with hundred or more parameters, required to explicitly describe each reaction of the metabolic network, which is too complicated for CFD integration purposes. Nikerel and coworkers (Nikerel et al., 2009) proposed the use of a lin-log type of approximation of the mechanistic kinetics, such that with much less computational effort, similar dynamics could be predicted. This approach was successfully implemented in

describing glycolysis in yeast. Later, the same authors proposed a systematic approach to further reduce the number of model reactions via lumping of reactions operating close to equilibrium, thereby increasing the parameter identifiability (Nikerel et al., 2011). However, due to the requirement of a reference state in the lin-log approach, it is difficult to directly apply those models to the millions of Lagrangian parcels during model integration. Recently, (La et al., 2020)a simplified dynamic yeast model where conserved moieties like ATP/ADP and NAD^+/NADH were included was published (La et al., 2020). The model was highlighted by using energy (ATP) and/or electron carriers (NADH) to trigger the different metabolic pathways and it showed a promising predictive capacity when applied to a batch process. However, due to a lack of model validation on intracellular compounds and short-term dynamics, e.g. glucose pulse or oscillation, it is difficult to evaluate the accuracy of such models within a CFD framework, where the parcels (representing microbial cells) experience a highly dynamic environment for up to hundreds of hours. To be able to guide the control and optimization of industrial bioprocesses using high-resolution CFD or a lower resolution compartment-based fluid dynamic model, a structured kinetic model is required that is built with the end of the final application in mind (Wang et al., 2020).

In this chapter, we present a structured kinetic model for *S.cerevisiae*, featured with intracellular energy (ATP-ADP) and redox carriers (NADH-NAD^+) and ethanol formation during excess glucose supply and/or oxygen limitation. We demonstrate the capability of the model in describing long-term Crabtree and Pasteur effects triggered by intracellular NADH accumulation (NAD^+ limitation). By adapting the glucose uptake capacity, this model can reproduce the ethanol production in a short-term glucose pulse event and, in a qualitative way, the cellular responses under glucose oscillation conditions. We also proved its suitability for CFD integration by testing it in a highly dynamic, glucose concentration profile simulated for a $22m^3$ pilot tank (Haringa, Deshmukh, et al., 2017a).

Data sources

All experimental data used for model fitting have been collected from publications. The type of data and their sources are detailed in Table 1. For experiment No.2, we only refer to the extracellular glucose and ethanol concentrations but not concentrations of intracellular intermediates. Because those intermediates were measured using a different sample quenching method, which might have led to a systematic deviation, we decided not to use these data (van Gulik et al., 2012).

Table 1 Experimental data used in this Chapter

No.	Type	Strain	Experimental Setting	Reference
1	Steady states	CEN.PK 113-7D	Steady states reached via Upward/downward step changes in dilution rate	(Canelas et al., 2011)
2	Glucose pulses		Glucose pulse experiment based on steady state cultivation with D= 0.05 h^{-1}. The glucose perturbation was executed by injection of concentrated glucose solution. Broth glucose concentration jumps from 0.11 to 2.8mM.	(Mashego et al., 2006)
3	Glucose oscillations		Chemostat with dilution rate of 0.1h^{-1}; feast famine cycles of 400s: fast feed (feast phase) in the first 20s and no feed (famine phase) in the remaining 380s	(Suarez-Mendez et al., 2014)

A 7-pool yeast kinetic model

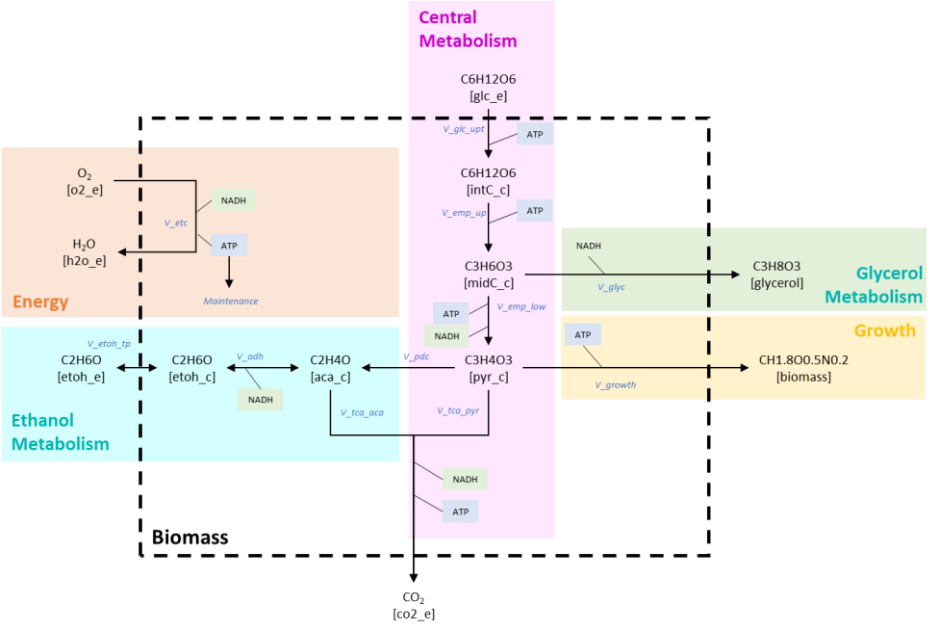

Figure 1. Schematic representation of the 7-pool yeast kinetic model. The 7 intracellular pools are: intC ($C_6H_{12}O_6$), midC($C_3H_6O_3$), Pyr($C_3H_4O_3$), Acetaldehyde (C_2H_4O), Ethanol (C_2H_6O) and two conserved moieties (ATP and NADH).

The simplified structured kinetic model is designed for integration with the computational fluid dynamics (CFD) approach and to provide a complete picture of the cell's lifelines under the dynamic conditions inside the large-scale bioreactor (Haringa et al., 2016). To begin with the end in mind, we applied the same philosophy as in case of an earlier published 9-pool model for *Penicillium chrysogenum* (Tang et al., 2017). We have lumped the glycolysis into three reactions and the TCA cycle into one single reaction. To cover the capability of *S. cerevisiae*'s to produce ethanol during carbon overflow (Crabtree effect (Pfeiffer & Morley, 2014)) or oxygen limited conditions (Pasteur effect (Krebs, 1972)), we describe the redox moiety (NADH/NAD) explicitly, thereby including the oxidation of NADH via the electron transfer chain. For its application (Noorman, 2011; Sarkizi Shams Hajian et al., 2020), the model further needs to be capable of using ethanol as carbon source. The detailed kinetics, stoichiometry and parameterization can be found in Supplementary: Model details. Figure 1 shows a schematic representation of the model. For simplicity purpose, we did not include storage in this model. However, the model structure is designed in such a way that the storage kinetics developed in the

Chapter 4 can be integrated. The ATP consumptions due to futile storage carbohydrates cycles are inherent in maintenance.

Glycolysis

The glycolysis pathway starts from the uptake of glucose and ends at cytosolic pyruvate. The whole glycolysis is split into the uptake of glucose, the upper glycolysis, and the lower glycolysis. The uptake of glucose also includes the first phosphorylation step by hexose kinase - therefore 1 ATP is consumed per glucose. The upper part of glycolysis describes the conversion of 6 carbon glycolytic intermediates (intC) to 3 carbon units (midC), as a result of lumping PGI, PFK and ALDO. Similarly, the lower part of glycolysis describes the conversion from those 3 carbon units (midC) to pyruvate, by lumping TPI, GAPDH, PGK, PGM, ENO, and PK. The stoichiometric coefficients of NADH and ATP are also summarized for both lumped reactions.

There have been several reports on the allosteric regulations for the control of central carbon metabolism of yeast (Banaszak et al., 2011; Cornish-Bowden, 2014; Hackett et al., 2016; Xu et al., 2012). Here, we highlighted two key allosteric regulations that fit our model's structure. The upper emp pathway representing PGI, PFK and ALDO, where ALDO is inhibited by ADP or AMP and PFK can be inhibited by ATP and citrate (Banaszak et al., 2011; Hackett et al., 2016). ATP has the roles of both activator (as a substrate) and inhibitor. As the ATP level in the cell is, in most conditions, rather stable, we only applied the ATP activation term with high affinity (i.e. low Km) to prevent ATP draining when the concentration is too low. Citrate is not an explicitly described species in this model, so we use midC, which has a similar pattern as citrate (decrease at higher growth rate) across different growth rates, to mimic the effect of citrate. This leads to an inhibition term of midC on v_emp_up. A well-known allosteric effector for the lumped lower emp pathway is the activation of PK by FBP (Jurica et al., 1998; Xu et al., 2012). This effect is included via the activation of v_emp_low by intC.

Biomass formation

For simplicity purposes, we implemented one single lumped reaction for biomass formation, derived from the central carbon intermediate pyruvate. We used the general biomass elemental composition $CH_{1.8}O_{0.5}N_{0.2}$ resulting in the following biomass formation reaction from pyruvate:

0.42 $C_3H_4O_3$(pyruvate) + 0.2 NH_4^+ + 1.82 ATP-> 1 $CH_{1.8}O_{0.5}N_{0.2}$(biomass) + 0.2 H^+ + 0.26 CO_2 + 0.24 H_2O

This is a highly lumped conversion without consumption/production of NAD(P)H. Its kinetic rate equation involves limitation by pyruvate, ATP and the 6C unit (intC) in the model. The presence of intC indicates the availability of the 5C unit which is for nucleic acid formation. Ammonia is considered to be never limited and not involved in any kinetics.

The initial guess of growth-associated energy requirement (GAM) and the non-growth associated maintenance energy requirement (NGAM) were 69 mmolATP/gDW (~1.82 molATP/CmolX) and 1 mmol/gDW/h, respectively (Famili et al., 2003). Both values were subjected to further parameterization for a better fit.

TCA cycle

The tricarboxylic acid cycle (TCA) describes the full oxidation of pyruvate to carbon dioxide, with the production of energy (ATP) and electrons (NADH). In this model, one molecule of pyruvate will produce 5 NADH and 1 ATP, assuming FADH2 is equivalent to NADH and GTP is equivalent to ATP.

In the alternative scenario, when the cells use ethanol as carbon source under aerobic conditions, ethanol is first oxidized to acetaldehyde (ACA) via alcohol dehydrogenase (ADH) and then eventually enters the TCA cycle. The ethanol is believed to freely diffuse across the cell membrane, and its consumption rate is limited by the reversible ADH of which the kinetics were referred to in (Kesten et al., 2015; Smallbone et al., 2013).

The flux of TCA is limited by the availability of substrates of this lumped reaction, namely pyruvate, acetaldehyde, and NAD^+. We are aware of the inhibitory effect of NADH on PDH, CS, ICDH and α-KGDH but as NAD+ and NADH serve as conserved moieties, a similar effect can also be realized from the substrate limitation of NAD^+. Therefore no additional product (NADH) inhibition applied to the kinetic of this lumped reaction.

Electron transfer chain (ETC) and NADH balance

Different from the previously published 9-pool kinetic model which describes growth and penicillin production of *Penicillium chrysogenum* (Tang et al., 2017), the current yeast model includes NADH and NAD as additional conserved moiety. It was designed as the signal of overflow metabolism and anaerobic metabolism (Pfeiffer & Morley, 2014). Under aerobic conditions, the majority of NADH is recycled via the electron transfer chain (ETC) and energy is produced according to an actual P/O ratio. The value of the P/O ratio was initially set at 1.04 (Famili et al., 2003). The dynamic behavior of the ETC depends on the

availability of NADH, ADP and O_2. Therefore, the effects of the concentrations of these species on the reaction rate were incorporated as Michaelis-Menten type activation terms.

Mechanism of triggering Crabtree and Pasteur effect

Different mechanisms have been proposed as triggers of aerobic ethanol production from fermentable substrates in *S.cerevisiae*. For convenience purposes, we used a high level of NADH to trigger aerobic ethanol formation. The accumulation of NADH, or more precisely, the limitation of NAD^+ will limit the TCA cycle flux, thus preventing full combustion of the substrate and forcing the cells to select an alternative, NADH-neutral catabolic pathway, i.e. fermentation to ethanol, at the expense of a lower energy efficiency. The accumulation of NADH could be caused by an overflow from the glycolysis that exceeds the capacity of NADH oxidization (ETC), or via a limitation of oxygen supply which reduces the ETC capacity. The accumulation of NADH will limit the availability of NAD^+ for the TCA cycle, leading to accumulation of pyruvate. In either possibility, the surplus of pyruvate is converted to acetaldehyde (ACA) by pyruvate decarboxylase (PDC). This pushes the reversible ADH towards ethanol production and the surplus of NADH is used to reduce ACA. Both ACA and ethanol passively diffuse across the cell membrane based on the concentration gradient. To calculate the intracellular ethanol concentration a cell volume of 2.5ml/gDW was assumed. The above-described mechanism appeared suitable to cover both the Crabtree and Pasteur effects. The current model is yet too simple to consider protein localization limitation or dynamic protein regulation which may also trigger the Crabtree/Pasteur effect (Xia et al., 2022; Zhang et al., 2022).

Predictivity of 7-pool model on multiple steady states

The constructed yeast kinetic model was first parameterized using published experimental data (Canelas et al., 2011), where specific rates and intracellular compound levels of 16 aerobic glucose limited steady state chemostat cultivations were collected at different dilution rates, ranging from D=0.05 h^{-1} to 0.35h^{-1}.

Figure 2 shows the predictivity of the 7-pool model across a wide range of growth rates with a clear shift on two distinct cell responses below and above the critical value. More specifically, the model successfully reproduced the reduced biomass yield on glucose above the critical growth rate which is apparent from the accelerated increase of the specific glucose consumption rate (steeper slope shown in Figure 2B), decreased specific O_2 consumption (Figure 2E) increased specific CO_2 production (Figure 2D) and most importantly, ethanol production (Figure 2E). Furthermore, the model predicted a steep increase of the specific glycerol production rate above a growth rate of 0.18 h^{-1} which did not match the experimental data. This indicates that above this growth rate additional NADH is produced via v_emp_low and v_tca_pyr, which can only be removed via glycerol formation after the limit of the ETC has been reached. Therefore glycerol is produced to balance the extra produced NADH for the partial of growth. The glycerol pattern doesn't match the experimental data probably because the incomplete byproduct profile due to simplicity.

We noticed that the residual glucose concentration at specific growth rates below the critical value was slightly underestimated (Figure 2A), indicating there is an overestimation of either the glucose uptake capacity or the glucose affinity. The best-fit of the K_m for glucose transport appeared to be 1.06 mM in the model, which is in good agreement with a previous report (Canelas, 2010). However, the V_{max} for glucose uptake was fitted to 818 mmol.CmolX^{-1}.h^{-1} which is almost three times higher than the previously reported 279 mmol.CmolX^{-1}.h^{-1} for the same host (Suarez-Mendez et al., 2014). Although the kinetics of this lumped glucose uptake and glucose phosphorylation step also involves the limitation by the intracellular ATP availability and feedback inhibition from its product IntC_c, we believe this fitted value is overestimated. Maier and coworkers summarized the kinetic features of multiple types of hexose transporters (HXT) that are present in the yeast cell and their expression profiles as a function of the specific growth rate (Maier et al., 2002). When the extracellular glucose concentration is low, the cell expresses high affinity transporters (e.g. HXT6, HXT7) of which's the K_m value is around 1 mM. On the

other side, when the cell experiences high glucose concentrations and grows at a higher rate, low glucose affinity transporters (e.g. HXT1) are induced. HXT1 has a K_m for glucose of around 107 mM but in contrast has a 5-time higher K_{cat} than previous high-affinity transporters (Maier et al., 2002).This feature enables yeast cells to transport more glucose with the same amount of protein at high glucose concentrations, in spite of the lower affinity. However, the expression of different hexose transporters based on the residual glucose concentration was not covered by this model. The fitting algorism tends to fit the model to an overestimated glucose uptake V_{max} to satisfy sufficient glucose uptake rate at high growth rates, with as consequence an underestimated glucose concentration at low growth rates. Our observation suggests that for adequately predicting the glucose uptake behavior of yeast across a wide range of growth rates, multiple glucose uptake kinetics (high and low affinity) is needed. Moreover, dynamic enzyme balances should be considered for a smooth shift from one to another. The details of such glucose uptake mechanism will be explained in the "OUTLOOK" chapter of this thesis.

Looking into the capability of the model to describe levels of intermediates of central carbon metabolism, we noticed a clear deviation on the prediction of the 6-carbon unit pool at higher growth rates (Figure 3A). This lumped pool consists of glucose-6-phosphate(G6P), fructose-6-phosphate(F6P) and fructose-1,6-bisphosphate(FBP). The relation between the intracellular levels of these metabolites and the growth rate have been shown to be different (Canelas et al., 2011): Both the levels of G6P and F6P initially increase at increasing growth rate but show a decline when the growth rate increases beyond the critical level. In contrast to this the level of FBP increases with increasing growth rate until the maximum growth rate is reached. FBP plays an important role in activating PK via allosteric regulation and this feature is also captured by this model for triggering overflow of pyruvate resulting in ethanol formation at high growth rates. An important role of this lumped 6-carbon pool is thus the activation of PK (included in v_emp_low) and therefore it shows a different profile compared to the lumped pool of the three measured metabolites. This discrepancy also suggests that lumping metabolites based on structure is not suitable here and an improvement of metabolite lumping can be made for this model based on functionality. For example, metabolites involved in allosteric regulation, such as FBP, should be modeled in a more dedicated way.

