Applications of computational modeling in metabolic engineering of yeast

Applications of computational modeling in metabolic engineering of yeast

ARTICLE *in* FEMS YEAST RESEARCH · AUGUST 2014

Impact Factor: 2.82 · DOI: 10.1111/1567-1364.12199

3 AUTHORS, INCLUDING:

Eduard Kerkhoven
Chalmers University of Technology
11 PUBLICATIONS **73** CITATIONS
SEE PROFILE

Petri-Jaan Lahtvee
Chalmers University of Technology
8 PUBLICATIONS **140** CITATIONS
SEE PROFILE

Received Date : 28-May-2014
Revised Date : 19-Aug-2014
Accepted Date : 19-Aug-2014
Article type : MiniReview
Editor : Hal Alper

Applications of computational modeling in metabolic engineering of yeast

Eduard J Kerkhoven[1,*], Petri-Jaan Lahtvee[1,2,*], Jens Nielsen[1,2,3,†]

[1] Department of Chemical and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden
[2] Novo Nordisk Foundation Center for Biosustainability, Chalmers University of Technology, Gothenburg, Sweden
[3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

* These authors contributed equally.

† Corresponding author: nielsenj@chalmers.se

Abstract

Generally, a microorganism's phenotype can be described by its pattern of metabolic fluxes. Although fluxes cannot be measured directly, inference of fluxes is well established. In biotechnology, often, the aim is to increase the capacity of specific fluxes. For this, metabolic engineering methods have been developed and applied extensively. Many of these rely on balancing of intracellular metabolites, redox and energy fluxes, using genome-scale models (GEMs) that in combination with appropriate objective functions and constraints can be used to predict potential gene targets for obtaining a preferred flux distribution. These methods point to strategies for altering gene expression, however, fluxes are often controlled by post-transcriptional events. Moreover, GEMs are usually not taking into account metabolic regulation, thermodynamics and enzyme kinetics. To facilitate metabolic engineering, tools from synthetic biology have emerged, enabling integration and assembly of naturally non-existent, but well characterized components into a living organism. To

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1567-1364.12199

describe these systems kinetic models are often used, and to integrate these systems with the standard metabolic engineering approach it is necessary to expand the modeling of metabolism to consider kinetics of individual processes. This review will give an overview about models available for metabolic engineering of yeast and discusses their applications.

Running head: Computational modeling in metabolic engineering of yeast

One-sentence summary: An overview of different types of computational models and their application in the field of metabolic engineering of yeast.

Keywords: genome-scale model; kinetic model; biotechnology; synthetic biology

Introduction

In nature, evolution takes place partly due to imperfect processes of DNA replication resulting in appearance of mutations followed by natural selection. Although relatively rare (approximately 1 in 50 million), this phenomenon has allowed (micro-)organisms to evolve during millions of years and adapt to varying environmental conditions. Using the same principles, adaptive laboratory evolution has become a valuable method in biotechnology to select desired traits in microorganisms by applying a selective pressure during multiple generations of growth. In addition to the improved phenotypes, *de novo* sequencing of the evolved strains gives valuable fundamental knowledge about the metabolic regulation. This method has been used to evolve strains with higher stress resistance (Cakar et al. 2005; Almario et al. 2013; Avrahami-Moyal et al. 2012; Dhar et al. 2011) or to increase the consumption of non-preferred substrates (Hong et al. 2011; Demeke et al. 2013; Yuan & Zhao 2013). However, adaptive evolution cannot always be directly used to increase the production of a desired metabolite or chemical which is often the aim of modern biotechnological applications. Therefore, together with the developments in whole genome sequencing, the field of metabolic engineering has emerged and rapidly developed within the past 25 years. Metabolic engineering has been defined as *directed modulation of metabolic pathways using methods of recombinant technology for the purpose of overproducing fuels and chemical and pharmaceutical products* (Bailey 1991).

To be able to predict the genome modifications necessary for the desired pathway fine-tuning, good knowledge about pathway patterns and flux values is needed. Based on the genome sequence of a microorganism, the presence of enzymes and functional pathways can be elucidated using bioinformatics. However, measurements of flux rates require relative timely and resource consuming experimental methods with labeled carbon and measurement of its distribution in proteinogenic amino acids (Choi & Antoniewicz 2011; Frick & Wittmann 2005; Blank et al. 2005; Gombert et al. 2001).

These methods for inference of metabolic fluxes have been thoroughly developed and used extensively. Although the results are usually highly informative, the outcome may be biased when the system is not in a steady state; enzymes differentiate between labeled and unlabeled substrates; or due to the incomplete network or imperfections in modeling the cell compartmentalization. Additionally, using the latter method flux distributions can only be collected from the central carbon metabolism and biomass formation pathways. As an alternative, genome scale metabolic models (GEMs) have been created and constrained based modeling methods have been developed to predict intracellular fluxes. These methods are described below in more detail and their applications together with shortcomings are discussed.