Besides the predicted profiles for the glucose uptake rate and the 6-carbon unit pool, the intracellular ATP level also showed slightly larger fluctuations

compared to the experimental data (Figure 3D). Canelas and coworkers (Canelas et al., 2011) showed that the adenylate energy charge of *S. cerevisiae* was around 0.85 with a standard deviation of 0.023 independent of the specific growth rate, suggesting a strong regulation of the intracellular energy availability. The measured decline of the ATP level at higher growth rates (>0.25 h^{-1}) was also observed for ADP and AMP, thus pointing at a slight decrease of the total adenine nucleotide pool for growth rates above the critical value. This observed decline of the total adenine nucleotides is not covered by the model because we fixed the size of the total adenine nucleotide pool and defined ATP and ADP as conserved moieties. By design, ATP and ADP are involved in certain reactions and supposed to play critical roles in the ability of the cells to respond rapidly to perturbations at a time scale of seconds but under (pseudo-)steady state conditions they shall remain at a 'normal' value independent of the growth condition. However, the current model showed a bigger change on the predicted ATP level at different growth rates. This indicates that some ATP dependent reactions in the 7-pool model lack some control factors (due to lumping of the network) and need to use ATP as a replacement. We believe that the current description of the dynamics of ATP is still acceptable, as a compromise of model structure simplification and the assumption of a fixed total adenine nucleotide pool. Nevertheless, the predictivity of ATP can be further improved by using the other moiety, ADP, as activator or inhibitor in the relevant kinetic expressions. As ADP represents around 20% of the total adenine nucleotide pool, the relative change of ADP is more significant than that of ATP and therefore able to reach the same dynamic effect with an even smoother ATP profile (Explanation in Supplementary: conserved moieties).

NADH and NAD^+ is the second conserved moiety pair in this model. From Figure 3E, we see a clear phase shift when the growth rate increases above the critical value. It is also one of the hypotheses for the Crabtree effect that the accumulation of NADH, caused by a capacity limit of either the TCA cycle or the electron transfer chain is, the trigger for ethanol production (Pfeiffer & Morley, 2014). As designed, the simplified model successfully used NADH as trigger for ethanol and glycerol production, although the glycerol produce profile is overestimated and should be improved.

In summary, successful simulation of various steady states across a wide range of specific growth rates proved the predictivity of this model for steady state conditions. The production of ethanol accompanied with glycerol production make this model suitable for describing the overflow conditions (Crabtree

effect). The trigger of ethanol production, namely NADH accumulation, gives confidence in predicting ethanol formation under oxygen limited conditions (Pasteur effect), in which v_etc will likely be limited by oxygen availability and causes NADH accumulation even when the glucose uptake rate does not exceed the critical rate. What's more, the analysis of the model prediction also pointed out several suggestions for further improvements: 1) A proper glucose uptake prediction requires at least two glucose uptake routes, whereby one relies on a high-affinity low-capacity enzyme and the other on a low-affinity but high-capacity enzyme. Both enzymes need to be regulated at expression level so that a smooth shift from low glucose concentration to high glucose concentration can be realized; 2) The lumping of G6P, F6P with FBP is not appropriate because of their different regulatory roles; 3) The assumption of a fixed total adenine nucleotide pool may not be appropriate if the model needs to cover a wide range of growth rates. For reaching a more stable ATP profile, using ADP as activator/inhibitor is suggested; 4) the glycerol production is triggered at the right timing but improvements is required for a quantitative matching to experimental data.

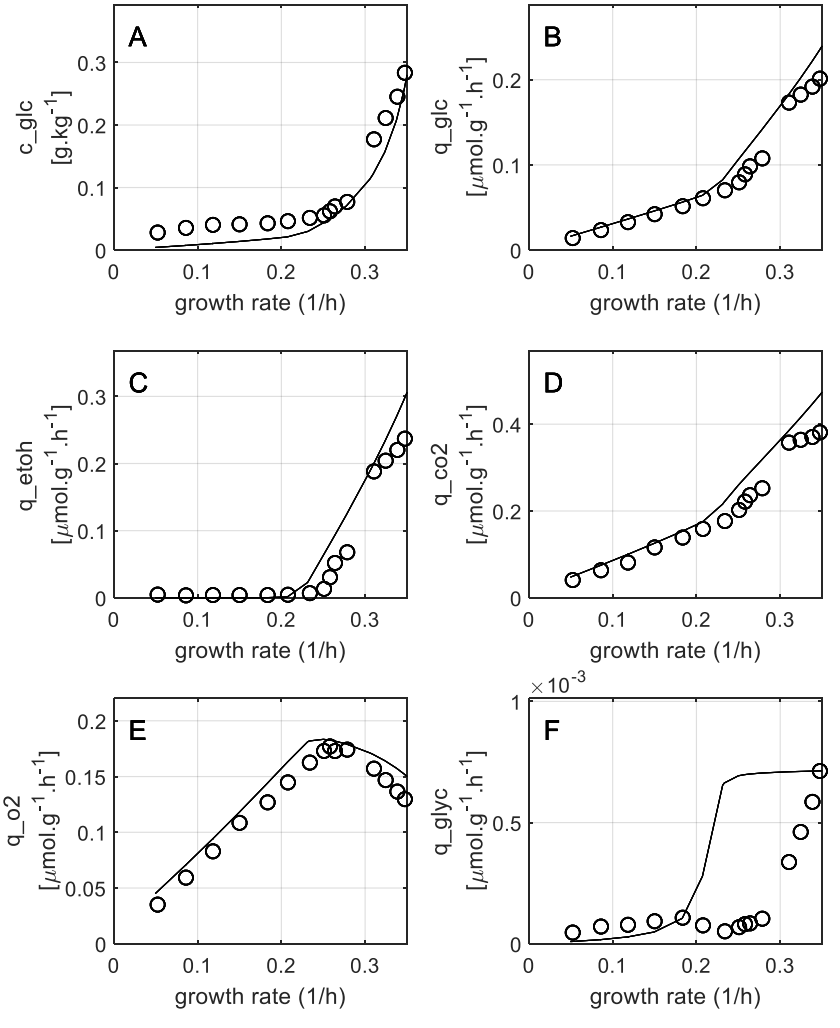

Figure 2. Structured yeast kinetic model validation on multiple steady sates. Glucose concentration and specific production/consumption rate against specific growth rates (solid lines) were plotted on top of experimental data (circle)

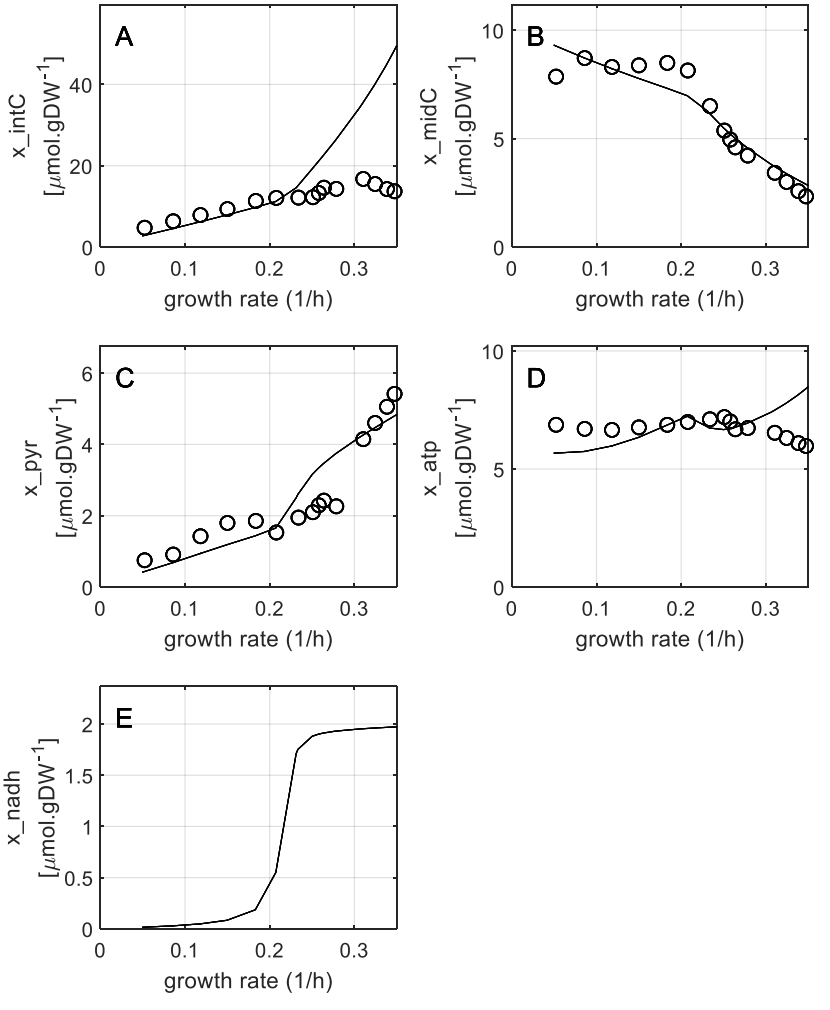

Figure 3. Structured yeast kinetic model validation on multiple steady sates. Intracellular compounds against growth rates (solid lines) were plotted on top of experimental data (circle). Intracellular NADH level is not available for this experiment (Canelas et al., 2011) but the trend matches the cytosolic free NADH/NAD+ ratio at high glucose concentration (Canelas et al., 2008)

Predictivity of a single glucose pulse with glucose uptake capacity adapted to chemostat steady state

After the calibration of the 7-pool model to describe steady state conditions, we tested the model against short-term dynamic conditions. The model was first subjected to a glucose pulse after the system reached a steady state. We compared the response of the model to published data from (Mashego et al., 2006), wherein glucose pulse experiments were executed in both the BioScope (a mini plug flow reactor coupled to a steady state chemostat) and directly in the chemostat for validating the representativeness of the BioScope. We selected the dataset obtained from a direct fermentor pulse for comparison because it fits best to our simulation setting. However, it is good to mention that in the original publications, it is concluded that there are no significant deviations between the response of the cells to a glucose pulse in the BioScope and to a direct pulse in the chemostat.

When applying the original 7-pool model to the pulse conditions, a significant overestimation of glucose consumption and ethanol production is observed (Figure 4, solid line). The pulse of glucose increases the glucose concentration from a limiting level (glucose-limited steady state, D=0.05 h^{-1}) to a sufficiently high, excess level (>>20 K_m). Therefore, the predicted glucose uptake rate immediately reached its maximum capacity. However, the cells were first adapted to a glucose-limited low growth rate and therefore the expressed glucose transporter should be mainly of the high-affinity, low-capacity type (Maier et al., 2002). A sudden high glucose concentration may trigger the cell to start expressing the low affinity transporter (e.g. HXT1) but the turnover time for expressing a protein is much longer than the 400s observation window. Therefore, because of the adaptation to a glucose-limited steady state before the pulse, the cell is still only 'equipped' with the high-affinity low-capacity transporter (e.g. HXT6 and/or HXT7) and the response of the cells should reflect the features of those transporters. To validate our hypothesis, we reduced the glucose uptake capacity of the 7-pool model from 818 $mMol.gDW^{-1}.h^{-1}$ to 279 $mMol.gDW^{-1}.h^{-1}$, a reported value (Suarez-Mendez et al., 2014) which is in agreement with previous reports ((van Dijken et al., 1993). The K_m value was kept the same as it agrees well with the previous reports. This adaptation resulted in a satisfactory description of the experimentally observed glucose and ethanol profiles (Figure 4, dashed lines), which confirmed our hypothesis about the impact of the glucose uptake capacity. Based on this, we would assume that if the pulse experiment would have been carried out after a steady state at a sufficiently high dilution rate or, alternatively, at the end of a batch

cultivation, the dynamics of the same pulse should be closer to the solid line from Figure 4.

Retrospectively, we also compared the model with adapted glucose uptake capacity against the steady state data (Canelas et al., 2011) shown in the previous section. We observed that after this adaptation the model still provides a reasonable description of the residual glucose level if the dilution rate is below 0.2 h^{-1}. When the dilution is increased further, the predicted residual glucose concentration exponentially increases until 50 g/kg (at D=0.3 h^{-1}) and the ODE system will not cover D>0.3h^{-1}. This further makes clear that an alternative glucose transporter with higher capacity is required for reaching the steady state at a high dilution rate.

Although the adaptation of the glucose uptake capacity validated our model in predicting short-term dynamics, we noticed a small discrepancy in the ethanol production profile for the first 50s (Figure 4B). Remarkably, the delayed detection of ethanol after the pulse is only observed from the pulse in fermentor but not in the BioScope. The authors suspect this difference can be attributed to the differences in biomass age, i.e., eight generations in BioScope perturbation and 13 generations in the fermentation perturbation (Mashego et al., 2006). Due to simplification of our 7-pool model, our model is not capable of reproducing this delay in ethanol production, which could be related to protein expression characteristics or cell aging. Still, we believe this has minor impact to the application of our model.

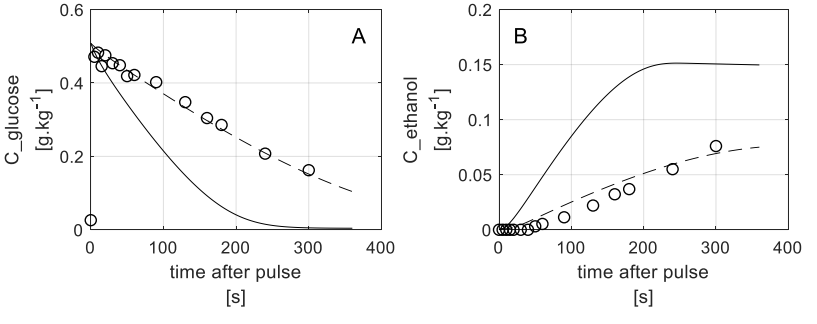

Figure 4. 7-pool kinetic model prediction under single glucose pulse conditions. A: residual glucose concentration after a pulse and B: ethanol production as a response to the glucose pulse. Open circle for experimental data (Mashego et al., 2006). Solid line for the 7-pool model parameterized by all steady states discussed above; dash line for a 7-pool model with reduced glucose uptake capacity: V_glc_upt_max = 279 mMol.CmolX-1.h-1 (instead of 818), which is determined under oscillation conditions with average D=0.1 h-1 (Suarez-Mendez et al., 2014).

Fair model predictivity in reproducing dynamics in short-term oscillation

The model was further challenged in a repeated glucose oscillation experiment, with periodic feeding of glucose (20s on, 380s off). The average dilution rate was kept at 0.1 h^{-1}. For a more detailed experimental setting please refer to the published paper (Suarez-Mendez et al., 2014). Referring to the previous learnings from the single pulse validation, we tested both versions of the 7-pool kinetic models under these glucose feast-famine conditions. Figure 5 shows key extracellular and intracellular pool profiles compared to the corresponding experimental data when the system produces reproducible cycles.

As can be seen from a plot of the residual glucose concentration during one cycle (Fig. 5A solid line) the glucose consumption rate predicted by the original model was too high, similar as was observed for the glucose pulse experiment. However, the prediction with the glucose uptake capacity estimated by Suarez-Mendez et al. (Suarez-Mendez et al., 2014) (279 $mMol.gDW^{-1}.h^{-1}$) fitted the measured residual glucose profile well (Figure 5A, dashed line). As with this value of the maximum glucose uptake capacity also a satisfactory description of the glucose uptake during the single pulse experiment was obtained, this suggests that even after a period of 50 hours of repeated oscillations no adaptation of the glucose uptake kinetics had occurred. This although the maximum residual glucose concentration (0.5 mM) which was reached shortly after the glucose addition can already support a cell growth rate beyond 0.25 h^{-1} and trigger ethanol production.

Due to the high glucose uptake rate in the original version of the model, the Crabtree effect is also triggered incorrectly and causes ethanol accumulation to 0.4 g/kg (Figure 5B). On the other side, almost no ethanol was produced according to model with the reduced glucose uptake capacity. This agrees with the published experimental result where no ethanol production was observed(Suarez-Mendez et al., 2014).

Regarding the intracellular compounds, both versions of the model captured the correct trend of primary carbon metabolites, featured by a high peak for the 6-carbon unit pool (intC) but a dip of the 3-carbon unit pool (midC) in the first 100s of each cycle (Figure 5C and D). The model with reduced glucose uptake capacity showed smaller changes in both intermediates, which makes perfect sense because of a smaller glucose uptake peak. However, both models missed the peak of the 3-carbon unit pool at around 200s. The increase of this lumped pool is mainly caused by the accumulation of 3-phosphoglycerate (3PG),

2-phosphoglycerate (2PG) and phosphoenolpyruvate (PEP) caused by yet incompletely known regulatory mechanisms at the metabolite level (Suarez-Mendez et al., 2014).

We observed a clear deviation on the pyruvate profile for both models during a cycle. Both models suggest a pyruvate profile that should be similar to the profile of extracellular glucose or intC. This makes sense from the model's perspective as the level of pyruvate is critical to determine the rate of either the TCA cycle or the ethanol production (PDC, under overflow conditions). The stable pyruvate level was explained by compartmentalization or a buffer capacity of alanine, a closely related amino acid that has a concentration about 100 times higher than pyruvate (Suarez-Mendez et al., 2014). None of those were considered in our highly lumped model. Depending on the application of the model, a buffer pool of pyruvate can be added to bring the pyruvate profile closer to experimental data. At this point, we don't think the mismatch of pyruvate will cause serious problems to the model performance.

The predicted ATP profiles of both models showed interesting results: the model with the high glucose uptake capacity predicted an initial steep decrease of the ATP concentration while the model with reduced glucose uptake rate didn't. Typically, one would expect an initial decrease of the ATP when there is a sudden increase in glucose concentration (Mashego et al., 2006; Theobald et al., 1997), which is in generally referred to as the 'ATP paradox' (Somsen et al., 2000). However, this paradox was not observed in the original glucose oscillation experiment we refer to here(Suarez-Mendez et al., 2014). It was explained by authors that 1) the trigger of the ATP paradox depends on the maximum glucose concentration during the oscillation or pulse and/or 2) the microorganism was 'trained' and has increased its oxygen uptake capacity to prevent the use of the ATP salvage pathway. Considering that the ATP paradox can also be observed in the previous single-pulse simulation by the model, even with a reduced glucose uptake capacity (Figure S1), we believe this ATP paradox is triggered due to the short-term imbalance between phosphates/ATP regeneration due to a sudden increase of glucose uptake rate, which was reported previously (van Heerden et al., 2014). Because our model did not cover the ATP salvage pathway and the total pool of adenine nucleotides is assumed constant, the initial glucose phosphorylation could be already sufficient to explain that the ATP paradox occurred during glucose oscillations/pulses. Unfortunately, the 7-pool model is not suitable to validate the second

hypothesis due to its simplified structure and missing gene regulation mechanisms.

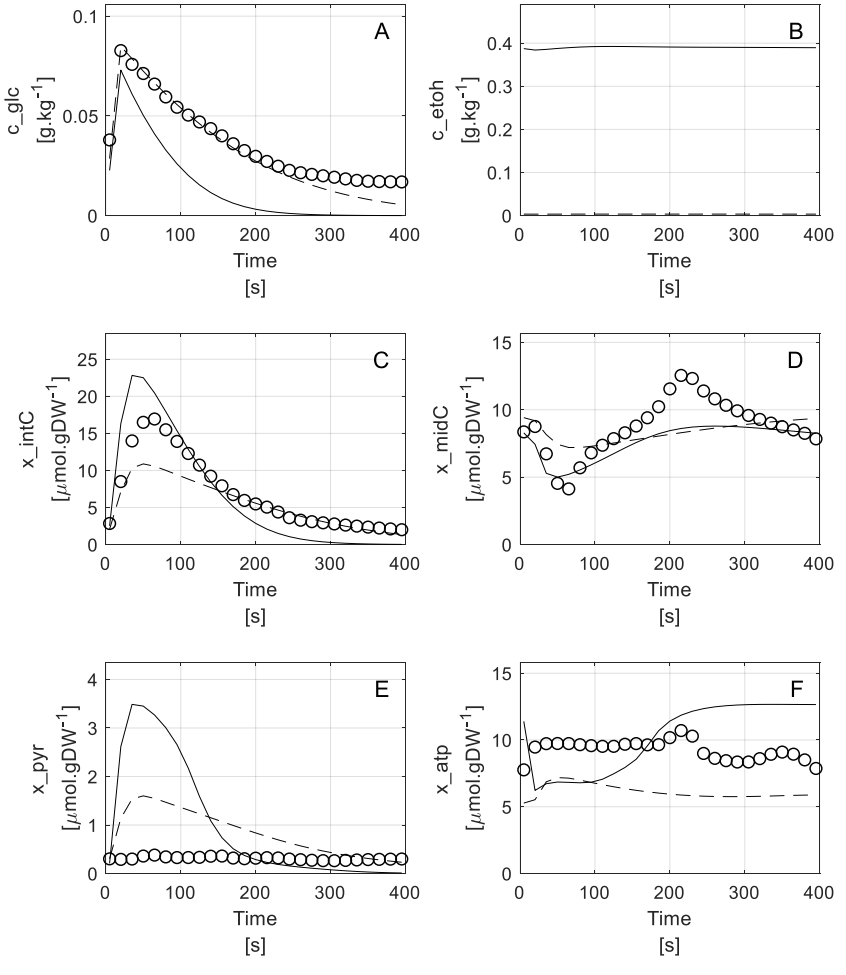

Figure 5. 7-pool kinetic model prediction under glucose oscillation conditions (20s feed on, 380s feed off). A and B: the extracellular glucose and ethanol concentration (mol.m^{-3}) in one glucose oscillation cycle; C,D,E and F: intracellular compound profiles of intC(C6), midC(C3), pyruvate and ATP. Open circle for experimental data collected from Camilo A, et.al. (2014), solid line for prediction from the 7-pool model parameterized by all steady states discussed above; dash line for a 7-pool model with reduced glucose uptake capacity: $V_{max,glcupt}$ = 279 mMol.CmolX^{-1}.h^{-1} (instead of 818), which is determined under oscillation condition with average D=0.1 h^{-1}(Suarez-Mendez et al., 2014). All extracellular concentrations with units of mol.m^{-3} and all intracellular concentration with units of mol.m^{-3}(cell), assuming 26gDW.CmolX^{-1} and 2.5ml.gDW^{-1}.

Suitability of the kinetic model for CFD simulation

Finally, we tested the suitability of our 7-pool model for integration with CFD-based reactor models. For this, we adapted the cell's lifelines extracted from a 22 m^3 large pilot bioreactor (Haringa, Deshmukh, et al., 2017b). Like the test in the previous Chapter, we connected 120 separate original lifelines into a long, single lifeline which lasts for about 40 hours, to enable the model to reach an unbiased dynamic state.

Here we only tested the 7-pool model with reduced glucose uptake capacity (279 $mmol.gDW^{-1}.h^{-1}$) because this glucose uptake capacity is able to provide a more realistic glucose profile for both the single glucose pulse and the glucose oscillation conditions (Figure 4A and Figure 5A). The glucose affinity constant fitted for the 7-pool model is very similar to the one used in the CFD simulation of Haringa and coworkers (Haringa, Deshmukh, et al., 2017b) and therefore no adaption of K_m was needed. The lifeline consists of glucose concentrations experienced by the Lagrange parcels in the CFD framework. Dissolved oxygen was proven to be higher than 0.08mM (30% of saturated dissolved oxygen under 1 bar) and therefore no Pasteur effect was expected.

All tests with the 20 lifelines went smoothly without any stability issues, which assured the fitness and robustness of the 7-pool model in handling this type of highly dynamic glucose concentrations. Figure 6 shows a snapshot of the responses of the cells to one tested lifeline in a 720s time window. The model showed a similar dynamic feature as in the controlled glucose oscillation: the intC and pyruvate pool showed a similar pattern as extracellular glucose but more dampened. The midC pool also showed a proper inverted pattern compared to intC. The predicted ATP showed a rather stable value around of 6 $\mu mol.gDW^{-1}$ and the NADH level showed a high peak as it rapidly responds when the cells were suddenly exposed to high glucose concentrations. Based on these local observations, we believe the 7-pool model performs according to our expectations in the provided set of cell lifelines.

When looking from a longer time scale perspective, we observed proper long term cell growth under the conditions of these lifelines with an average growth rate around 0.07 h^{-1}. We also do not see significant discrepancies among different lifelines, suggesting a consistent model response to these long-term glucose profiles. Ethanol production was observed to occur at a low level (maximum concentration of about 0.1 g/kg at a biomass intensity of around 90gDW/kg). This indicates that the short-term Crabtree effect may be triggered due to local glucose accumulation in the 22 m^3 large pilot scale tank. Higher

ethanol peaks were observed in the later phase of the simulation which is a result of the higher biomass concentration. Nevertheless, the prediction of glycerol production is problematic as it keeps accumulating till the end of the fermentation. This is a defect of the 7-pool model because of the absence of a glycerol uptake mechanism. This will not only cause incorrect predictions on the glycerol profile but also an underestimation of CO_2 production and biomass formation due to unrealistic carbon loss. Despite this defect, we confirmed that the 7-pool model showed stable and reasonable cell responses when facing long term rapid glucose oscillations. This successfully proved its fitness for a complete CFD integration.

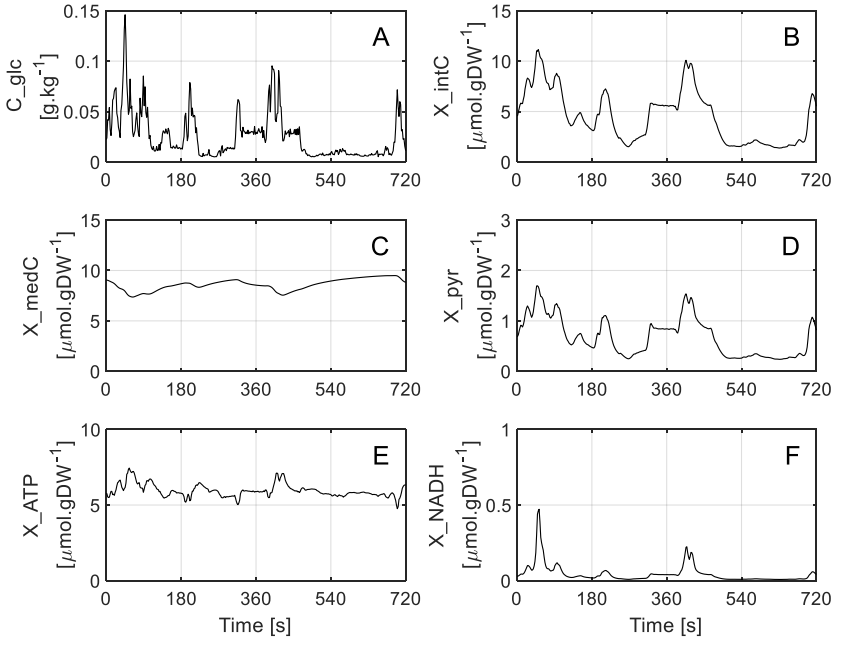

Figure 6 A snapshot of intracellular cell responses during a 720s cell lifeline. A: the experienced glucose profile, simulated by (Haringa, Deshmukh, et al., 2017b) for a 22m^3 large pilot scale bioreactor. B-F: corresponding intracellular cell response to the experienced glucose concentration.

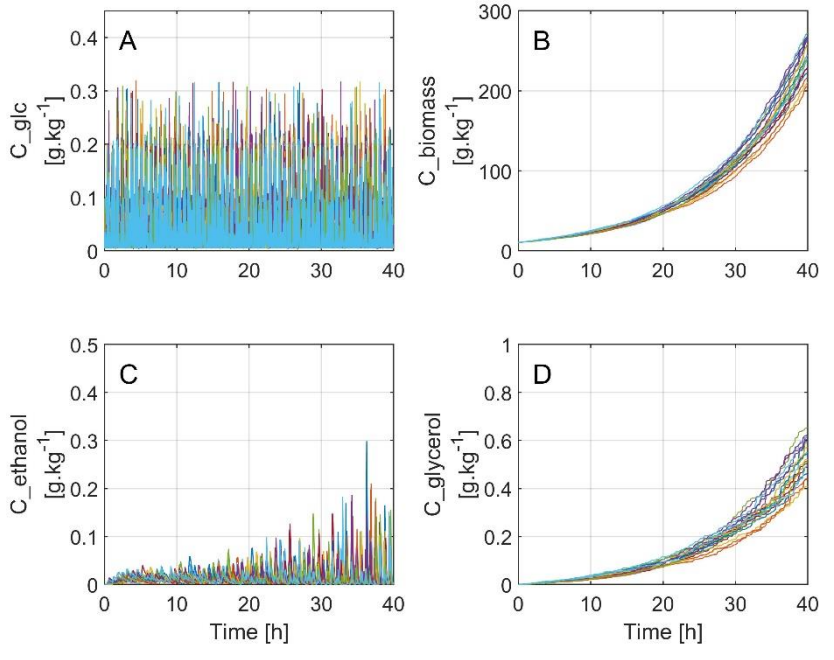

Figure 7 Extracellular compounds' profile in 20 long-term (40h) lifelines simulated by the 7-pool model. A: Experienced glucose profile; B: biomass growth under such glucose oscillation regime. The biomass profile follows an exponential growth with a specific growth rate of 0.07 h^{-1} C&D: accumulation of two main byproducts.

Conclusions

In this Chapter, the possibility of using a compact kinetic model to describe the dynamics behavior of *Saccharomyces cerevisiae* as response to glucose gradients is explored. A 7-pool lumped kinetic model covering glycolysis, TCA cycle, electron transfer chain and overflow metabolism was first proposed and showed a reasonable prediction under steady state conditions, including at specific growth rates that are beyond the critical value for fully respiratory metabolism. The ethanol formation was successfully triggered by NADH accumulation caused by either overflow (glucose excess, Crabtree effect) or limited ETC capacity (oxygen limitation, Pasteur effect) in this model. The model was further tested for single glucose pulse conditions and periodic glucose oscillations. It showed good predictivity under glucose pulse conditions and qualitatively predicted the glucose oscillation conditions. Finally, the model was tested with highly dynamic glucose profiles simulated for a 22 m^3 large pilot bioreactor. This 7-pool model showed sufficient stability in handling such dynamics and maintained proper kinetic features throughout multiple long lifelines. This proves that the 7-pool model is suitable for integrating into a complete CFD framework.

Nevertheless, we also noticed that for reproducing short-term dynamics, the glucose uptake capacity has to be adapted to a high-affinity low-capacity glucose transporter which is normally expressed at low residual glucose conditions. This adaptation will not impact the predictivity of steady states below a dilution rate of 0.2 h^{-1} but sacrifices the model's predictivity at higher growth rates. A proper solution is to introduce two or more parallel glucose uptake mechanisms and regulate the use of the glucose uptake channels via enzyme balancing (Tang et al., 2017). We also suggest several other possible improvements including a better design for lumped intracellular species and additional mechanisms for consuming produced byproducts such as glycerol. In the Outlook Chapter of this thesis, we will address those opportunities in more detail and propose an extended model structure for the future work.

References

Banaszak, K., Mechin, I., Obmolova, G., Oldham, M., Chang, S. H., Ruiz, T., Radermacher, M., Kopperschläger, G., & Rypniewski, W. (2011). The crystal structures of eukaryotic phosphofructokinases from Baker's yeast and rabbit skeletal muscle. *Journal of Molecular Biology*, *407*(2), 284–297. https://doi.org/10.1016/j.jmb.2011.01.019

Buijs, N. A., Siewers, V., & Nielsen, J. (2013). Advanced biofuel production by the yeast saccharomyces cerevisiae. In *Current Opinion in Chemical Biology* (Vol. 17, Issue 3, pp. 480–488). https://doi.org/10.1016/j.cbpa.2013.03.036

Canelas, A. B. (2010). Towards quantitative metabolomics and in vivo kinetic modeling in S. cerevisiae. *TU Delft - Biotechnology Department - Bioprocess Technology Group*.

Canelas, A. B., Ras, C., ten Pierick, A., van Gulik, W. M., & Heijnen, J. J. (2011). An in vivo data-driven framework for classification and quantification of enzyme kinetics and determination of apparent thermodynamic data. *Metabolic Engineering*, *13*(3), 294–306. https://doi.org/https://doi.org/10.1016/j.ymben.2011.02.005

Canelas, A. B., van Gulik, W. M., & Heijnen, J. J. (2008). Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions. *Biotechnology and Bioengineering*, *100*(4), 734–743.

Cornish-Bowden, A. (2014). Understanding allosteric and cooperative interactions in enzymes. *FEBS Journal*, *281*(2), 621–632. https://doi.org/10.1111/febs.12469

Dai, Z., Huang, M., Chen, Y., Siewers, V., & Nielsen, J. (2018). Global rewiring of cellular metabolism renders Saccharomyces cerevisiae Crabtree negative. *Nature Communications*, *9*(1), 3059. https://doi.org/10.1038/s41467-018-05409-9

Famili, I., Forster, J., Nielsen, J., & Palsson, B. O. (2003). Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(23), 13134–13139. https://doi.org/10.1073/pnas.2235812100

Hackett, S. R., Zanotelli, V. R. T., Xu, W., Goya, J., Park, J. O., Perlman, D. H., Gibney, P. A., Botstein, D., Storey, J. D., & Rabinowitz, J. D. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science*, *354*(6311). https://doi.org/10.1126/science.aaf2786

Haluk Resat, Linda Petzold, and M. F. P. (2010). *Kinetic Modeling of Biological Systems*. 9–21. https://doi.org/10.1007/978-1-59745-243-4

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017a). Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669.

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017b). Euler-Lagrange analysis towards representative down-scaling of a 22m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669. https://doi.org/https://doi.org/10.1016/j.ces.2017.01.014

Haringa, C., Mudde, R. F., & Noorman, H. J. (2018). From industrial fermentor to CFD-guided downscaling: what have we learned? *Biochemical Engineering Journal*. https://doi.org/https://doi.org/10.1016/j.bej.2018.09.001

Haringa, C., Tang, W., Deshmukh, A. T., Xia, J., Reuss, M., Heijnen, J. J., Mudde, R. F., & Noorman, H. J. (2016). Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: An analysis of organism lifelines. *Engineering in Life Sciences*, *16*(7), 652–663. https://doi.org/doi:10.1002/elsc.201600061

Haringa, C., Tang, W., & Noorman, H. J. (2022). Stochastic parcel tracking in an Euler-Lagrange compartment model for fast simulation of fermentation processes. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.28094

Haringa, C., Tang, W., Wang, G., Deshmukh, A. T., van Winden, W. A., Chu, J., van Gulik, W. M., Heijnen, J. J., Mudde, R. F., & Noorman, H. J. (2017). Computational fluid dynamics simulation of an industrial P.

chrysogenum fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization. *Chemical Engineering Science*. https://doi.org/https://doi.org/10.1016/j.ces.2017.09.020

Hynne, F., Danø, S., & Sørensen, P. G. (2001). Full-scale model of glycolysis in Saccharomyces cerevisiae. *Biophysical Chemistry*, *94*(1), 121–163. https://doi.org/https://doi.org/10.1016/S0301-4622(01)00229-0

Jem, K. J., Fateen, S., & Michaels, J. (1994). Mixing phenomena in industrial bioreactors with perfusion spin filters. In *Animal Cell Technology* (pp. 392–396). Elsevier.

Jurica, M. S., Mesecar, A., Heath, P. J., Shi, W., Nowak, T., & Stoddard, B. L. (1998). The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure*, *6*(2), 195–210. http://biomednet.com/elecref/0969212600600195

Kesten, D., Kummer, U., Sahle, S., & Hübner, K. (2015). A new model for the aerobic metabolism of yeast allows the detailed analysis of the metabolic regulation during glucose pulse. *Biophysical Chemistry*, *206*, 40–57. https://doi.org/https://doi.org/10.1016/j.bpc.2015.06.010

Krebs, H. A. (1972). The Pasteur effect and the relations between respiration and fermentation. *Essays Biochem*, *8*, 1–34.