While providing a good overview and starting point, the knowledge about intracellular flux distribution may not be enough to design successful strategies to control and modify metabolic flux rates, as fluxes are seldom regulated at the transcription level (Chubukov et al. 2013; Bordel et al. 2010; Daran-Lapujade et al. 2007; Valgepea et al. 2013). Specially, this is the case for central carbon metabolism where enzymes have been demonstrated to be in abundance and working far from their maximal enzyme activity rates but being ready to increase their activities when environmental conditions improve (Oliveira et al. 2012). In those cases enzyme activities are often controlled by allosteric regulation or through post-translational modifications and modifications affecting the transcriptional level may have very limited effect on actual flux rates (Daran-Lapujade et al. 2004; Hauf et al. 2000; Feder & Walser 2005). Additionally the observation that flux patterns can be altered within the second scale indicates that they are not exclusively regulated at the transcriptional level (Link et al. 2013). More and more understanding is gained on allosteric control and post-translational modifications (Oliveira & Sauer 2012; Henriksen et al. 2012; Zhang et al. 2013). These regulatory analyses suggest that the global control over the metabolic regulation is distributed between several layers of metabolism and alternative back-up mechanisms exist to increase the robustness of the cell (Gerosa & Sauer 2011), potentially hindering the designed metabolic engineering approaches.

To improve the throughput and success rate in metabolic engineering approaches, synthetic biology methods have been used to produce efficient and well characterized tools. The field synthetic biology represents more fundamental analysis where the aim is to engineer biological components of the cell to improve their performance. Parallels have been drawn with synthetic organic chemistry, a field that blossomed in the second half of 19th century where synthesis of naturally occurring chemicals was discovered (Yeh & Lim 2007). One of the goals of synthetic biology would be to create a fully synthetic platform cell with minimal maintenance and energy costs as the desired trait. So far synthetic biology is contributing the most for synthetic DNA production and modifications of protein structures (Annaluru et al. 2014; Dymond et al. 2011; Dantas et al. 2006). Additionally, toggle switches, cell oscillators and biosensors for instant feedback of transcript, protein or metabolite

abundances are appearing with the aim to understand and control the regulation of metabolic fluxes (Gardner et al. 2000; Elowitz & Leibler 2000; Zhang et al. 2012; Pfleger et al. 2007; Paddon et al. 2013), but applications of these in industrial production strain is still to be seen. This kind of approach presumes more input from dynamic modeling as constraint-based models lack information about regulatory functions. This is particularly important as kinetic modeling is used extensively in synthetic biology to gain a better understanding of the system, to see what components are responsible for certain dynamic behavior and integration of synthetic biology components into complete metabolic networks therefore will require a better understanding of the kinetics of these systems.

The use of mathematical models for quantitative description of biological systems has been used for long time, but with the development of high-throughput analytical techniques for analysis the cellular transcriptome, proteome, metabolome etc. mathematical modeling has become essential for integration of these data, and today the use of mathematical models for integrative analysis of biological systems is generally referred to as systems biology. Furthermore, the combination of comprehensive mathematical models of metabolism, often referred to as genome-scale metabolic models, with omics type data has shown to be very valuable for identification of novel metabolic engineering targets (Lee et al. 2007; Otero et al. 2013)

Here we review different computational modeling methods in metabolic engineering of *S. cerevisiae*. We explain the principles of constraint-based genome-scale metabolic models and kinetic models, the various ways they can be applied to address metabolic engineering questions and whether they can handle regulatory processes, while highlighting examples of their use in the metabolic engineering of *S. cerevisiae*. Finally, we give our perspectives about where the field is moving.

Yeast GEMs

Description

The development of constraint-based genome scale metabolic models (GEMs) started in the late 1990s after the first whole genome sequences became available. GEMs are comprehensive descriptions of all known metabolic reactions in an organism, based on the available genome sequence. Automated bioinformatics tools have been created for fast reconstruction of metabolic networks for any desired organism for which the genome sequence is available (Agren, Liu, et al. 2013; Schellenberger et al. 2011; Henry et al. 2010; Pitkänen et al. 2014). The reconstructed networks can be converted into mathematical stoichiometric matrices (**S**), where rows represent metabolites and columns represent individual reactions (Figure 1). This matrix will determine the solution space of the metabolic network. Genome scale models are constrained by (i) the stoichiometry of the network; (ii) pre-set upper and lower boundaries for selected reactions (usually substrate uptake reactions, but can

also describe thermodynamics); and (iii) the assumption of a steady state. The latter assumption presumes that all the flux rates and metabolite concentrations are constant over time. This state can be reached in chemostat cultivation experiments and, in theory, during the controlled batch logarithmic growth phase when organisms grow at their maximum specific growth rate and the change of neither substrate concentration nor product formation will affect the growth regulation.

Constraint-based modeling

Due to the large number of reactions, GEMs are underdetermined and additional constraints are necessary to determine the flux distribution. Flux Balance Analysis (FBA) is commonly used for this purpose, where an assumed biological objective is applied in the form of maximizing (or minimizing) a certain flux. This results in a solution in the solution space that satisfies the presumed objective (Figure 1). Selection of correct optimization function has been the subject of investigation for some time as a universal objective for all cells and growth environments seems to be absent. The most commonly used objective functions maximize the specific growth rate, ATP generation or product formation (Schuetz et al. 2007; Burgard & Maranas 2003). However, the selection of the objective function is selected by the user and the result may therefore be biased. Furthermore, a single objective function might not be sufficient, as metabolism seems to operate at an optimal surface in multi-dimensional space which is defined by competing objectives (Schuetz et al. 2012). Methods for selecting objective functions for specific conditions based on the analyzed data have been developed. As an example Biological Objective Solution Search (BOSS) has been created to find the most suitable objective function by minimizing the difference between experimental flux data and *in silico* flux distributions (Gianchandani et al. 2008); or minimalization of metabolic adjustment (MOMA) where metabolic phenotype is approximated by performing distance minimization in flux space (Segre et al. 2002). Additionally, unbiased methods have been developed where the whole solution space is scanned by using random sampling or Markov chain Monte Carlo techniques (Schellenberger et al. 2011; Bordel et al. 2010). These methods use a large number of randomly chosen reactions as objective functions, and therefore results in a distribution of results for each intracellular flux, representing a flux range that is allowed according to given constraints.

To overcome the optimization issues additional constraints can be added to the GEMs, for example thermodynamics and ranges of intracellular metabolite concentrations. Addition of Gibbs free energy constants of the reactions and measurements of metabolite concentrations helps to describe the direction and reversibility of the metabolic reactions and therefore a more feasible solution space is obtained. Such an approach has been used on smaller constraint-based models where, first the reactions reversibilities at different conditions are determined with Flux Variability analysis (FVA), subsequently all thermodynamically feasible reactions are calculated, combined with

intracellular metabolomics data, and finally fluxes are determined using Elementary Flux Modes (EFM; (Jol et al. 2012)). Therefore, EFM is an alternative methodology to FBA where optimization of one selected reaction is not needed.

For another constraint based modeling approach – metabolic flux analysis (MFA) – a fully determined network is generated with an equal number of metabolites and reactions. This allows finding a unique solution without using any optimization algorithms. As MFA models are simplified in terms of network complexity and usually smaller in size compared to GEMs, they lack the global perspective of the flux patterns. At the same time, the combination of MFA with stabile isotope labelling experiments allows more accurate measurements of *in vivo* metabolic fluxes (Gombert et al. 2001; Blank et al. 2005; Frick & Wittmann 2005).

History of the yeast GEMs

After the development of the first smaller scale constraint-based models of yeast (van Gulik & Heijnen 1995; Vanrolleghem et al. 1996; Nissen et al. 1997), the first genome-scale yeast model iFF708 was reported in 2003 (Förster et al. 2003). The model contained 708 metabolic genes associated with 1175 reactions and 733 metabolites, divided between three cellular compartments. Within the following years the model was improved by addition of more compartments (Duarte et al. 2004), removal of dead end reactions (Kuepfer et al. 2005) and addition of fatty acid metabolism (Nookaew et al. 2008). With the increase of contributors and complexity of the models, inconsistency between different models grew, impeding the merger and comparison of different models. To overcome those issues, which could hinder progress, a consensus metabolic network called Yeast 1.0 was created. Yeast 1.0 followed standardized rules, and in particular ensured strict naming conventions for metabolites (Herrgård et al. 2008). Further improvements in terms of increased number of genes and reactions and compartmentalization have been carried out and the most updated information about the models can be found from the web page http://yeast.sf.net. Meanwhile, additional yeast genome-scale models have been constructed that are not part of the yeast consensus model effort (Osterlund et al. 2012).

Applications

GEMs represent a data driven approach where typically the objective is to achieve a global picture of possible flux patterns. FBA is often used for estimating the biotechnological potential of microorganisms and pinpoint genetic manipulations that could improve the performance of a cell. The main applications of FBA could be listed as follows:

- Instructions for metabolic engineering purposes
- Biological interpretation and discovery through contextualizing high throughput data

- Development of a computational framework
- Evolutionary elucidation
- Description of multi-species communities

The first two subjects will be elaborated in the current review. More information about the other applications can be found elsewhere (Osterlund et al. 2012; Lewis et al. 2012).

Instructions for metabolic engineering

GEMs and FBA can be applied to identify potential metabolic engineering strategies, or even select the best performing network among alternatives. There have been several successful examples where genome-scale constraint-based modeling was applied for strain improvement (Table 1). The increasing crude oil price has created a demand for bio-based chemicals and succinic acid, an intermediate chemical for a broad range of applications, represents a lucrative target for biobased production. Through genome-scale metabolic network analysis of *S. cerevisiae in silico* metabolic engineering targets for improved succinic acid production were identified (Otero et al. 2013). Potential deletion targets were identified by using the evolutionary programming method OptGene (Patil et al. 2005). Production was improved by blocking succinic acid conversion into fumarate by knocking out an essential subunit of succinate dehydrogenase and re-directing serine and glycine such that it was derived from glyoxylate. As glyoxylate is derived from the glyoxylate cycle, which also results in production of succinic acid, production of this organic acid was coupled to growth. Followed by adaptive evolution experiments and transcriptional characterization of the evolved strain, the need to overexpress the TCA cycle enzyme ICL1 was identified, and by over-expressing the ICL1 gene a strain was obtained that had a more than 30-fold improvement in succinic acid production. A recent constraint-based modeling study, where the RAVEN toolbox (Agren, Liu, et al. 2013) was used to test all the combinations of single and double deletions while optimizing the growth under a limiting glucose uptake rate, reported an alternative pathway, where by deletion of the inner dicarboxylate mitochondrial transporter (DIC1) it was possible to obtain a low-level succinic acid production at anaerobic conditions (Agren, Otero, et al. 2013).