La, A., Du, H., Taidi, B., & Perré, P. (2020). A predictive dynamic yeast model based on component, energy, and electron carrier balances. *Biotechnology and Bioengineering*, *117*(9), 2728–2740. https://doi.org/10.1002/bit.27442

Lao-Martil, D., Verhagen, K., Schmitz, J., Teusink, B., Wahl, S., & van Riel, N. (2022). Kinetic Modeling of Saccharomyces cerevisiae Central Carbon Metabolism: Achievements, Limitations, and Opportunities. *Metabolites*, *12*(1), 74. https://doi.org/10.3390/metabo12010074

Lapin, A., Müller, D., & Reuss, M. (2004). Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler- Lagrange methods: Traveling along the lifelines of single cells. *Industrial & Engineering Chemistry Research*, *43*(16), 4647–4656.

Maier, A., Völker, B., Boles, E., & Fuhrmann, G. F. (2002). Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Res*, *2*(4), 539–550. https://doi.org/10.1111/j.1567-1364.2002.tb00121.x

Mashego, M. R., van Gulik, W. M., Vinke, J. L., Visser, D., & Heijnen, J. J. (2006). In vivo kinetics with rapid perturbation experiments in Saccharomyces cerevisiae using a second-generation BioScope. *Metabolic Engineering*, *8*(4), 370–383.

Miskovic, L., Béal, J., Moret, M., & Hatzimanikatis, V. (2019). Uncertainty reduction in biochemical kinetic models: Enforcing desired model properties. *PLOS Computational Biology*, *15*(8), e1007242. https://doi.org/10.1371/journal.pcbi.1007242

Nadal-Rey, G., McClure, D. D., Kavanagh, J. M., Cassells, B., Cornelissen, S., Fletcher, D. F., & Gernaey, K. v. (2021). Development of dynamic compartment models for industrial aerobic fed-batch fermentation processes. *Chemical Engineering Journal*, *420*, 130402.

Niebel, B., Leupold, S., & Heinemann, M. (2019). An upper limit on Gibbs energy dissipation governs cellular metabolism. *Nature Metabolism*, *1*(1), 125–132. https://doi.org/10.1038/s42255-018-0006-7

Nielsen, J., Larsson, C., van Maris, A., & Pronk, J. (2013). Metabolic engineering of yeast for production of fuels and chemicals. In *Current Opinion in Biotechnology* (Vol. 24, Issue 3, pp. 398–404). https://doi.org/10.1016/j.copbio.2013.03.023

Nikerel, I. E., Canelas, A. B., Jol, S. J., Verheijen, P. J. T., & Heijnen, J. J. (2011). Construction of kinetic models for metabolic reaction networks: Lessons learned in analysing short-term stimulus response data. *Mathematical and Computer Modelling of Dynamical Systems*, *17*(3), 243–260. https://doi.org/10.1080/13873954.2010.548167

Nikerel, I. E., van Winden, W. A., Verheijen, P. J. T., & Heijnen, J. J. (2009). Model reduction and a priori kinetic parameter identifiability analysis using metabolome time series for metabolic reaction

networks with linlog kinetics. *Metabolic Engineering*, *11*(1), 20–30. https://doi.org/https://doi.org/10.1016/j.ymben.2008.07.004

Noorman, H. (2011). An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnology Journal*, *6*(8), 934–943.

Otero, J. M., & Nielsen, J. (2010). Industrial systems biology. *Biotechnology and Bioengineering*, *105*(3), 439–460.

Parapouli, M., Vasileiadis, A., Afendra, A. S., & Hatziloukas, E. (2020). Saccharomyces cerevisiae and its industrial applications. In *AIMS Microbiology* (Vol. 6, Issue 1, pp. 1–31). AIMS Press. https://doi.org/10.3934/microbiol.2020001

Pfeiffer, T., & Morley, A. (2014). An evolutionary perspective on the Crabtree effect. *Frontiers in Molecular Biosciences*, *1*(OCT). https://doi.org/10.3389/fmolb.2014.00017

Rizzi, M., Baltes, M., Theobald, U., & Reuss, M. (1997). In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model. *Biotechnology and Bioengineering*, *55*(4), 592–608. https://doi.org/doi:10.1002/(SICI)1097-0290(19970820)55:4<592::AID-BIT2>3.0.CO;2-C

Sander, R. (2015). Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmos. Chem. Phys.*, *15*(8), 4399–4981. https://doi.org/10.5194/acp-15-4399-2015

Sarkizi Shams Hajian, C., Haringa, C., Noorman, H., & Takors, R. (2020). Predicting By-Product Gradients of Baker's Yeast Production at Industrial Scale: A Practical Simulation Approach. *Processes*, *8*(12), 1554. https://www.mdpi.com/2227-9717/8/12/1554

Schumacher, R. (2018). *Metabolic trade-offs arising from increased free energy conservation in Saccharomyces cerevisiae* [Doctoral thesis, Delft University of Technology]. https://doi.org/10.4233/uuid:177e9f4c-f847-436d-9fd4-9ed97ba709d9

Smallbone, K., Messiha, H. L., Carroll, K. M., Winder, C. L., Malys, N., Dunn, W. B., Murabito, E., Swainston, N., Dada, J. O., Khan, F., Pir, P., Simeonidis, E., Spasic, I., Wishart, J., Weichart, D., Hayes, N. W., Jameson, D., Broomhead, D. S., Oliver, S. G., … Mendes, P. (2013). A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. *FEBS Lett*, *587*(17), 2832–2841. https://doi.org/10.1016/j.febslet.2013.06.043

Somsen, O. J. G., Hoeben, M. A., Esgalhado, E., Snoep, J. L., Visser, D., van der Heijden, R. T. J. M., Heijnen, J. J., & Westerhoff, H. v. (2000). Glucose and the ATP paradox in yeast. In *Biochem. J* (Vol. 352).

Sonnleitner, B., & Käppeli, O. (1986). Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis. *Biotechnol Bioeng*, *28*(6), 927–937. https://doi.org/10.1002/bit.260280620

Suarez-Mendez, C. A., Sousa, A., Heijnen, J. J., & Wahl, A. (2014). Fast "Feast/Famine" Cycles for Studying Microbial Physiology Under Dynamic Conditions: A Case Study with Saccharomyces cerevisiae. *Metabolites*, *4*(2), 347. http://www.mdpi.com/2218-1989/4/2/347

Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J. J., Xia, J., Chu, J., & others. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by Penicillium chrysogenum. *Biotechnology and Bioengineering*, *114*(8), 1733–1743.

Theobald, U., Mailinger, W., Baltes, M., Rizzi, M., & Reuss, M. (1997). In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae : I. Experimental observations. *Biotechnology and Bioengineering*, *55*(2), 305–316. https://doi.org/doi:10.1002/(SICI)1097-0290(19970720)55:2<305::AID-BIT8>3.0.CO;2-M

van Dijken, J. R., Weusthuis, R. A., & Pronk, J. T. (1993). Kinetics of growth and sugar consumption in yeasts. In *Antonie van Leeuwenhoek* (Vol. 63).

van Gulik, W. M., Canelas, A. B., Taymaz-Nikerel, H., Douma, R. D., Jonge, L. P. de, & Heijnen, J. J. (2012). Fast sampling of the cellular metabolome. In *Microbial Systems Biology* (pp. 279–306). Springer.

van Heerden, J. H., Wortel, M. T., Bruggeman, F. J., Heijnen, J. J., Bollen, Y. J. M., Planqué, R., Hulshof, J., O'Toole, T. G., Wahl, S. A., & Teusink, B. (2014). Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. *Science*, *343*(6174). https://doi.org/10.1126/science.1245114

Verduyn, C., Postma, E., Scheffers, W. A., & van Dijken, J. P. (1990). Physiology of Saccharomyces cerevisiae in anaerobic glucose-limited chemostat cultures. *Journal of General Microbiology*, *136*(3), 395–403. https://doi.org/10.1099/00221287-136-3-395

Wang, G., Haringa, C., Noorman, H., Chu, J., & Zhuang, Y. (2020). Developing a Computational Framework To Advance Bioprocess Scale-Up. *Trends in Biotechnology*, *38*(8), 846–856. https://doi.org/https://doi.org/10.1016/j.tibtech.2020.01.009

Wang, G., Haringa, C., Tang, W., Noorman, H., Chu, J., Zhuang, Y., & Zhang, S. (2019). Coupled Metabolic-hydrodynamic Modeling Enabling Rational Scale-up of Industrial Bioprocesses. *Biotechnology and Bioengineering*, *n/a*(n/a). https://doi.org/10.1002/bit.27243

Wang, G., Tang, W., Xia, J., Chu, J., Noorman, H., & van Gulik, W. M. (2015). Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. *Engineering in Life Sciences*, *15*(1), 20–29.

Xia, J., Sánchez, B. J., Chen, Y., Campbell, K., Kasvandik, S., & Nielsen, J. (2022). Proteome allocations change linearly with the specific growth rate of Saccharomyces cerevisiae under glucose limitation. *Nature Communications*, *13*(1), 2819. https://doi.org/10.1038/s41467-022-30513-2

Xie, T., Chen, M., Nielsen, J., & Xia, J. (2022). Multi-omics analyses of the transition to the Crabtree effect in S. cerevisiae reveals a key role for the citric acid shuttle. *FEMS Yeast Research*. https://doi.org/10.1093/femsyr/foac030

Xu, Y. F., Zhao, X., Glass, D. S., Absalan, F., Perlman, D. H., Broach, J. R., & Rabinowitz, J. D. (2012). Regulation of Yeast Pyruvate Kinase by Ultrasensitive Allostery Independent of Phosphorylation. *Molecular Cell*, *48*(1), 52–62. https://doi.org/10.1016/j.molcel.2012.07.013

Zhang, Y., Su, M., Wang, Z., Nielsen, J., & Liu, Z. (2022). Rewiring regulation on respiro-fermentative metabolism relieved Crabtree effects in Saccharomyces cerevisiae. *Synthetic and Systems Biotechnology*. https://doi.org/10.1016/j.synbio.2022.06.004

Supplementary: Model details

Model Stoichiometry

Table S1 Stoichiometry matrix of the 7-pool kinetic model

			v_glc_upt	v_emp_up	v_emp_low	v_tca_pyr	v_tca_ac	v_etc	v_sto_in*	v_sto_out*	v_glyc	v_pdc	v_adh	v_etoh_tp	v_growth
Intracellular	Int_C	$C_6H_{12}O_6$	1	-1					-1	1					
	Mid_C	$C_3H_6O_3$		2	-1						-1				
	Sto_C	$C_{12}H_{22}O_{11}$							0.5	-0.5					
	Pyr	$C_3H_4O_3$			1	-1						-1			-0.42
	Aca	C_2H_4O					-1					1	-1		
	EtOH	C_2H_6O											1	-1	
	NADH	NAD-H			1	5	5	-2			-1		-1		
	ATP	-	-1	-1	2	1	-1	2*PO	-1	-1					GAM
Extracellular	Glc	$C_6H_{12}O_6$	-1												
	Fru	$C_6H_{12}O_6$													
	Glycerol	$C_3H_8O_3$									1				
	EtOH	C_2H_6O												1	
	Biomass	$CH_{1.8}O_{0.5}N_{0.2}$													1
	Ammonia	NH_4													-0.2
	Proton	H			1	5	5	-2			-1		-1		0.2
	water	H_2O				-3	-3	2	0.5	-0.5					0.24
	oxygen	O_2						-1							
	Carbon dioxide	CO_2				3	2					1			0.26

* The storage cycle is excluded in this kinetic model.

Kinetics

Table S2 Core kinetics applied in 7-pool kinetic model

Conversion	Kinetics
glucose uptake	$v_{glcupt} = v_{max,glcupt} \cdot \frac{C_{EC,glc}}{C_{EC,glc}+K_{m,glcupt,glc}} \cdot \frac{C_{IC,atp}}{C_{IC,atp}+K_{m,glcupt,atp}} \cdot \frac{C_{IC,intC}}{C_{IC,intC}+K_{m,glcupt,intC}}$
higher glycolysis	$v_{empup} = v_{max,empup} \cdot \frac{C_{IC,intC}}{C_{IC,intC}+K_{m,empup,intC}} \cdot \frac{C_{IC,atp}}{C_{IC,atp}+K_{m,empup,atp}} \cdot \frac{K_{i,empup,midC}}{C_{IC,midC}+K_{i,empup,midC}}$
lower glycolysis	$v_{emplow} = v_{max,emplow} \cdot \frac{C_{IC,intC}}{C_{IC,intC}+K_{m,emplow,intC}} \cdot \frac{C_{IC,midC}}{C_{IC,midC}+K_{m,emplow,midC}}$
TCA cycle (pyruvate)	$v_{tcapyr} = v_{max,tcapyr} \cdot \frac{C_{IC,pyr}}{C_{IC,pyr}+K_{m,tcapyr,pyr}} \cdot \frac{C_{IC,nad}}{C_{IC,nad}+K_{m,tcapyr,nad}}$
PDC	$v_{pdc} = v_{max,pdc} \cdot \frac{{C_{IC,pyr}}^3}{{C_{IC,pyr}}^3+{K_{m,pdc,pyr}}^3}$
TCA cycle (acetaldehyde)	$v_{tcaaca} = v_{max,tcaaca} \cdot \frac{C_{IC,aca}}{C_{IC,aca}+K_{m,tcaaca,aca}} \cdot \frac{C_{IC,nad}}{C_{IC,nad}+K_{m,tcaaca,nad}}$
ETC	$v_{etc} = v_{max,etc} \cdot \frac{{C_{EC,o2}}^2}{{C_{EC,o2}}^2+{K_{m,etc,o2}}^2} \cdot \frac{C_{IC,nadh}}{C_{IC,nadh}+K_{m,etc,nadh}} \cdot \frac{C_{IC,adp}}{C_{IC,adp}+K_{m,etc,adp}}$
maintenance (non-growth associate)	$v_{ngam} = v_{max,ngam} \cdot \frac{{C_{IC,atp}}^5}{{C_{IC,atp}}^5+{K_{m,ngam,atp}}^5}$
glycerol formation	$v_{glyc} = v_{max,glyc} \cdot \frac{C_{IC,midC}}{C_{IC,midC}+K_{m,glyc,midC}} \cdot \frac{C_{IC,nadh}}{C_{IC,nadh}+K_{m,glyc,nadh}}$
ADH	$v_{adh} = v_{max,adh} \cdot \frac{\frac{1}{K_{m,adh,aca} \cdot K_{m,adh,nadh}} \left(C_{IC,aca} C_{IC,nadh} - \frac{C_{IC,etoh} C_{IC,nad}}{K_{eq,adh}} \right)}{\left(1+\frac{C_{IC,aca}}{K_{m,adh,aca}}\right) \cdot \left(1+\frac{C_{IC,nadh}}{K_{m,adh,nadh}}\right) + \left(1+\frac{C_{IC,etoh}}{K_{m,adh,etoh}}\right) \cdot \left(1+\frac{C_{IC,nad}}{K_{m,adh,nad}}\right) - 1}$
biomass formation	$v_{growth} = v_{max,growth} \cdot \frac{C_{IC,intC}}{C_{IC,intC}+K_{m,growth,intC}} \cdot \frac{C_{IC,pyr}}{C_{IC,pyr}+K_{m,growth,pyr}} \cdot \frac{{C_{IC,atp}}^3}{{C_{IC,atp}}^3+{K_{m,growth,atp}}^3}$
ethanol transport	$v_{etoh,transport} = k_{transport,etoh} \cdot \left(C_{IC,etoh} - C_{EC,etoh}\right)$

Parameters

Table S3 Parameter values of the 7-pool kinetic model fitted to multiple steady states

Parameter	Unit	Value	Remark	Source
P/O	dimensionless	0.950	amount of atp produced per O atom	(Famili et al., 2003)
vmax_glcupt	molintC/CmolX/h	0.279	max glucose uptake rate	(Suarez-Mendez et al., 2014)
K_glcupt_glc	mol/m3(broth)	1.062	affinity constant glc	(Suarez-Mendez et al., 2014)
K_glcupt_atp	mol/m3(cell)	0.101	affinity constant atp	Fitted
Ki_glcupt_intC	mol/m3(cell)	20.44	inhibition constant c6	Fitted
vmax_empup	molintC/CmolX/h	1.289	max conversion rate from c6 to c3	Fitted
K_empup_intC	mol/m3(cell)	21.02	affinity constant c6	Fitted
Ki_empup_midC	mol/m3(cell)	0.448	inhibition constant c6	Fitted
K_empup_atp	mol/m3(cell)	0.025	affinity constant atp	Fitted
vmax_emplow	molmidC/CmolX/h	3.740	max conversion rate from c3 to pyr	Fitted
K_emplow_intC	mol/m3(cell)	121.3	affinity constant c6	Fitted
K_emplow_midC	mol/m3(cell)	0.111	affinity constant c3	Fitted
vmax_tcapyr	molpyr/CmolX/h	0.069	max conversion rate for pyr entering tca	Fitted
K_tcapyr_pyr	mol/m3(cell)	0.702	affinity constant pyruvate	Fitted
K_tcapyr_nad	mol/m3(cell)	0.127	affinity constant nad+	Fitted
vmax_tcaaca	molaca/CmolX/h	0.180	max conversion rate for aca entering tca	Fitted
K_tcaaca_aca	mol/m3(cell)	0.006	affinity constant acetaldehyde	Fitted
K_tcaaca_nad	mol/m3(cell)	0.061	affinity constant nad+	Fitted
vmax_etc	molo2/CmolX/h	0.426	max conversion rate of electron transfer chain	Fitted
K_etc_o2	mol/m3(broth)	0.006	affinity constant oxygen	Fitted
K_etc_nadh	mol/m3(cell)	0.022	affinity constant nadh	Fitted
K_etc_adp	mol/m3(cell)	3.181	affinity constant adp	Fitted
GAM_baseline	molatp/CmolX	2.050	growth related maintenance parameter	(Famili et al., 2003)
v_max_ngam	molatp/CmolX/h	0.028	non-growth related maintenance	(Famili et al., 2003)
K_atpm_atp	mol/m3(cell)	0.009	affinity constant atp	Fitted
vmax_glyc	molglyc/CmolX/h	0.002	max conversion rate from c6 to glycerol	Fitted
K_glyc_midC	mol/m3(cell)	0.001	affinity constant c3	Fitted
K_glyc_nadh	mol/m3(cell)	1.200	affinity constant nadh	Fitted
vmax_pdc	molpyr/CmolX/h	5.465	max conversion rate from pyruvate to acetaldehyde	Fitted
K_pdc_pyr	mol/m3(cell)	4.843	affinity constant pyruvate	Fitted

vmax_adh	moletoh/CmolX/h	1.208	max conversion rate from acetaldehyde to ethanol	(Kesten et al., 2015)
Keq_adh	dimentionless	8319	equilibrium constant for adh	(Kesten et al., 2015)
K_adh_aca	mol/m3(cell)	0.059	affinity constant acetaldehyde	(Kesten et al., 2015)
K_adh_nadh	mol/m3(cell)	0.069	affinity constant nadh	(Kesten et al., 2015)
K_adh_etoh	mol/m3(cell)	54.00	affinity constant ethanol	(Kesten et al., 2015)
K_adh_nad	mol/m3(cell)	0.075	affinity constant nad	(Kesten et al., 2015)
Ktp_etoh	m3/h	123.8	ethanol export parameter	Fitted
vmax_growth	molC/CmolX/h	0.526	max growth rate	Fitted
K_growth_intC	mol/m3(cell)	1.289	affinity constant c6	Fitted
K_growth_pyr	mol/m3(cell)	0.300	affinity constant pyruvate	Fitted
K_growth_atp	mol/m3(cell)	2.050	affinity constant atp	Fitted
Hcc_o2	-	0.030	Henry solubility defined as ca/cg	(Sander, 2015)
MW_cell	-	26.00	molecular weight of cell	assumed
V_cell	ml/gDW	2.500	cell volume	assumed
Total_nadX	mol/m3(cell)	0.800	total available nadh+nad in the cell	(Canelas et al., 2008)
Total_axp	mol/m3(cell)	4.200	total available axp in the cell	(Canelas et al., 2011)

Supplementary: conserved moieties

In the 7-pool kinetic model, 2 conserved moiety pairs were applied, namely ATP-ADP and NADH-NAD^+, which implies that their total amounts were assumed constant and only one of the moieties (in this case, ATP and NADH) was subjected to the kinetics defined in the model.

Taking the ATP-ADP pair as an example, the amount of ATP would normally take around 80% of the total AXP. This leads to different dynamics in applying ATP activation or ADP inhibition:

Assuming:

$$C_{ATP} + C_{ADP} = 1$$

In a normal healthy yeast cell, the energy charge level is around 0.85. From which we can calculate ATP =0.7 and ADP =0.3. This means a stimulus causing a short period of ATP increase at 0.1 (14.3%) will lead to a 0.1 drop of ADP (33%).

Take the following ATP activation and ADP inhibition kinetic as examples:

$$Acti_{ATP} = \frac{C_{ATP}}{C_{ATP} + K_{m,ATP}}$$

$$Inhi_{ADP} = \frac{K_{i,ADP}}{C_{ADP} + K_{i,ADP}}$$

We assign $K_{m,ATP} = 0.7$ and $K_{i,ADP} = 0.3$ so that for the reference state both terms are equal to 0.5. In the above stimulus condition, the ATP term will increase from 0.5 to 0.53 (+6% more activation) while the ADP term will increase from 0.5 to 0.6 (+20% less inhibition).

Therefore, applying the ADP inhibition instead of ATP activation, with a proper value of the kinetic constant, can realize the same level of rate change with less metabolite fluctuation. Essentially, one could expect a more stable ATP profile but still provide a sufficient dynamic role in the kinetic model.

Supplementary: ATP paradox after single glucose pulse

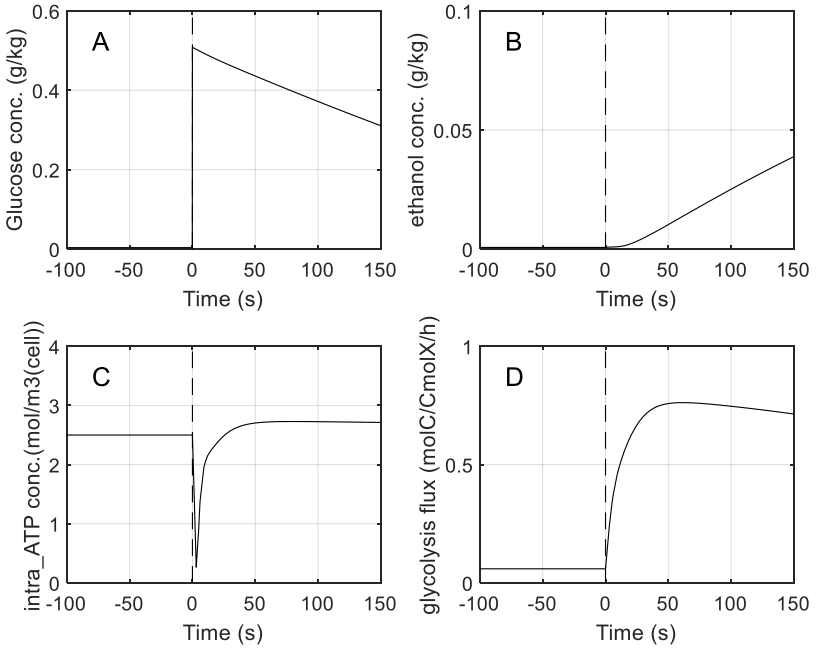

Figure S1 Simulated ATP paradox during a glucose pulse on 7-pool kinetic model (with a reduced Vmax,glcupt at 279 mmol.CmolX^{-1}.h^{-1}).The level before time 0 indicates a steady state with D=0.05 h^{-1}.

Chapter 6

Reflections on further improvements of the yeast kinetic model & Outlook

Introduction

In previous Chapters, we used two industrially relevant model strains, *Penicillium chrysogenum* and *Saccharomyces cerevisiae,* as examples to demonstrate the feasibility and power of the integration of metabolic kinetic models and CFD.

The 9-pool (**Chapter 3**) for *P. chrysogenum* was the first model designed according to the principle of CFD integration from the very beginning. The design of the model largely relied on the pathway interactions knowledge from well-developed intracellular metabolite quantifications and metabolic flux estimations (de Jonge et al., 2014; Nasution et al., 2008) as well as existing kinetic models (Deshmukh et al., 2015; Douma et al., 2010). The model was able to successfully describe all experimental data, including the productivity loss during oscillation conditions. Nevertheless, stiff nodes remain to be adapted to achieve a successful and stable two-way integration in a CFD framework (**Chapter 4**).

The modelling work for the other strain, *S.cerevisiae*, was targeted at incorporating an interesting natural phenotype of this yeast: the Crabtree effect. First, we paid a closer look at the storage carbohydrate cycle, which proved that to properly model the dynamical process, a certain degree of complexity is required (i.e. enzyme level regulation). The central metabolism of yeast was then modelled in a similar way as the 9-pool model for *P.chrysogenum*. Key metabolites/nodes were extracted and an additional conserved moiety, the NAD^+/NADH pair, was included as trigger for a smooth shift from respiration to fermentation. In the development of both yeast models, the model fitness for integration with CFD was assessed via a one-way coupling to lifelines from a 22 m^3 tank fermentation process (Haringa et al., 2017b).

During the analysis of the yeast model, we spotted several shortcomings in the whole yeast model (**Chapter 5**) and acknowledge the still missing storage kinetics (**Chapter 4**). I here take the opportunity for the first part of this Chapter to propose and design an extended yeast kinetic model structure which is more promising regarding dynamic response predictions in all previous conditions, but yet compact enough to allow efficient CFD simulation. The proposed model structure thus fits our goal of integration into a full CFD framework. In the second part, we look back to the path we went through, highlighting two important aspects to bear in mind before starting any modeling construction work: “begin with the end in mind” and “fit for purpose”, which, I believe, are the determining factors of making a useful model. We also foresee a future for

model-guided smart biomanufacturing, via the integration of the traditional factory (real world) and model-based 'digital twins' (digital world). We are convinced that this is the way towards a better-controlled, optimized, efficient and smart bioprocess industry.

Proposal for an improved kinetic model structure for *Saccharomyces cerevisiae*

The 7-pool model for *S. cerevisiae* already showed its capability in reproducing major kinetic responses in both short- and long-term scenarios. In addition, it also proved to be stable when tested against highly dynamic glucose concentration profiles simulated for the 22 m^3 pilot scale tank. All of this demonstrates that it is a suitable yeast kinetic model for integration in a full-scale CFD simulation framework and can serve as a tool for evaluating industrial scale fermentation performance. Nevertheless, several issues were identified as structural defects that could not be solved by improving the parameter fitting quality.

In this outlook Chapter, we propose possible solutions to these issues by extending the yeast model with more intracellular details. Beginning with the end in mind, we propose to keep a relatively simple model structure but to include 1) a clearer classification of central metabolism, including two kinetic mechanisms for glucose uptake featured by high and low glucose affinities; 2) a full capability to produce and re-consume byproducts, e.g. glycerol, acetate and ethanol; 3) incorporation of storage carbohydrate synthesis and consumption as developed in **Chapter 4**; 4) 8 enzyme pools to describe cellular adaptation at higher specific growth rates, using acetate or glycerol as carbon source and storage dynamics at different time scales. The proposed yeast model still has all the features from the 7-pool yeast model, such as the trigger of ethanol formation as well as the concept of lean biomass. Figure 1 shows a schematic representation of the improved, 21-pool yeast kinetic model (11 metabolite pools, 2 pairs of conserved moieties, 8 enzyme pools). Table S1 and Table S2 explain all intermediates and conversions that have been included in the proposed yeast model and Table S3 contains the stoichiometry matrix of the model. The rationale and details of the structured model are explained in the following parts.

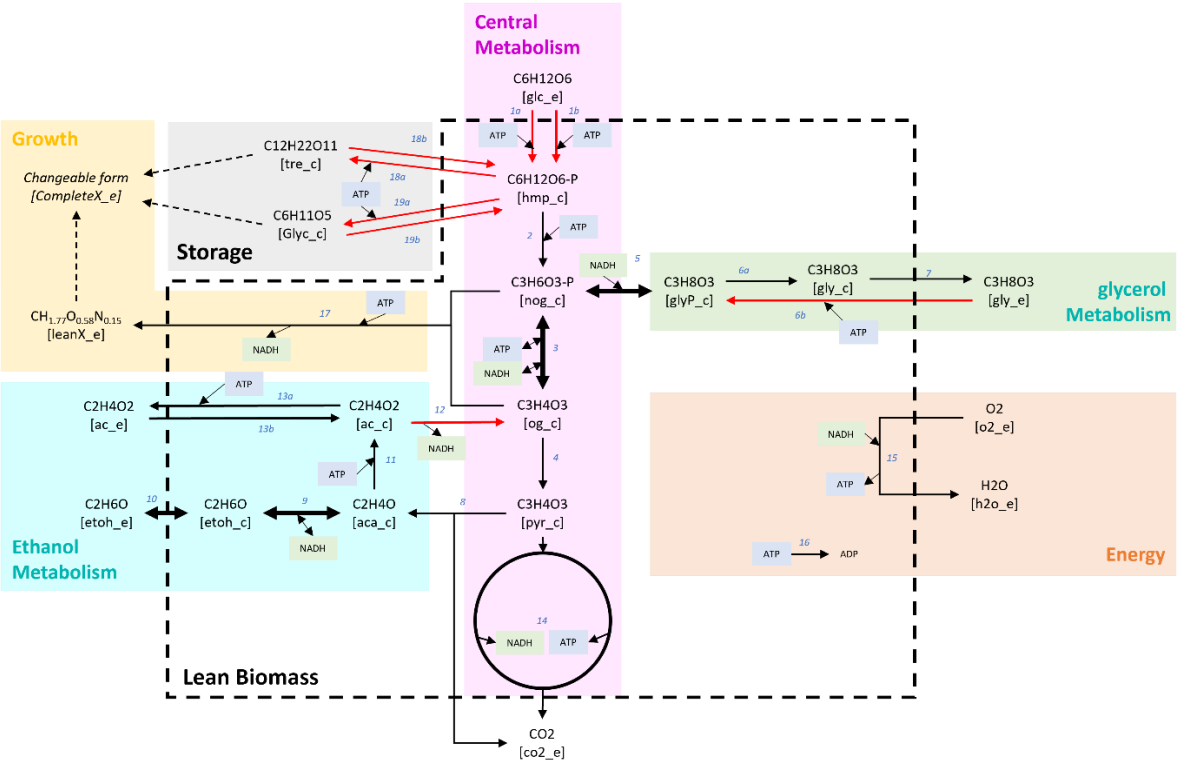

Figure 1 Schematic of the proposed yeast full-cell kinetic model structure. The model summarizes 6 metabolic network modules at high level and introduced the concept of lean biomass, where the complete biomass requires the lean biomass added with storage carbon consisting of trehalose and glycogen. The red arrow stands for conversions with gene regulation. The bold arrow stands for reversible reactions of which the direction is determined by their rate law. Reaction ID with 'a' and 'b' denote conservations interacting with the same compounds but in different directions, following different rate laws.

Glucose uptake

It is known that *Saccharomyces cerevisiae* has multiple hexose transport (HXT) systems to accommodate for different growth conditions(Maier et al., 2002; Ozcan & Johnston, 1999). From the behavior of the 7-pool model under steady state and dynamic conditions it was inferred that a single glucose uptake mechanism was not sufficient to capture the glucose uptake kinetics across the full range of specific growth rates (Chapter 5). As the first step, we here propose to incorporate two types of hexose transport systems to improve the yeast kinetic model. Referring to the HXT expression profile (Maier et al., 2002), the high affinity HXT ($K_m \approx 1$ mM) is expressed at low glucose concentrations but repressed at high glucose concentrations. In contrast, the low glucose affinity transporter ($K_m > 10$ mM) is expected to be activated only at high glucose concentrations (Ozcan & Johnston, 1999). The turnover times of the high and low affinity transporters are assumed to be about one half to several hours. This feature is embedded in two gene-regulated kinetic expressions (e_1a and e_1b) via balancing of the synthesis rate, degradation rate and growth dilution term for both transport proteins.

The uptake process is further lumped with the first phosphorylation step in the glycolysis and thus produces hexose monophosphate (**hmp_c**), representing glucose-6-phosphate (G6P) and fructose-6-phosphate (F6P). One ATP is consumed for the lumped hexokinase and the redox level remains unchanged.

Central carbon metabolism (Glycolysis and TCA cycle)

The central carbon metabolism is responsible for catabolism of glucose and provides metabolites for storage carbohydrate synthesis, biomass formation, and energy production. In the 7-pool model, a discrepancy between the measured levels of the individual intermediates G6P, F6P and FBP and their lumped pool (intC) was observed at specific growth rates beyond the critical value (Chapter 5). Here, we propose the following new structure for modeling the central carbon metabolism:

1. Non-oxidative part. This represents the first half of glycolysis, resulting in a lumped reaction including phosphofructokinase (PFK), fructose bisphosphate aldolase (ALDO) and triosephosphate isomerase (TPI). Therefore, one ATP is consumed for the kinase (PFK step) and the product of this lumped step is called **nog_c** (non-oxidized glycolytic intermediate). This lumped pool consists of fructose-1,6-biphosphate (FBP), glyceraldehyde 3-phosphate (G3P) and dihydroxyacetone

phosphate (DHAP). The lumped pool is represented as a three-carbon unit with the same redox level as glucose.

2. Oxidative part. The second part of glycolysis includes oxidation via glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and ATP production via phosphoglycerate kinase (PGK), which are both reversible reactions. This lumped step includes also the other two reversible steps: phosphoglycerate mutase (PGM) and enolase (ENO). All lumped steps are biochemically reversible which makes this lumped step reversible too. NADH is produced in this lumped lower part of glycolysis. The product is called **og_c** (oxidized glycolytic intermediates) and includes 1,3-biphosphoglycerate (1,3-BPG), 3-phosphoglycerate (3PG), 2-phosphoglycerate (2PG) and phosphoenolpyruvate (PEP).
3. Pyruvate kinase. In the 7-pool model, the allosteric regulation of pyruvate kinase (PK) by intC ensured a good prediction of midC dynamics during both short- and long-term dynamic conditions (Chapter 5). This feature will be kept in the proposed model by explicitly modeling the PK step using nog_c as allosteric regulator. This is an irreversible step and produces pyruvate (**pyr_c**) from og_c together with one ATP.
4. Tricarboxylic acid cycle. Because we are not aiming at incorporating all details of the central metabolic pathways and the model application doesn't require specific intermediate(s) of the TCA cycle, we decided to lump the whole cycle into one reaction and neglect the compartmentation. This results in one reaction for the full oxidization of pyruvate into carbon dioxide, producing 5 NADH and 1 ATP per molecule of pyruvate. Similar as for the 7-pool model, FADH2 and GTP are considered equivalent to respectively NADH and ATP for simplicity purposes.

Glycerol metabolism

Glycerol formation is not coupled to ethanol formation due to the fact that the fermentation of glucose to ethanol is NADH neutral (Nordström, 1966). Under anaerobic conditions, NADH produced from anabolism needs to be recycled via glycerol-3-phosphate dehydrogenase (GPDH) (Geertman et al., 2006).