Alternatively, GEMs and FBA has been used for modeling the protein secretory mechanisms (Feizi et al. 2013). A genome-scale secretory network was constructed, although using a bottom-up approach rather than directly from the genome annotation. Being connected with different cellular processes the model allowed integration of genome-scale datasets. The model can assist to design engineering strategies for high level protein production by (i) determine the secretory class the protein of interest belongs to; (ii) predicting the metabolic demands of secretory mechanism and (iii) estimating catalytic activities of the machinery components.

Biological interpretation and discovery through contextualizing high throughput data

To acquire more comprehensive understanding of the flux regulation mechanisms and to detect the potential network interactions that likely determine the fluxes, global datasets can be overlaid with the changes in fluxes as estimated from the GEMs. For example, firstly, integration of transcriptomics data into constraint-based models is to increase the predictive capability of GEMs (various published methods are compared in Machado & Herrgård, 2014). Secondly, the comparison of flux change and gene expression allows identification of fluxes that could be regulated at the transcriptional level. In contrary to this, poor correlations between flux rates and transcriptional changes indicate other regulation levels, like post-transcriptional, post-translational or allosteric control (Bordel et al. 2010). Adding the global scale data from proteome analysis which has recently appeared mainly for single conditions would help to give an additional layer to distinguish between post-transcriptional and post-translational regulation levels. Comparison of global proteome and flux data have been carried out for *E. coli* and *B. subtilis* (Valgepea et al. 2013; Buescher et al. 2012; Ishii et al. 2007; Kohlstedt et al. 2014; Peebo et al. 2014), but is currently still carried out only in small scale for yeast (Oliveira et al. 2012; Fendt et al. 2010) .

Recently, Zelezniak *et al.* (2014) used genome-wide mRNA fold changes from three perturbation studies in *S. cerevisiae* to estimate metabolite fold changes. The estimated metabolite changes were subsequently compared to measured metabolite fold changes and it was found that higher correlations were found when a larger part of the metabolic network was taken into account for the estimation. It will be interesting to see whether inclusion of protein abundance or activity data in this analysis would improve the observed correlations.

Challenges

Although GEMs are widely used as standard tools for data analysis and interpretation, they encounter a number of challenges and limitations. One limitation is related to the required steady state of the system. Often, industrial processes are run in fed-batch or batch mode where the cultures are not strictly in a steady state. Additionally, stressful conditions such as high temperatures, high concentrations or organic acids and osmotic stress that are of high industrial interest can perturb the system, but these adaptations are typically out of the scope of general genome-scale constraint-based modeling. Although it is usually presumed that the homeostasis at the level of metabolites will be established within seconds, accompanied changes in transcriptome and protein levels may take much longer and therefore introduce a significant error.

While GEMs can predict the best performing engineering strategies based on the metabolic network, the application of those changes may not be as straight forward, as metabolic networks are highly regulated at various levels to ensure the robustness of the functional system. Predictions of

allosteric regulation, functional state of enzymes (post-transcriptional modifications) are currently not taken into account in genome-scale models. Additionally, enzyme kinetics and performance of nonnative genes/proteins are currently not covered in GEMs. Although the first genome-scale and reduced constraint-based models that take regulation into account have emerged for *E. coli* (Kotte et al. 2010; Covert et al. 2001) they have so far only been created for certain regulatory systems present in yeast (Christensen et al. 2009). Although the lack of regulatory predicting power in GEMs hinders the quality of predictions for metabolic engineering purposes, the combination of modeling data with real experiments can precisely identify the regions which are highly regulated at post-transcriptional level and allows one to investigate those regions more thoroughly using bottom-up approaches (described below).

Kinetic modeling

GEMs are top-down, kinetic modeling is bottom-up

GEMs follow a top-down approach of systems biology, where information from the different layers of the system (genome, transcriptome, metabolome etc.) are integrated to construct a comprehensive view of metabolism, and insight is gained by studying the behavior of this system as a whole (Bruggeman & Westerhoff 2007). Contrastingly, a bottom-up approach assumes that the behavior of a system can be deduced from the characteristics of its constitutive parts (Bruggeman & Westerhoff 2007). The interactions of these parts give rise to the behavior of the system, and therefore a good understanding of the parts of the system is paramount. Models of metabolism that follow this paradigm are called kinetic models, and the constitutive parts of this system are the enzymes, their metabolites, other effectors, and their interactions. The separation of models of metabolism between top-down and bottom-up is not completely black and white as it is now also possible to construct kinetic models on a genomic scale, as detailed below.