One of the main shortcomings of the 7-pool model is an overestimation of the glycerol production at high growth rates and in lifeline simulations. To correct for this, the capacity to consume glycerol as carbon source is required for this model. Four reactions can be lumped to represent the whole glycerol metabolism as a branch of central carbon metabolism from nog_c (Figure 1, green section). Glycerol-3-phosphate (G3P) is the first intermediate which is

reduced from nog_c via GPDH. It is known that GPDH consumes NADH for reducing DHAP to G3P while reducing quinone when converting G3P back to DHAP. In this model, the process was simplified as a reversible conversion where the reducing agent is always the NADH/NAD pair. The produced G3P is sequentially converted to glycerol via phosphoglycerate phosphatase (PGP). When the cell starts using glycerol as carbon source, glycerol kinase (GK) is activated and phosphorylates the intracellular glycerol instead. This step is a gene-regulated step and the expression of GK is inhibited by the presence of extracellular glucose and limited by low oxygen availability (Klein et al., 2017). During the fermentation of glucose, glycerol is produced and diffused passively via the Fps1p channel (Oliveira et al., 2003).However, when glucose is depleted and secreted glycerol is being reused, the *STL1* gene will be induced and encode an H^+ symporter for active glycerol uptake (Ferreira et al., 2005). For simplification purpose, we would suggest lumping the induction of *STL1* and GK and therefore transportation and phosphorylation of glycerol will be modelled in one step, at a cost of 2 ATP per glycerol.

Ethanol formation

The 7-pool kinetic model provided a good prediction of ethanol production under carbon overflow conditions and oxygen limitation (Chapter 5). Therefore, we would propose to keep the same ethanol production pathway as in the 7-pool model with similar kinetics. In short, the designed model will still use the NADH as the trigger for stimulating ethanol production (Vemuri et al., 2007). The irreversible pyruvate decarboxylase reaction (PDC) will be modelled as a single conversion step and alchohol dehydrogenase (ADH) acts as a reversible step that either reduces acetaldehyde or oxidizes ethanol using NADH or NAD as cofactor. The exchange flux of ethanol follows simple diffusion without energy costs (Kotyk & Alonso, 1985).

Ethanol uptake and gluconeogenesis

When using C2 carbon sources, e.g. ethanol or acetate, the cell requires the gluconeogenesis pathway (de Jong-Gubbels et al., 1995). In this model, part of the gluconeogenesis (acetate to PEP) is lumped into one reaction which consists of the glyoxylate cycle from acetate (which is oxidized from ethanol via ADH and ALDH) and phosphoenolpyruvate carboxykinase (PEPCK). These conversions will end in phosphoenolpyruvate (PEP) and join the central metabolism (detailed stoichiometry of single steps shown in Table 1). Due to the lumping of central carbon metabolism in this model, we treat the product of this part of the conversions, phosphoenolpyruvate, as og_c. In the absence of glucose, the

og_c from ethanol can go to nog_c for biomass synthesis and to pyruvate for energy formation. In this way the whole model is capable of describing growth using ethanol or acetate as the sole carbon source.

The lumped gluconeogenesis reaction is a gene-regulated conversion as the key enzyme, PEPCK, is known to be rapidly inactivated by the presence of glucose, fructose and mannose (de Jong-Gubbels et al., 1995; Gancedo & Schwerzmann, 1976), irreversibly. To reproduce such rapid inhibition dynamics, a high protein degradation rate activated by the presence of glucose is required, next to a constant protein degradation rate and dilution by cell division.

Table 1 Stoichiometry of glyoxylate cycle and PEPCK. All reactions can be lumped into one in the proposed model

Reactions[1] / Metabolites	ACS[2]	CS & ACONT	ICL	MS	SDH & FH	MDH	PEPCK	SUM
Acetate	-2							**-2**
Acetyl-CoA	2	-1		-1				
Isocitrate		1	-1					
Glyoxylate			1	-1				
Succinate			1		-1			
Malate				1	1	-2		
Oxaloacetate		-1				2	-1	
Phosphoenolpyruvate							1	**1**
H_2O		-1		-1	-1			**-3**
CO_2							1	**1**
CoA-SH	-2	1		1				
NADH					1	2		**3**
ATP	-4						-1	**-5**

[1] ACS: Acetyl-CoA synthase; CS: citrate synthase; ACONT: Aconitase; ICL: isocitrate dehydrogenase; MS: malate synthase; SDH: succinate dehydrogenase; FH: fumarate hydratase; MDH: Malate dehydrogenase; PEPCK: phosphoenolpyruvate carboxykinase

[2] Producing 1 Acetyl-CoA via ACS will consume 1 ATP and produce 1 AMP which is equivalent to consuming 2 ATP to 2 ADP in this model

Acetate transport

Different from the transport of ethanol which is driven by simple diffusion, acetate, as a weak acid, requires more attention in cross-membrane transport. Casal and coworkers (Casal et al., 2016) reported that undissociated carboxylic acids can cross the plasma membrane by passive diffusion and more specfically, acetic acid can also be transported by facilitated diffusion via the Fps1 Channel (Mollapour & Piper, 2007). Other transporters, for example Jen1 and Ady2 can import anionic forms of these substrates by a H^+-symport mechanism and Pdr12, an ATP-binding cassette (ABC) transporter, can export the anionic form of acetate which prevails at the physiological pH of the cytoplasm (Casal et al., 2016; Holyoak et al., 1999; Verduyn et al., 1990).

Following those reports, we suggest passive diffusion of the undissociated acetic acid into the cells but an ATP consuming process in pumping acetate out. This simplification is structurally similar to the transport kinetics for (PAA) in the *P. chrysogenum* 9-pool model (Deshmukh et al., 2015; Tang et al., 2017).

Electron transfer chain and maintenance energy requirements

As the prediction of the respiration and the energy balance of the 7-pool model were satisfactory, we suggest maintaining the structure and kinetics of the electron transfer chain and maintenance in the 7-pool model. The initial value of the P/O ratio (ATP production per electron pair) would be set at 1 (Famili et al., 2003; Verduyn et al., 1991) which is close to the optimized value (0.95 molATP/molO) for the 7-pool model.

The initial guess of the non-growth associated maintenance energy rate (NGAM) was 0.028 molATP.CmolX^{-1}.h^{-1}, referring to the 7-pool model. This value is in agreement with the previously reported value of 0.026 molATP.CmolX^{-1}.h^{-1} (Famili et al., 2003) but higher than the value (0.017 molATP.CmolX^{-1}.h^{-1}) determined from near-zero growth experiments in an aerobic retentostat (Vos et al., 2016). The growth associated maintenance (GAM) is defined as ATP consumption for producing one mole lean biomass. The fitted value for the 7-pool model was 2.05 molATP.CmolX^{-1}, which is close to the reported 1.83 molATP.CmolX-1 (Famili et al., 2003) in the constraint-based genome-scale metabolic model analysis. We will adopt the value determined by the 7-pool model as initial guess for the proposed model.

Growth (lean biomass)

Because the biomass storage carbohydrate content (here we refer to trehalose and glycogen) can change dramatically at different growth rates (Suarez-Mendez, 2015), we propose the biomass of the extended model to be divided in two parts, namely lean biomass and storage carbohydrates, similar to the model presented in Chapter 4.

The lean biomass is basically referring to the biomass without any intracellular trehalose and glycogen content. This is similar to the definition of 'active biomass' by (Aboka et al., 2012). At a specific growth rate of 0.1 h^{-1}, a complete yeast cell contains 12.4%(w/w) of trehalose and glycogen and has a molar composition of $CH_{1.761}O_{0.581}N_{0.155}$. The composition of biomass after eliminating the storage carbohydrates (i.e. lean biomass) then becomes $CH_{1.773}O_{0.548}N_{0.176}$. In the proposed extended yeast kinetic model, we will only use the lean biomass formula in the simulation. The molar composition and the weight of complete

biomass will be subject to the prediction of trehalose and glycogen content. It needs to be mentioned that when calibrating and validating the model predictions against experimental data, any specific rates and intracellular concentrations should be re-calculated based on the complete biomass rather than the rates directly coming out of the kinetic model. All kinetics were set up based on lean biomass (active biomass) with units of mol.CmolX(lean)$^{-1}$.h^{-1}.

Storage carbohydrates

The kinetics of the synthesis and degradation of storage carbohydrates, that is trehalose and glycogen, were separately developed and calibrated in Chapter 4. It was inferred that to predict the dynamics of those two storage compounds during both long-term steady state conditions and short-term perturbations, all four steps of conversions required gene-regulation based enzyme kinetics. We propose to incorporate the storage kinetics developed in Chapter 4 into this model, but some parameters may require further optimization.

Further steps with the extended yeast model

The re-designed yeast kinetic model featured with more intracellular components is expected to be able to capture more complex cell responses under various conditions. The proposed adaptations should solve the structural defects of the kinetic model described in Chapter 5. Furthermore, it should contain storage carbohydrate synthesis and degradation as described in Chapter 4. However, the increased number of kinetic expressions and intermediates also means an increased demand for experimental data for model training, which should cover all necessary parts of the lumped network, including, but not limited to:

- Ethanol/acetate formation and (co-)consumption
- Glycerol production and (co-) consumption
- Storage carbon profile beyond critical growth rates
- Short-term response and long-term adaption of glucose transporters
- Ideally, quantitative proteomics data for PEPCK related enzyme and glycerol metabolism (esp. GK)

To produce reliable data on these parts of the metabolism requires well-designed experiments and a high-quality analytical platform, let alone all the remaining *in-scilio* work like parameter estimation, parameter confidence interval (CI) calculation, sensitivity analysis, cross validation, and full-scale simulation integrated with CFD.

So far, we proposed improvements to the 7-pool yeast model which has been constructed and validated in Chapter 5, to solve certain shortcomings. Before actually constructing the model and designing experiments to provide data for parameter estimation questions should be raised such as:

- Does this “expensive” model really fit our purpose?
- Is it necessary to solve our problem with such an “expensive” model?

In the next section we will address these questions in a more general way, as they are crucial to make proper choices in the development, construction, and further optimization of mathematical models for (fermentation) process development.

“Begin with the end in mind” and “fit for purpose”

The field of developing a kinetic model to describe the behavior of microorganisms is very broad and diverse, and highly dependent on the purpose of the model. Simple black box models based on hyperbolic glucose uptake kinetics and the Herbert-Pirt equation for substrate consumption together with the proper material balances can already provide guidance in solving many engineering problems (Haringa et al., 2017a). This simple type of kinetic models can also be a baseline for more dedicated pathway kinetics (Douma et al., 2010). On the other side of the spectrum there are also research works targeting at a detailed representation of microbial metabolism, e.g. using genome-scale models (dFBA) (Vargas et al., 2011) or via a complex kinetic models (Kesten et al., 2015; Smallbone & Mendes, 2013). Such detailed approaches can help to understand the detailed dynamic responses of the cells to stimuli: figuring out the mechanisms behind the cell’s phenotype, identifying the engineering target for strain improvements, etc.

Because the number of possible variants of kinetic metabolic models and their level of detail is infinite, it is crucial to ‘begin with the end in mind’, where ‘begin’ can even refer to the conceptual phase of a project. All three models (one for *P.chrysogenum* and two for *S.cerevisiae*) presented in this thesis, and the improved version of the *S. cerevisiae* model proposed above, were developed with the purpose of integrating them into a CFD simulation framework, trying to solve the environmental heterogeneity problem at the industrial scale and provide guidance for the design of representative scale-down simulators (Table 2). Some key features of reaction kinetic models to be combined within a CFD framework are:

1. The model takes environmental variables (concentrations) as input and provides cellular responses (rates) including but not limited to substrate uptake, biomass growth and product formation.
2. A relatively simple and robust model structure that will not add heavy computational burden when integrated into CFD simulation platforms.
3. Capability of reproducing dynamics of the host as a result of different environmental perturbations at different time scales. The perturbation is defined quantitatively from the specific industrial scale bioprocess under study.

The last point largely determines the structure and complexity of the model. For example, **Chapter 2** and **3** present a compact mathematical model for growth and product formation in *P.chrysogenum* which has as main purpose to describe the dynamics of the formation of one specific product, penicillin G (PenG). The model has a relatively simple central metabolism (one carbon pool only) but was extended with a lumped amino acid pool which is considered as the precursor of PenG. The kinetics of the transport of the side chain precursor phenylacetic acid (PAA) and a gene regulation model for the PenG biosynthesis pathway were obtained from previous publications to ensure a proper kinetic performance of PenG production. In **Chapter 4**, a model was developed aimed at reproducing the cellular dynamic response of storage metabolism of *S. cerevisiae*. The central metabolism was therefore highly simplified. In this specific case, directly adopting a black box model for substrate uptake and cell growth appeared sufficient. In **Chapter 5**, another yeast kinetic model was designed but for a completely different purpose. This model was expected to reflect a more complete yeast physiology, by reproducing the ethanol production pattern under various conditions (Crabtree effect and Pasteur effect) at different time scales. It was also required for this model to trigger ethanol production with the proper metabolic mechanism (in our case, we used NADH as trigger) rather than mathematically mimic the ethanol production profile. To reach this target, the central metabolism of the yeast was described at a certain resolution. But still, in the design of this model we also paid attention to the first and second requirement for combination with CFD modeling, stated above. This ensured a successful application of this model with a highly dynamic lifeline family and proves its fitness to CFD integration.

The conceptional phase of developing a kinetic model is the first important step. A clear definition of the problem to be solved is required and one needs to be able to further translate it into a proper metabolic network structure that

highlights the major conflicts in the problem. The next step is to secure that the model is fit for its purpose.

One example of model fitness for purpose is the adaption of glucose uptake kinetics in **Chapter 5**. We first tried one single Michaelis-Menten kinetic equation to describe the glucose uptake pattern under steady state conditions throughout the whole range of specific growth rates. The kinetics, despite some underestimations on the residual glucose concentration at low growth rates, could describe the overall phenotype good enough. However, a closer look at the responses of the cells during short-term perturbations suggested that at least two different types of glucose uptake mechanisms were required, depending on the residual glucose concentration. One solution is to equip the model with two different kinetic mechanisms for glucose uptake. Moreover, their capacity (i.e., the expression of corresponding transporter) should be regulated at the gene level. This solution is the most 'physiologically correct' one and should ensure a proper cell response when shifting from low to high glucose concentrations and back. Nevertheless, this solution will introduce a total of six kinetic equations (2 for glucose uptake and 4 for gene regulation of the transporters, which is exactly proposed in the above yeast kinetic model). Additional data, ideally quantitative proteomics data on different types of transporters, is needed to properly fit the parameters of these equations. A temporary solution, applied in **Chapter 5,** is a straightforward adaption of V_{max} based on the glucose uptake rate that makes the glucose uptake fit to only a low growth rate range. This adaptation compromised the model's predictivity (or more precisely, the long-term adaptation of the cells) at very high growth rates or glucose excess conditions. But still, it gives a robust and simple solution to increasing the predictability of the model under single and periodic glucose pulse conditions, without increasing the degree of freedom of the model and additional challenges in parameter fitting. It needs to be mentioned that it should be checked if this adaptation and hypothesis remain correct for the 'final' simulation on the large scale and if not, it indicates that not only the application boundary needs extension, but also raises doubts whether the scale-down design (feast-famine) is a good representation of large-scale gradients.

Table 2 An overview of all four models developed/proposed in this thesis with their key technical features

Chapter	model name	Central metabolism	Biomass growth	conserved moieties	product/byproduct formation	enzyme regulations	CFD integration
2	9-pool	1 glycolytic pool 1 amino acid pool	unified biomass	ATP-ADP	Penicillin synthesis PAA transportation	glucose transportation Penicillin formation storage synthesis storage degradation	two-way coupling after model adaption (**Chapter 3**)
4	storage model	black-box (no pool)	lean biomass variable storage carbon	n.a.[1]	storage regulation - metabolite level - enzyme level	trehalose synthesis trehalose degradation glycogen synthesis glycogen degradation	one-way coupling validated
5	yeast model _v1	3 glycolytic pools	unified biomass	ATP-ADP NAD^+-NADH	ethanol glycerol	n.a. [1]	one-way coupling validated
6 (This Chapter)	yeast model _v2[2]	4 glycolytic pools	lean biomass variable storage carbon	ATP-ADP NAD^+-NADH	ethanol glycerol acetate storage regulation - metabolite level - enzyme level	glucose transportation trehalose synthesis trehalose degradation glycogen synthesis glycogen degradation gluconeogenesis glycerol utilization	n.a. [1]

[1] n.a.: not available

[2] conceptional phase

An era of digital twins (DTs)

The concept of creating a virtual representation of real objects was first publicly introduced in 2002. The first practical definition of 'digital twins' was launched at NASA later in 2010. It originated from a model but was extended extensively for describing/monitoring/controlling a whole system. Digital twins (DTs) are believed to be the ideal tool for the rapid and cost-effective development, realization and optimization of control and automation strategies for bioprocesses. They may be utilized for the development and implementation of conventional controllers (e.g. temperature, dissolved oxygen, etc.), for monitoring non-measurable biological indexes, i.e. using soft sensors (Kadlec et al., 2009), as well as for advanced control strategies (e.g. control of substrate or metabolite concentrations, multivariable controls), and the development of complete bioprocess control (Appl et al., 2021). Ultimately, the DTs leads to the smart biomanufacturing under the scope of industry 4.0.

Mathematical process models are central and basic components of DTs (Moser André et al., 2020). The model should be capable of describing and predicting the microorganism's cultivation process, including product formation, with high fidelity. Udugama and coworkers (Udugama et al., 2020) highlighted 5 levels of digital twin application in industry, from basic steady state models to DT-assisted process control. A validated model (such as the models developed in this thesis) is positioned at level 3 (validated kinetic model) and "digital shadow" is lying between this level and the final realization of DTs. As an intermediate step, a fairly broadly accepted definition of digital shadow is that, even though digital shadows share a lot of properties with DTs, the dataflow of digital shadows is only one-way (Gargalo et al., 2020). This is further being explained as the capability of real-time model re-parameterization and curation using the real-time dataflow from the actual bioprocess. Unlike the forward simulation using a validated kinetic model, the digital shadow can fuse the data from the two worlds and predict the development of the target process with higher accuracy. This builds a solid foundation towards the final step: DTs, where data flows are communicated in two-ways between physical instances and its digital twin representation such that the real process is always running under the most optimal control supervised and forecasted by the DT. This thesis provides several model examples for *Penicillium chrysogenum* and *Saccharomyces cerevisiae* to demonstrate the feasibility, the value, and the challenges in constructing the core part of DTs. We are convinced that model-guided, DT-supported bioprocess control, problem diagnosis, optimization and innovation

will be the future for a smarter bio-industry. This thesis can be viewed as an early, but vital step towards this vision.

During our development of the mechanistic-based kinetic models in this thesis, we realized that there are constraints when using pure mechanistic kinetic models to describe everything we observed, especially when the phenotype of the microorganism is clear while the underlying mechanism is complicated, hypothesized without validation or even remains unknown. An example is that, in **Chapter 4**, we tried to reproduce the fast/slow responses of carbohydrate storage. We applied a total of 6 kinetic equations and fitted 11 kinetic parameters for just describing the trehalose concentration profile. Although we explained in **Chapter 4** that those are all built on top of our inference and necessity to mimic cellular regulation (the actual regulation mechanism can only be more complex), we are still in doubt whether it is still feasible when more conversions in the interested network require this level of complexity and may largely increase the computation burden of the whole model, especially when integrating into a CFD framework. An alternative way to tackle the system's complexity is the application of a hybrid model, i.e., the synthesis of data-driven and mechanistic models.

Lopez and coworkers (Lopez et al., 2020) showed a simple DT example via integration of a mechanistic and a partial-least-squares (PLS) model. This hybrid model was used to forecast the cellulose fermentation process with complex interactions, including potential inhibitory effects emerging from lignocellulosic feedstocks. The PLS model is used to estimate the glucose concentration from spectroscopic data and its output feeds the mechanistic model to predict long-horizon glucose, xylose and ethanol profiles. The data-driven model, including a machine learning model, requires a large amount of data to calibrate and is particularly suitable for solving problems in a complex system that lacks thoughtful understanding. In such a way, the hybrid model can take the advantages of both mechanistic models and data-driven models and represent the most suitable 'high-fidelity' model set-up (Gargalo et al., 2020), fulfill the requirement of DTs and maintain the computational burden of DTs at an acceptable level.

Other efforts are being made to reduce the computational burden of DTs via the simplification of hydrodynamic simulation (Tajsoleiman et al., 2019) and novel ways of parcel tracking in simplified compartment models (Haringa et al., 2022). The later methodology uses the adapted 9-pool model of *P.chrysogenum* (**Chapter 3**) as an example and demonstrated the possibility of solving 80h of

flow time within just a few minutes, way faster than a full Euler-Lagrange CFD simulation that usually takes weeks. The fact that the solving time (~several minutes) is much shorter the real process time (~80h) enables further model applications, like model parameter adaptation during process, real-time quality check or even inline real-time process optimization, etc.

Other than a 'hi-fi' mathematical model as the core of DT system, its implementation in industry requires more support elements, for example, a mature digitalization infrastructure. As mentioned above, the dataflow towards digital shadows and digital twins is not necessarily limited to the live data stream from available online measurements but may also include various data types such as in-line, at-line, and offline datasets, or even manual records (e.g. batch number of seed, raw materials) upon request. This can be particularly true for those DT systems using hybrid/data-driven models as the core model. Besides, the scope of digital shadow/twin models in biomanufacturing is not limited to monitoring/controlling the bioprocess but can further be extended to upstream (e.g. seed cultivation/preparation, raw material quality, robust strain re-design, etc.) and downstream processing (e.g. separation, purification, etc.). A more ambitious scope is to take logistics, scheduling of unit operations, and material quantity/quality management also into account. In such a way, various models present a digital replica at the whole plant level which eventually enables the realization of industry 4.0 smart biomanufacturing (Son et al., 2022).

References

Aboka, F. O., van Winden, W. A., Reginald, M. M., van Gulik, W. M., van de Berg, M., Oudshoorn, A., & Heijnen, J. J. (2012). Identification of informative metabolic responses using a minibioreactor: a small step change in the glucose supply rate creates a large metabolic response in Saccharomyces cerevisiae. *Yeast*, *29*(3–4), 95–110. https://doi.org/doi:10.1002/yea.2892

Appl, C., Moser, A., Baganz, F., & Hass, V. C. (2021). Digital Twins for Bioprocess Control Strategy Development and Realisation. In C. Herwig, R. Pörtner, & J. Möller (Eds.), *Digital Twins: Applications to the Design and Optimization of Bioprocesses* (pp. 63–94). Springer International Publishing. https://doi.org/10.1007/10_2020_151

Casal, M., Queirós, O., Talaia, G., Ribas, D., & Paiva, S. (2016). Carboxylic acids plasma membrane transporters in saccharomyces cerevisiae. In *Advances in Experimental Medicine and Biology* (Vol. 892, pp. 229–251). https://doi.org/10.1007/978-3-319-25304-6_9

de Jonge, L., Buijs, N. A. A., Heijnen, J. J., van Gulik, W. M., Abate, A., & Wahl, S. A. (2014). Flux response of glycolysis and storage metabolism during rapid feast/famine conditions in Penicillium chrysogenum using dynamic 13C labeling. *Biotechnology Journal*, *9*(3), 372–385.

de Jong-Gubbels, P., Vanrolleghem, P., Heijnen, S., van Dijken, J. P., & Pronk, J. T. (1995). Regulation of carbon metabolism in chemostat cultures of Saccharomyces cerevisiae grown on mixtures of glucose and ethanol. *Yeast*, *11*(5), 407–418. https://doi.org/https://doi.org/10.1002/yea.320110503

Deshmukh, A. T., Verheijen, P. J. T., Maleki Seifar, R., Heijnen, J. J., & van Gulik, W. M. (2015). In vivo kinetic analysis of the penicillin biosynthesis pathway using PAA stimulus response experiments. *Metabolic Engineering*, *32*, 155–173. https://doi.org/http://dx.doi.org/10.1016/j.ymben.2015.09.018

Douma, R. D., Verheijen, P. J. T., de Laat, W. T. A. M., Heijnen, J. J., & van Gulik, W. M. (2010). Dynamic gene expression regulation model for growth and penicillin production in Penicillium chrysogenum. *Biotechnology and Bioengineering*, *106*(4), 608–618.

Famili, I., Forster, J., Nielsen, J., & Palsson, B. O. (2003). Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(23), 13134–13139. https://doi.org/10.1073/pnas.2235812100

Ferreira, C., Frank Van Voorst, †, Nio Martins, A., Neves, L., Oliveira, R., Kielland-Brandt, M. C., Lucas, C., & Brandt, A. (2005). A Member of the Sugar Transporter Family, Stl1p Is the Glycerol/H Symporter in Saccharomyces cerevisiae. *Molecular Biology of the Cell*, *16*, 2068–2076. https://doi.org/10.1091/mbc.E04

Gancedo, C., & Schwerzmann, K. (1976). Inactivation by Glucose of Phosphoenolpyruvate Carboxykinase from Saccharomyces cerevisiae *. In *Arch. Microbiol* (Vol. 109).

Gargalo, C. L., de las Heras, Simoneta Caño Jones, Mark Nicholas Udugama, Isuru Mansouri, Seyed Soheil, Krühne, U., & Gernaey, K. v. (2020). Towards the Development of Digital Twins for the Bio-manufacturing Industry. In *Digital Twins Tools and Concepts for Smart Biomanufacturing* (Vol. 176, pp. 1–34). Springer, Cham. http://www.springer.com/series/10

Geertman, J.-M. A., van Dijken, J. P., & Pronk, J. T. (2006). Engineering NADH metabolism in Saccharomyces cerevisiae: formate as an electron donor for glycerol production by anaerobic, glucose-limited chemostat cultures. *FEMS Yeast Research*, *6*(8), 1193–1203. https://doi.org/10.1111/j.1567-1364.2006.00124.x

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017a). Euler-Lagrange analysis towards representative down-scaling of a 22 m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669.

Haringa, C., Deshmukh, A. T., Mudde, R. F., & Noorman, H. J. (2017b). Euler-Lagrange analysis towards representative down-scaling of a 22m3 aerobic S. cerevisiae fermentation. *Chemical Engineering Science*, *170*, 653–669. https://doi.org/https://doi.org/10.1016/j.ces.2017.01.014

Haringa, C., Tang, W., & Noorman, H. J. (2022). Stochastic parcel tracking in an Euler-Lagrange compartment model for fast simulation of fermentation processes. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.28094

Holyoak, C. D., Bracey, D., Piper, P. W., Kuchler, K., & Coote, P. J. (1999). The Saccharomyces cerevisiae weak-acid-inducible ABC transporter Pdr12 transports fluorescein and preservative anions from the cytosol by an energy- dependent mechanism. *Journal of Bacteriology*, *181*(15), 4644–4652. https://doi.org/10.1128/jb.181.15.4644-4652.1999

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven Soft Sensors in the process industry. In *Computers and Chemical Engineering* (Vol. 33, Issue 4, pp. 795–814). https://doi.org/10.1016/j.compchemeng.2008.12.012

Kesten, D., Kummer, U., Sahle, S., & Hübner, K. (2015). A new model for the aerobic metabolism of yeast allows the detailed analysis of the metabolic regulation during glucose pulse. *Biophysical Chemistry*, *206*, 40–57. https://doi.org/https://doi.org/10.1016/j.bpc.2015.06.010

Klein, M., Swinnen, S., Thevelein, J. M., & Nevoigt, E. (2017). Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities. In *Environmental Microbiology* (Vol. 19, Issue 3, pp. 878–893). Blackwell Publishing Ltd. https://doi.org/10.1111/1462-2920.13617

Kotyk, A., & Alonso, A. (1985). Transport of ethanol in baker's yeast. *Folia Microbiologica*, *30*(1), 90–91. https://doi.org/10.1007/BF02922503

Lopez, P. C., Udugama, I. A., Thomsen, S. T., Roslander, C., Junicke, H., Mauricio-Iglesias, M., & Gernaey, K. v. (2020). Towards a digital twin: a hybrid data-driven and mechanistic digital shadow to forecast the evolution of lignocellulosic fermentation. *Biofuels, Bioproducts and Biorefining*, *14*(5), 1046–1060. https://doi.org/10.1002/bbb.2108

Maier, A., Völker, B., Boles, E., & Fuhrmann, G. F. (2002). Characterisation of glucose transport in Saccharomyces cerevisiae with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Res*, *2*(4), 539–550. https://doi.org/10.1111/j.1567-1364.2002.tb00121.x

Mollapour, M., & Piper, P. W. (2007). Hog1 Mitogen-Activated Protein Kinase Phosphorylation Targets the Yeast Fps1 Aquaglyceroporin for Endocytosis, Thereby Rendering Cells Resistant to Acetic Acid. *Molecular and Cellular Biology*, *27*(18), 6446–6456. https://doi.org/10.1128/mcb.02205-06

Moser André, Appl Christian, Brüning Simone, & Hass Volker C. (2020). Mechanistic Mathematical Models as a Basis for Digital Twins. In *Digital Twins Tools and Concepts for Smart Biomanufacturing* (Vol. 176, pp. 133–180). http://www.springer.com/series/10

Nasution, U., van Gulik, W. M., Ras, C., Proell, A., & Heijnen, J. J. (2008). A metabolome study of the steady-state relation between central metabolism, amino acid biosynthesis and penicillin production in Penicillium chrysogenum. *Metabolic Engineering*, *10*(1), 10–23.

Nordström, K. (1966). Yeast growth and glycerol formation. *Acta Chem Scand*, *20*(4), 1016–1025.

Oliveira, R., Lages, F., Silva-Graça, M., & Lucas, C. (2003). Fps1p channel is the mediator of the major part of glycerol passive diffusion in Saccharomyces cerevisiae: Artefacts and re-definitions. *Biochimica et Biophysica Acta - Biomembranes*, *1613*(1–2), 57–71. https://doi.org/10.1016/S0005-2736(03)00138-X

Ozcan, S., & Johnston, M. (1999). Function and regulation of yeast hexose transporters. *Microbiology and Molecular Biology Reviews : MMBR*, *63*(3), 554–569. https://pubmed.ncbi.nlm.nih.gov/10477308

Smallbone, K., & Mendes, P. (2013). Large-Scale Metabolic Models: From Reconstruction to Differential Equations. *Industrial Biotechnology*, *9*(4), 179–184. https://doi.org/10.1089/ind.2013.0003

Son, Y. H., Kim, G. Y., Kim, H. C., Jun, C., & Noh, S. do. (2022). Past, present, and future research of digital twin for smart manufacturing. In *Journal of Computational Design and Engineering* (Vol. 9, Issue 1, pp. 1–23). Oxford University Press. https://doi.org/10.1093/jcde/qwab067

Suarez-Mendez, C. a. (2015). *Dynamics of Storage Carbohydrates Metabolism in Saccharomyces cerevisiae: A Quantitative* (Issue december).

Tajsoleiman, T., Spann, R., Bach, C., Gernaey, K. v, Huusom, J. K., & Krühne, U. (2019). A CFD based automatic method for compartment model development. *Computers & Chemical Engineering*, *123*, 236–245.

Tang, W., Deshmukh, A. T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J. J., Xia, J., Chu, J., & others. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by Penicillium chrysogenum. *Biotechnology and Bioengineering*, *114*(8), 1733–1743.

Udugama, I. A., Gargalo, C. L., Yamashita, Y., Taube, M. A., Palazoglu, A., Young, B. R., Gernaey, K. v., Kulahci, M., & Bayer, C. (2020). The Role of Big Data in Industrial (Bio)chemical Process Operations. *Industrial and Engineering Chemistry Research*, *59*(34), 15283–15297. https://doi.org/10.1021/acs.iecr.0c01872

Vargas, F. A., Pizarro, F., Pérez-Correa, J. R., & Agosin, E. (2011). Expanding a dynamic flux balance model of yeast fermentation to genome-scale. *BMC Systems Biology*, *5*. https://doi.org/10.1186/1752-0509-5-75

Vemuri, G. N., Eiteman, M. A., McEwen, J. E., Olsson, L., & Nielsen, J. (2007). Increasing NADH oxidation reduces overflow metabolism in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences*, *104*(7), 2402–2407.

Verduyn, C., Postma, E., Scheffers, W. A., & van Dijken, J. P. (1990). Physiology of Saccharomyces cerevisiae in anaerobic glucose-limited chemostat cultures. *Journal of General Microbiology*, *136*(3), 395–403. https://doi.org/10.1099/00221287-136-3-395

Verduyn, C., Stouthamer, A. H., Scheffers, W. A., & van Dijken, J. P. (1991). A theoretical evaluation of growth yields of yeasts. In *Antonie van Leeuwenhoek* (Vol. 59). Kluwer Academic Publishers.

Vos, T., Hakkaart, X. D. V., Hulster, E. A. F., Maris, A. J. A., Pronk, J. T., & Daran-Lapujade, P. (2016). Maintenance-energy requirements and robustness of Saccharomyces cerevisiae at aerobic near-zero specific growth rates. *Microbial Cell Factories*, *15*(1), 1–20. https://doi.org/10.1186/s12934-016-0501-z

Supplementary

Table S1 List of metabolites/enzymes of the proposed yeast full-cell kinetic model

No.	Metabolites or Enzymes[1]	Formula[2]	Remarks
1	hmp_c	$C_6H_{12}O_6$-P	intracellular hexose mono phosphate, including: G6P, F6P
2	nog_c	$C_3H_6O_3$-P	non-oxidized glycolysis intermediates, including: FBP, GAP, DHAP
3	og_c	$C_3H_4O_3$-P	oxidized glycolysis intermediates, including: BPG, 3PG, 2PG, PEP
4	pyr_c	$C_3H_4O_3$	intracellular pyruvate
5	glyP_c	$C_3H_8O_3$-P	glycerol-3-phosphate
6	gly_c	$C_3H_8O_3$	intracellular glycerol
7	aca_c	C_2H_4O	intracellular acetaldehyde
8	etoh_c	C_2H_6O	intracellular ethanol
9	ac_c	$C_2H_4O_2$	intracellular acetate
10	tre_c	$C_{12}H_{22}O_{11}$	intracellular trehalose
11	glyc_c	$C_6H_{10}O_5$	intracellular glycogen
12	e_1a	n.a.	low affinity glucose transporter
13	e_1b	n.a.	high affinity glucose transporter
14	e_6b	n.a.	glycerol kinase
15	e_11	n.a.	acetaldehyde dehydrogenase
16	e_18a	n.a.	lumped trehalose synthesis kinetics
17	e_18b	n.a.	lumped trehalose degradation kinetics
18	e_19a	n.a.	lumped glycogen synthesis kinetics
19	e_19b	n.a.	lumped glycogen degradation kinetics
20	atp_c	AP3	adenosine triphosphate (ATP), conserved moiety
21	nadh_c	DH2	NADH and H^+, conserved moiety
22	nad_c	D	NAD^+, conserved moiety
23	adp_c	AP2	adenosine diphosphate (ADP), conserved moiety
24	p_c	P	phosphate (neutral), ignored in balances
25	glc_e	$C_6H_{12}O_6$	extracellular glucose
26	etoh_e	C_2H_6O	extracellular ethanol
27	ac_e	$C_2H_4O_2$	extracellular acetate
28	gly_e	$C_3H_8O_3$	extracellular glycerol
29	nh4_e	NH_4^+	ammonia, charged
30	leanX_e	$CH_{1.77}O_{0.58}N_{0.15}$	lean biomass, excluding storage compounds
31	o2_e	O_2	oxygen
32	co2_e	CO_2	carbon dioxide
33	h2o_e	H_2O	water, ignored in balances
34	h_e	H^+	proton, charged, ignored in balances

[1] "_c" and "_e" refer to intracellular and extracellular compounds, respectively.

[2] n.a.: not applicable.