History

Kinetic modeling of metabolism has essentially evolved out of the fields of molecular biology and enzymology, moving from characterizing one enzyme in isolation to investigating them in the biological context of a metabolic pathway. Fundamentals were laid in the 1960s and 70s, through the development of biochemical systems theory (BST) and metabolic control analysis (MCA), which are discussed in more detail below. Since then, different parts of yeast metabolism have been subjected to kinetic modeling, but focus has been on glycolysis.

The structure of a kinetic model

The structure of a kinetic model, but also the process of generating a new kinetic model, can be divided into three parts which will be discussed here briefly: (i) description of the structure of the

metabolic pathway; (ii) description of kinetic rate expressions of the metabolic reactions; and (iii) parameterization of said rate expressions. An overview of this process is shown in Figure 2, while kinetic models in metabolic engineering have recently been reviewed in detail by Almquist *et al.* (2014).

(i) The network structure (stoichiometry) of the metabolic pathway to be modeled is described in the **S**-matrix, similar as in GEMs (Figure 2). Kinetic models typically only included one or a few pathways, in contrast to GEMs, primarily due to the difficulty of identifying the large number of model parameters that would appear in comprehensive models. Yeast metabolic pathways that have been modeled include sphingolipid and ergosterol metabolism (Alvarez-Vasquez et al. 2004, 2011), xylose utilization (Parachin et al. 2011), the pentose phosphate pathway (Vaseghi et al. 1999; Messiha et al. 2014), but most focus has been on modeling glycolysis (Smallbone et al. 2013; Hynne et al. 2001; Galazzo & Bailey 1990; Teusink et al. 2000; van Eunen et al. 2012)).

(ii) The metabolic reactions from (i) that are represented as the rows of the **S**-matrix are described using kinetic rate expressions. The detail of these rate expressions can vary significantly from elementary mass-action kinetics, describing the association and disassociation of each metabolite with the enzyme-metabolite complex, to familiar approximations as Michaelis-Menten kinetics and Hill equations. Alternatively, more abstract rate expressions can be used to describe the enzyme kinetics, typically with the argument that the kinetics for most of the enzymes are unknown or complex and unreliable, such that further abstraction of the rate equations does not signify a loss of information while reducing the complexity of the model. As more abstract rate expressions typically contain less parameters, inductive parameterization (discussed below) also becomes easier to perform (Heijnen 2005). Examples of these approximative rate expressions are generalized mass action and S-systems (both extensively reviewed by Voit, 2013), log-linear (Hatzimanikatis & Bailey 1996), lin-log (Visser & Heijnen 2003), convenience kinetics (Liebermeister & Klipp 2006) and modular rate laws (Liebermeister et al. 2010). These abstract rate expressions are often not able to fully describe the kinetic behavior of enzymes: some can only handle non-zero metabolite concentrations, while others do not demonstrate the typical saturation that is observed when metabolite concentrations are high (Heijnen 2005). Instead of using deterministic equations one can also describe the kinetics using stochastic rate expressions. However, the number of metabolites in yeast metabolism are typically high enough (more than 10^2–10^3 as a rule of thumb (Chen et al. 2010)) that fluctuations due to stochasticity are negligible.

(iii) Parameter values need to be given for the rate expressions described in (ii). The combination of rate equations and parameters forms the *v*-vector (Figure 2). Parameter values of rate equations, such as K_m and v_{max} values, have been determined experimentally one-by-one from cell-

free extracts for over 100 years. One can perform these measurements themselves or obtain kinetic parameters from literature, with databases that collect these values, such as BRENDA (Schomburg *et al.*, 2013; 888 unique EC-number associated reactions for *S. cerevisiae*, with many duplicates of the same reactions) and SabioRK (Wittig *et al.*, 2012; 115 reactions for *S. cerevisiae*). Whether kinetic parameters are measured experimentally or obtained from literature, it is important that these values have been determined in relevant conditions. Eunen *et al.* (2012) updated a kinetic model of yeast glycolysis with enzyme kinetics that had been measured in an *in vivo*-like buffer, and only results from the improved model were in correspondence with various steady-state and dynamic conditions.

Parameter values can alternatively be induced by measuring other variables of the system (Mendes & Kell 1998). By reversing the bottom-up assertion that a system's behavior emerges from interactions of its constitutive parts, detailed knowledge of the behavior of a metabolic pathway can be used to induce the kinetics of the enzymes. The systemic behavior that can be measured and used for inductive parameter estimation includes the dynamics of metabolite concentrations over time, or upon perturbation of the system. Not only can multiple parameters be estimated simultaneously, parameter estimation has the additional advantage that experimental data can be derived from *in vivo* measurements, where the enzymes are in their native environment. However, kinetic models are often too complex and over-parameterized for parameter estimation to assign unique values to each parameter (Gutenkunst et al. 2007). This problem is somewhat reduced by using more abstract rate expressions, which typically require less parameters.