Table S2 Reactions included in the proposed yeast full-cell kinetic model

Reaction ID	Remarks[1]
v_1a	glucose uptake with low-affinity glucose transporter (uniport), e.g. HXT1, lumped with HK which is presumed to be at pseudo steady state
v_1b	glucose uptake with high-affinity glucose transporter (H-symport), e.g. HXT2, lumped with HK which is presumed to be at pseudo steady state
v_2	non-oxidative part of glycolysis (upper part), including: PGI, PFK, ALDO, TPI
v_3	lower part of glycolysis (oxidative), including: GAPDH, PGK, PGM, ENO
v_4	PK
v_5	GPDH (cytosolic and mitochondrial)
v_6a	PGP
v_6b	GK
v_7	glycerol transportation
v_8	PDC
v_9	ADH
v_10	ethanol transportation
v_11	ALDH
v_12	lumped reaction that allows ethanol entering central metabolism via glyoxylate cycle + PEPCK
v_13a	ac import (undissociated passive Fps1)
v_13b	ac export (ABC transporter Pdr12)
v_14	TCA
v_15	ETC
v_16	lumped reaction for lean biomass formation
v_17	maintenance as ATP dissipation
v_18a	trehalose formation
v_18b	trehalose degradation
v_19a	glycogen formation
v_19b	glycogen degradation
vE_1a_syn	synthesis rate of low-affinity glucose transporter
vE_1a_deg	degradation rate of low-affinity glucose transporter
vE_1b_syn	synthesis rate of high-affinity glucose transporter
vE_1b_deg	degradation rate of high-affinity glucose transporter
vE_6b_syn	synthesis rate of GK
vE_6b_deg	degradation rate of GK
vE_11_syn	synthesis rate of ALDH
vE_11_syn	degradation rate of ALDH
vE_18a_syn	synthesis rate of trehalose synthesis capacity
vE_18a_deg	degradation rate of trehalose synthesis capacity
vE_18b_syn	synthesis rate of trehalose degradation capacity
vE_18b_deg	degradation rate of trehalose degradation capacity
vE_19a_syn	synthesis rate of glycogen synthesis capacity
vE_19a_deg	degradation rate of glycogen synthesis capacity
vE_19b_syn	synthesis rate of glycogen degradation capacity
vE_19b_deg	degradation rate of glycogen degradation capacity

[1] HK: hexose kinase; PGI: glucose-6-phosphate isomerase; PFK: phosphofructokinase; ALDO: frutose-bisphosphate aldolase; TPI: triosephophate isomerase; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; PGP: phosphoglycerate phosphatase; PGM: phosphoglycerate mutase; ENO: phosphopyruvate hydratase(enolase); PK: pyruvate kinase; GPDH: glycerol-3-phosphate dehydrogenase; PGP: glycerol-3-phosphate phosphatase; GK: glycerol kinase; PDC: pyruvate decarboxylase; ADH: alcohol dehydrogenase; ALDH: acetaldehyde dehydrogenase; PEPCK: phosphoenolpyruvate carboxykinase; TCA: tricarboxylic acid cycle; ETC: electron transport chain

Table S3 Stoichiometry of the full-cell yeast kinetic model

		v_1a	v_1b	v_2	v_3	v_4	v_5	v_6a	v_6b	v_7	v_8	v_9	v_10	v_11	v_12	v_13a	v_13b	v_14	v_15	v_16	v_17	v_18a	v_18b	v_19a	v_19b
intracellular	hmp_c	1	1	-1																		-2	2	-1	1
	nog_c			2	-1		-1													-0.2					
	og_c				1	-1									1					-0.26					
	pyr_c					1					-1							-1							
	glyP_c						1	-1	1																
	gly_c							1		-1															
	aca_c										1	-1		-1											
	etoh_c											1	-1												
	ac_c													1	-2	1	-1								
	tre_c																					1	-1		
	glyc_c																							1	-1
	adp_c	1	2	1	-1	-1			2						5		2	-1	**-2.5**	**1.54**	1	1	2	1	1
	atp_c	-1	-2	-1	1	1			-2						-5		-2	1	**2.5**	**-1.54**	-1	-1	-2	-1	-1
	nadh_c				1		-1					-1		1	3			5	-2	0.42					
	nad_c				-1		1					1		-1	-3			-5	2	-0.42					
	p_c		1		-1			1							4		1	-1	-2.5	2	1	3		2	
extracellular	glc_e	-1	-1																						
	etoh_e												1												
	ac_e															-1	1								
	gly_e								-1	1															
	nh4_e																			-0.15					
	leanX_e																			1					
	o2_e																		-1						
	co2_e										1				1			3		0.38					
	h2o_e														-3			-3	2	0.04		1	-1	1	-1
	h_e																			0.15					

Publications

1. **Tang, W**., Deshmukh, A.T., Haringa, C., Wang, G., van Gulik, W., van Winden, W., Reuss, M., Heijnen, J.J., Xia, J., Chu, J., et al. (2017). A 9-pool metabolic structured kinetic model describing days to seconds dynamics of growth and product formation by *Penicillium chrysogenum*. Biotechnology and Bioengineering 114, 1733-1743.
2. **Tang, W**., Pan, A., Lu, H., Xia, J., Zhuang, Y., Zhang, S., Chu, J., and Noorman, H. (2015). Improvement of glucoamylase production using axial impellers with low power consumption and homogeneous mass transfer. Biochemical Engineering Journal 99, 167-176.
3. Haringa, C., **Tang, W**., Wang, G., Deshmukh, A.T., van Winden, W.A., Chu, J., van Gulik, W.M., Heijnen, J.J., Mudde, R.F., and Noorman, H.J. (2017). Computational fluid dynamics simulation of an industrial *P. chrysogenum* fermentation with a coupled 9-pool metabolic model: towards rational scale-down and design optimization. Chemical Engineering Science.
4. Haringa, C., **Tang, W.**, Deshmukh, A.T., Xia, J., Reuss, M., Heijnen, J.J., Mudde, R.F., and Noorman, H.J. (2016). Euler-Lagrange computational fluid dynamics for (bio)reactor scale down: An analysis of organism lifelines. Engineering in Life Sciences 16, 652-663.
5. Haringa, C., ***Tang, W.***, & Noorman, H. J. (2022). Stochastic parcel tracking in an Euler-Lagrange compartment model for fast simulation of fermentation processes. Biotechnology and Bioengineering. https://doi.org/10.1002/bit.28094
6. Wang, G., **Tang, W.**, Xia, J., Chu, J., Noorman, H., and Gulik, W.M. (2015). Integration of microbial kinetics and fluid dynamics toward model-driven scale-up of industrial bioprocesses. Engineering in Life Sciences 15, 20-29.
7. Wang, G., Haringa, C., **Tang, W.**, Noorman, H., Chu, J., Zhuang, Y., and Zhang, S. (2019). Coupled Metabolic-hydrodynamic Modeling Enabling Rational Scale-up of Industrial Bioprocesses. Biotechnology and Bioengineering n/a.
8. Lu, H., Li, C., **Tang, W.**, Wang, Z., Xia, J., Zhang, S., Zhuang, Y., Chu, J., and Noorman, H. (2015). Dependence of fungal characteristics on seed morphology and shear stress in bioreactors. Bioprocess and Biosystems Engineering 38, 917-928.
9. Wang, G., Zhao, J., Haringa, C., **Tang, W.**, Xia, J., Chu, J., Zhuang, Y., Zhang, S., Deshmukh, A.T., van Gulik, W., et al. (2017). Comparative performance of different scale-down simulators of substrate gradients in Penicillium chrysogenum cultures: the need of a biological systems response analysis. Microbial Biotechnology 11, 486-497.
10. Wang, G., Chu, J., Noorman, H., Xia, J., **Tang, W.**, Zhuang, Y., and Zhang, S. (2014). Prelude to rational scale-up of penicillin production: a scale-down study. Applied Microbiology and Biotechnology 98, 2359-2369.

11. Janoska, A., Verheijen, J. J., ***Tang, W.,*** Lee, Q., Sikkema, B., & van Gulik, W. M. (2023). Influence of oxygen concentration on the metabolism of *Penicillium chrysogenum*. Engineering in Life Sciences 23:e2100139.
12. ***Tang, W.****, Xia, J., Chu, J., Zhuang, Y., and Zhang, S. (2015). Development and application of morphological analysis method in Aspergillus niger fermentation. Chinese journal of biotechnology 31, 291-299.*
13. *Zhuang, Y., Chen, H., Xia, J.,* ***Tang, W.****, and Zhao, Z. (2015). Progress in industrial bioprocess engineering in China. Chinese journal of biotechnology 31, 778-796.*
14. *Li, Y.,* ***Tang, W.****, Bai, C., Mingzhi, H., Chu, J., and Zhuang, Y. (2019). Optimization of the feeding process and metabolism analysis to improve the yield of penicillin. Chinese Journal of Antibiotics 44, 679-686.*

Acknowledgement

The Ph.D. journey I have experienced is a special one. It can be described as a super long journey which started all the way back in 2013 and can also be considered as a 24-month sprint towards the Ph.D. degree from May 2020. No matter what, it is really a pleasant for me to start writing this part of my disseration, indicating the finalization of my Ph.D. trajectory. There are so many people that helped me in the past during my journey and I would like to take this opportunity to share my sincere appreciation to them.

Prof. Henk Noorman. It is said that one's life can be completely changed when he or she meets the right person at the right time. You are definitely the person and literally changed me and my family's lifeline. Not only it gave me the chance to work and collaborate with DSM as early as my master study back in 2011, it also introduced me to a closer collaboration with TU Delft from 2013, encouraged me to apply for the job in DSM and eventually enabled me to start my life in the Netherlands. Moreover, it also provided me the opportunities and challenges to start my Ph.D. journey, while being my promoter to guide me through all the difficulties while I was lost in the balance between my DSM work and Ph.D. tasks. A 'thank you' weighs far too little for all what you have done for me, intentionally or unintendedly, which made me be the one I am right now. Your role is more than just a professor and promoter who is supposed to guide me solving scientific problems, but also a master who inspired me during my life.

Prof. Walter van Gulik. I would like to first thank you for willing to be my supervisor during my Ph.D. trajectory and helped me so much in discussing all the aspects of my work, till the detail of a single parameter, including its biological meaning behind and reference values in previous publications etc. On top of that, I regard your 'supervision' started as early as in 2013. You are THE professor, who helped me design the experiment, and guided me personally in the lab through my first chemostat-fermentation experiment abroad (which was in Sept 2013) with a high fermentation quality. Together with Guan Wang, Xiaoyun Liu, Prof. Jianye Xia, we visited your lab four times and each time you have been such a kind host. The most unforgettable experience was in 2015 and together with Guan Wang, we shared the same office (central position of the old biotechnology building on the Julianalaan) for two whole months. You are the most excellent teacher in guiding us how to execute a high-quality experiment, do properly data postprocessing and make us aware of its importance.

Dr. Cees Haringa. We are (almost) the same age, but I also treat you as my personal development example. During your Ph.D. trajectory, we carried out fermentations together and your excellent work brought the value of the *P.chrysogenum* kinetic model to a new level (and in the following years, the photo where I was performing rapid sampling becoming more well-known to the community). This also enabled that we continued to pursue in this direction by thinking deeper and making higher quality models to realize the real "digital-twin" concept. To some extent, you are my 'client' and your recognition of the kinetic model that I delivered is better than anything. You are the driving force for me to keep improving the model with a clear target. Together with Guan Wang, Xiaoyun Liu and Chao Li and other ECUST colleagues, we had a nice time together in Shanghai and becoming not only a collaboration partner but also my first Dutch friend. Thank you for your nice work and I believe we will keep collaborating in the coming years.

My comrades in the ComRaDes project: Prof. Ralf Takors, Christopher Sarkizi Shams Hajian and Steven Minden from University of Stuttgart, Prof. Frank Delvigne and Mathéo Delvenne from University of Liège and Dr. Peng Wei and Dr. Agnes Janoska from Delft University of Technology. Desipte the fact that I am not an official member of this project consortium, I started being involved in this Public-Private Partnership Project by helping to explain my previous *P. chrysogenum* experiments and extending the 9-pool model. Later I got more involved in developing and sharing the progress of Michaelis-Menten based kinetic models for *S. cerevisiae*. I am really amazed by all the discussions/debates/agreements we had during the project meetings. It is really a pity that we were not able to have a face-to-face meeting shortly after the project started, due to the global pandemic. Nevertheless, I would like to thank all professors, post-doc researchers and Ph.D. candidates in this consortium and valuable discussions that made this dissertation possible.

My TU Delft colleagues: Prof. Sef Heijnen, Dr. Reza Seifar, Cor Ras, Angela ten Pierick, Dirk Geerts, Yi Song. All of you were my entry point to my life abroad, in 2013. Prof. Sef Heijnen, you are one of the most attractive and friendly professors I have ever met. You are actually the one that led me into the world of modeling, and I really like the way you teach others, either in a course or in a discussion. I am really grateful for your insightful suggestions, critical comments, and generous sharing. You are also the one who makes me believe that modeling does not necessarily needs a computer (or even calculator). Dr. Reza Seifar, Cor Ras, Angela ten Pierick, Dirk Geerts, and Yi Song, thank you for

all your supports in setting up the fermentation system and all the efforts of sample analysis. Your hospitality largely relieved my upsets during my visit 11,000km away from my home and helped me quickly get used to a foreign lab during several short visits, without losing a single fermentation experiment over the years!

And the supports from DSM, from my work group, cannot be forgotten. My director and manager, Marieke Overeijnder, Dr. Christopher Bayer and Dr. Joep Schmitz. Thank you for managing the resources that enabled this Ph.D. trajectory. And my colleagues, Dr. Liang Wu, Dr. Wouter van Winden, Dr. Zheng Zhao, Dr. Amit Deshmukh, Dr. Sergio Rossell, thank you all for the help in modeling concepts and insightful discussions/suggestions on the *S.cerevisiae* kinetic model.

Also, I would like to thank all my colleagues working at the East China University of Science and Technology (ECUST), where I obtained my MSc degree and continued my work on the Hé Project. Prof. Siliang Zhang, Prof. Ju Chu, Prof. Yingping Zhuang, Prof. Jianye Xia, your daily supervision, and in-depth discussion helped me entering the world of bioscience. I am so fortunate and glad that I have made the most correct decision and started my biotechnology career in this multi-disciplinary group, to be able to work in the National Center of Bioengineering and Technology (NCBIO). This is the place where I decided to devote myself to biotechnology and serve in industry. Prof. Hongzhong Lu, we together started our first project about the morphology of *Aspergillus niger* in a small hidden lab in ECUST, starting from cleaning up everything. Even though you decided to stay in the academia (and now titled “Prof.”!) while I selected to work in industry, your hard-working attitude and self-discipline still inspires me nowadays. Dr. Guan Wang, Dr. Xiaoyun Liu, you were my closest working mates during the Hé project, the first part of this thesis. Thank you for all the hard work in carrying out complex but precise, high-quality fermentations, cutting-edge analysis works for the project and the enthusiasm in writing.

Last but not least, I would like to share my gratefulness to my parents, my family and my friends. My parents have a clear character of their generation: fight against poverty and had little chance to access higher education. However, like many parents of that generation, based on their life experience, they know better than anyone else about the importance of education. And because of that, my parents have been willing to spend any cost to support me further on my way to pursuing a higher degree, even though they may not have the knowledge to understand what exactly I am doing. I thank my parents and their

selflessness dedication that made me the current me. And, special thanks to my beloved wife, Xiangfen Chen and my son, He Tang. You are the reason I fight for, and you are the ultimate driving force to stimulate me overcome all hinders. Special thanks to my wife for her sacrifice during my Ph.D. trajectory period. Supporting a busy husband is easier to say than do, especially when the tiny family is in shape in a foreign country. My son, without you, ~~I probably can finish this thesis faster~~ (no, this is absolutely not true), I would not be able to enjoy such an interesting and colorful Ph.D. period. Your innocent smile and funny while sharp comments on my work deserve you another bag of candy.

The final appreciation is for you, the reader. Thank you for reading my work and I hope it will inspire you regarding your work and/or study in the future. Or, at least, preventing you step into a dead-end.

Wenjun Tang

2022-11-23

Delft, the Netherlands

Curriculum Vitae

Wenjun Tang was born on August 1988 in Shanghai, China. After he finished his high school education, he starts his BSc program in East China University of Science and Technology (ECUST) in Shanghai, majored in Biotechnology. He received 3 times of scholarship during his bachelor study and finished his bachelor study in 2010.

By the end of his 4-year study, he was recommended for admission to school of biotechnology, ECUST as a postgraduate with an exemption of examation. He is superivised by Prof. Siliang Zhang and Prof. Ju Chu in the State Key Labortary of Bioreactor Engineering. His master study focuses on a collaboration project between the University and Royal DSM, which is about *Aspergillus niger*'s morphology control and its correlation to the glucoamylase production. He developed a method to fix the cell pellet in 3D and an image analysis protocol to determine several KPIs for evaluating pellet features.

After obtaining his master degree in 2013, Wenjun joined the National Center of Bioengineering and Technology (NCBIO) in Shanghai and has the role of researcher, responsible for both academic and industry collaboration projects. During this period, he is one of the main contributors to a national collaboration project (Hé project) between China and the Netherlands, where ECUST, Delft University of Technology (TU Delft), Royal DSM and NCBIO are four collaboration partners. Wenjun focused on developing a structured kinetic for *Penicillum chrysogenum* producing pencillin G and simulate proper kinetic feature of the host at various time scales. The collaboration project finished by the end of 2016.

Since 2017, Wenjun started his career in DSM Biotechnology Center in Delft and now he is a senior associate scientist in the Center for Digital Innovation (CDI) in Biodata andTranslational Sciences (BD&T), DSM. While continuing his career in the Netherlands, he started his Ph.D. trajectory in TU Delft in 2020, under the supervision of Prof. Henk Noorman and Dr. Walter van Guilk. The topic of his Ph.D. trajectory follows his work in the Hé project and extends the structural kinetic modeling workframe to another important strain that widely used in industry: *Saccharomyces cerevisiae*.

Part of the works in the Hé project and follow up research works on *S.cerevisiae* are presented and discussed in this desseration.