Uncertainty of parameter values

The parameter values that are used in the rate equations are single, precise values. However, the knowledge of these values is often uncertain, due to measurement errors or lack of data. Also, parameter estimation methods often do not provide singular values, but rather distributions of possible parameter values. This uncertainty can be taken into account by generating a set of models, each with a different set of parameter values (Achcar et al. 2012). Assuming even more uncertainty, an ensemble of models can be described where each model has a different combination of parameter values (Tran et al. 2008). This ensemble can be narrowed down by only keeping a subset that is capable to reproduce experimental data, *i.e.* subset that fits with experimental data. In this sense, ensemble modeling is comparable to parameter estimation of the whole metabolic network, but ensemble modeling also has the ability to take alternative stoichiometries into account.

Size and modularity of the model

Kinetic models are typically covering only one or a few metabolic pathways. The complexity of these models increases significantly when the scope is enlarged to include more pathways, resulting in more rate expressions and more parameter values that need to be measured or estimated. Kinetic

models therefore typically only investigate one particular pathway, outside of the context of the whole metabolic network. Efforts have been made to generate genome-scale kinetic models. Smallbone *et al*. (2010) generated a genome-scale kinetic model of yeast metabolism using linlog-kinetics. While this model heavily depends on estimation due to absence of parameter values, it does demonstrate a proof-of-principle of genome-scale kinetic models. As some reactions are more important than others in the regulation of metabolic flux (see Metabolic Control Analysis below), it might not be necessary to model all the reactions in great detail. This principle has been used in a pipeline to generate a genome-scale model of kinetics where only the reactions that are the most important have their kinetics described (Stanford et al. 2013).

As an intermediary to genome-scale kinetic modeling, existing kinetic models can be extended in a modular fashion, where each module can represent a certain pathway, biological function or level of regulation. It has been proposed that such a modular approach can be used to build a SiliconCell (Snoep et al. 2006), as the bottom-up philosophy dictates that integration of functional sub-networks will result in a network with emerging properties from the interactions of the individual sub-networks. In a very interesting example, Klipp *et al*. (2005) modelled the response of yeast to osmotic shock by integrating modules that described osmotic sensing (a phosphorelay system and MAP kinase cascade are involved in signal transduction), gene expression (glycerol kinase and glycerol-3-phosphate dehydrogenase are upregulated) and metabolism (glycerol production is induced).

Applications of kinetic modeling

The combination of the **S**-matrix and the *v*-vector that describe the kinetic model results in a set of ordinary differential equations (ODEs) that describe the dynamics of the metabolite (dX) over time (dt) (Figure 2). These models can subsequently be used in a variety of ways to address different questions regarding metabolism, such as how to increase or reduce the yield of a particular (by-)product (Table 1). In the light of metabolic engineering, of particular interest is the analysis of flux control, while these models can also be used for the evaluation of different engineering strategies. Two theoretical frameworks can be used to investigate kinetic models:

- Metabolic control analysis
- Biochemical systems theory

These two frameworks will be briefly discussed below.

Metabolic Control Analysis (MCA)

The enzymes that make up a metabolic pathway all have a certain level of control over the metabolic flux. As metabolic engineering is often focused on increasing or decreasing a particular

flux, it is of interest to know which enzymes control the flux. The distribution of this control can be studied using the concept of MCA (Kacser & Burns 1973, 1981, 1979), reviewed by Moreno-Sánchez *et al.* (2008), which provides the means to quantify the control by calculating flux control coefficients. A metabolic pathway typically contains many non-zero control coefficients, which means that the control of the flux is shared over the different enzymes rather than one rate-limiting enzyme (Kacser & Burns 1979). It is likely, however, that the values of some of the control coefficients will be larger than those of the other enzymes. This indicates that those reactions are having a larger control over the metabolic flux and suggest potential targets for metabolic engineering.

MCA has been used by Cronwright *et al*. (2002), who produced a small model of glycerol synthesis in *S. cerevisiae* using both measured and published parameter values for glycerol 3-phosphate dehydrogenase (GPD) and glycerol 3-phosphatase. *S. cerevisiae* produces glycerol as a major by-product of ethanol fermentation, and a decreased glycerol production flux would be of industrial interest. MCA analysis of this model showed that GPD exerts the majority of the control over the glycerol synthesis flux, with a flux control coefficient of 0.85, suggesting this as a potential target for metabolic engineering. Unfortunately these results were not evaluated experimentally.

MCA has also showed that the glucose transporter has most of the control over the glycolytic flux in suspended cells (Galazzo & Bailey 1990) and that the flux through the pentose phosphate pathway is highly controlled by the growth rate and oxidative stress (Messiha et al. 2014). Interestingly, Smallbone *et al*. (2013) used MCA to guide how a model of yeast glycolysis should be refined: in iterative rounds the reaction that had the highest control over the glucose uptake and which kinetics had not been described in detail were characterized, resulting in a model that improved by each round.

Biochemical Systems Theory (BST)

BST is a mathematical framework for modeling of complete metabolic pathways that was introduced by Savageau (1969). Within this framework the rate of each reaction is defined as power law expressions, and in these expressions the rate is a function of the product of the metabolite concentrations imposed to a certain power. Parameters in the rate expressions are a rate constant and the power coefficients. Despite the empirical nature of the rate expression used in BST, this modeling approach has shown to efficiently describe the kinetics of complex metabolic pathways, e.g. spingolipid metabolism of yeast (Alvarez-Vasquez, 2004). An interesting aspect of the BST is that power coefficients in the rate expressions are identical to the so-called elasticity coefficients of MCA and there is hereby a link between these two approaches to analyze the kinetics of metabolic pathways. The elasticity coefficients of MCA specify the sensitivity of the individual reaction rates to each of the metabolites in the pathway, and hence the power coefficients have a direct biochemical

interpretation. Alvarez-Vasquez (2004) generated three models of yeast sphingolipid metabolism, one with Michaelis-Menten kinetics and the other two with two types of power law rate expressions according to BST, and it was found that the power law abstractions were able to show the same dynamic behavior compared to a model based on detailed Michaelis-Menten kinetics, but with fewer fitted parameters. Polisetty *et al.* (2008) generated a BST model of glycolysis and demonstrated *in silico* that increasing the glucose uptake rate is the most beneficial strategy to improve ethanol production. This result was obtained by optimizing a model for ethanol production, while each of the reaction rates were separately allowed to deviate within fivefold around the nominal steady-state values. While no experimental validation was provided others have showed that overexpression of the hexose transporter increases ethanol production (Gutiérrez-Lomelí et al. 2008).

Evaluating engineering strategies

While MCA allows one to calculate the control each enzyme has on the flux, and as such suggest potential metabolic engineering targets, these control coefficients are based on small perturbations of the parameter values. The control coefficients do not generally hold for larger changes in enzymes levels, as obtained through metabolic engineering strategies such as genetic manipulation. The control tends to be redistributed if the parameter values change, especially when the expression of an enzyme with high flux control is increased. This is at least partly due to changes in metabolite concentrations around an enzyme upon changing its expression level, these changes does not only affect the reaction rates around these metabolites, but can also induce other homeostatic responses via for instance gene expression.

Nonetheless, kinetic models can readily be used to study larger perturbations in enzyme kinetics: in principle any perturbation to the system can be tested *in silico* by changing parameter values and comparison of the original with the moderated model. Parachin *et al.* (2011) made a kinetic model to evaluate different strategies to allow *S. cerevisiae* to use xylose as a carbon source. Parameterization was performed as a combination of literature data and estimation with fermentation data. The *in silico* model was used to simulate different enzyme activities, showing that a further overexpression of xylulokinase is required to increase the xylose intake. Overexpression of an *E. coli* xylulokinase confirmed this experimentally.

Chen *et al.* (2012) also generated a kinetic model of glycolysis that was parameterized using literature and estimation with fermentation data. Production of the fine chemical building block DHAP was increased *in silico* by removing the triose-phosphate isomerase (TPI) reaction from the model, simulating a TPI knockout. No experimental validation was provided with this work.

As mentioned above, genetic manipulation of metabolic pathways often leads to changes in metabolite concentrations, which can subsequently have other, undesired, effects. In a proof-of-principle, Adamczyk and Westerhoff (2012) demonstrated *in silico* that the phosphotransferase system from *Lactococcus lactis* can be replaced with a glucose transporter and hexokinase from *S. cerevisiae*, while keeping the concentrations of glycolytic intermediates unchanged. They showed with time-course simulations that fine-tuning of the pyruvate kinase activity is required to keep the concentrations stable. This was, however, not validated experimentally.

Why dynamic models are not used extensively in metabolic engineering

A number of examples are given above, where kinetic models of yeast metabolism were used to address metabolic engineering questions. Theoretically there are significant advantages to use kinetic models to investigate and evaluate strategies for genetic manipulation, such as the easy tinkering with the system by effortlessly changing parameter values *in silico*. Nonetheless, the number of successful uses of kinetic models of yeast metabolism all the way to actual increased production are slim, most of the approaches strand with the *in silico* suggestion of potential targets. This is potentially due to some disadvantages of kinetic modeling: it relies on correct description of the system, in the form of realistic rate expressions, parameter values, but also inclusion of relevant regulatory processes. The determination of parameter values can be very laborious and time-consuming, values can vary significantly depending on condition, but also expression levels, phosphorylation states and allosteric regulation can cause major changes to the enzyme activity. Additionally, transcriptomic changes are not necessarily reflected in the proteome and in enzyme activities, such that these methods cannot be used to easily estimate v_{max} values in different experiments. Also other parameters such as affinity constants can change depending on the environments' pH, ionic strength or levels of other regulators. This variety can partly be mitigated by taking uncertainty into account during the simulations, however, larger parameter distributions also result in less precise predictions, reducing the power of modeling. Taking uncertainty into account, however, does not address the adaptation and regulation that can occur upon genetic manipulation, which is central in metabolic engineering. Moisset *et al.* (2012) generated a promising model of glucose repression in yeast, where the expression of genes was dependent on the level of glucose, and this model was able to reproduce metabolic fluxes during exponential growth on glucose or ethanol. Such an approach could be used to model other types of regulation.

Even when *in silico* parameter values correspond to real *in vivo* parameter values, dynamic models typically cover only or a few reaction steps or pathways. This means that large parts of the cell are not taken into account, which can for instance react to changes in co-factor ratios that are observed in the pathway under study (Chen et al. 2014). Although efforts are made to generate genome-scale kinetic models of metabolism, these models heavily rely on parameter estimation methods, typically

in combination with the use of abstract approximate rate expressions. As the use of models is to gain understanding of the system, it is better not to make the model more complicated than necessary.

Perspectives

Computational models of metabolism play an important role in metabolic engineering, as metabolism is a complex network that requires a systemic view. Application of systems biology becomes more and more important in metabolic engineering as "lower hanging fruits have already been collected" (Almquist et al. 2014) and better understanding of metabolism and its regulation is required to further advance metabolic engineering and synthetic biology. GEMs and kinetic models both have their advantages and drawbacks. GEMs are comprehensive, but are dependent on a pseudo-steady state and in their current form do not take regulation, such as gene-protein and protein-protein level interactions, allosteric regulation or regulation at post-translational level, into account. At the same time kinetic models require parameter values that are difficult to estimate at the global scale. As neither of them can fully replace the other, there has been a considerable effort on combining these two approaches in a singular model or applying them in succession. GEMs in combination with global datasets are invaluable for detecting the metabolic bottlenecks, but due to lack of kinetic data about enzyme activities or regulation mechanisms they are limited in their predictive power. Applying kinetic models on metabolic bottlenecks previously detected with GEMs can help to understand the regulation or kinetics of these specific enzymatic steps, as one would not need model parameters for the whole system (Figure 3).

Metabolic models that describe a sub system of a network and take basic metabolic regulation into account have been developed for bacterial systems. Regulatory and integrated FBA models have been developed for *E. coli* (Covert et al. 2001, 2004, 2008) and it has found that those models that take into account transcriptional regulation can well predict the diauxic growth of *E. coli* which is not possible with conventional constraint-based models. Regulation in terms of four well described global regulators of *E. coli* was introduced into a medium size central carbon metabolism model and, similarly, diauxic shifts were characterized. In another study thermodynamics and enzyme kinetics were taken into account for reconstructing a genome-scale model for *E. coli* and it represents a large scale nonlinear metabolic model that takes into account changes in enzyme activities while being used for designing metabolic engineering strategies (Chakrabarti et al. 2013). Larger scale whole cell modeling approach where genome-scale metabolic networks were combined with cell cycle and protein-DNA association have been developed for small bacteria *Mycoplasma genitalium* (Karr et al. 2012). Another similar model where the bacterial cell cycle is combined with DNA replication and cell volume using Cooper-Helmstetter theory have been developed (Abner et al. 2013) and should bring us closer to the understanding how different cellular sub-systems are working together. The current challenges in yeast metabolic modeling include catching up with bacterial models in terms of

combining ODEs with constraint-based models as eukaryotic cells contain additional challenges related with compartmentalization. Therefore, inclusion of a detailed description of protein synthesis and translation machinery and their connection to metabolism, instead of simple inclusion of their cost in the biomass equation, would represent the next steps in advancing yeast GEMs and would have similarities to the model developed for *Mycoplasma genitalium* (Karr et al. 2012). This allows investigation of these processes, how they affect metabolism and how changes in their characteristics can affect the whole system. Furthermore, by considering the cell cycle it would be possible to shift from analysis at the population level to analysis of single cells and hereby take us closer to understand cellular mechanisms and therefore improve strategies for synthetic biology applications.

Although larger scale and more complex integrated models represent a way to understand the global regulation of the cell and hereby drive synthetic biology towards creation of synthetic cells, the adage 'more-is-better' does not necessarily hold for the size of models in all cases. The purpose of a model is to aid in the understanding of a system, and larger and more complex model do not necessarily obtain this but might rather make it much harder to identify the core principles of a system. Very large and complex models might eventually be able to simulate the cell perfectly, proposed as a SiliconCell (Snoep et al. 2006; Westerhoff 2001), and constructed for *Mycoplasma genitalium* (Karr et al. 2012), but this does not necessarily mean that we *understand* the cell any better. For metabolic engineering, however, a near-perfect SiliconCell would be extremely valuable for testing different strategies *in silico* before initiating *in vivo* engineering.

As models can have different levels of complexities and scopes, all with their own purpose, it might be more important that resources and knowledge in these models is shared efficiently. Model building should be a community effort, where knowledge from all directions is incorporated. The yeast consensus networks are examples of this (Herrgård et al. 2008). A connection between these models and kinetic models, stored on databases such as BioModels database, would allow one to truly build a model in a modular fashion, gather information about all important processes and particularly focus on (model kinetically) those related to the area of interest.

Finally, while models are valuable tools by themselves in metabolic engineering and synthetic biology, they also represent a mindset of doing research. Kinetic modeling can provide very exact numbers for the control coefficient and final concentrations, but because of the dynamics and changing behavior of the cell, these values are not that interesting as the system can adapt anyways, and it may sometimes not be worth the effort to determine a large set of kinetic parameters. On the other hand large scale whole cell models will demonstrate how different subsystems of metabolism (e.g. DNA replication, ribosome synthesis, protein turnover, etc.) function together in time and space and would enable us to better design cell factories in the future.

Acknowledgements

We acknowledge financial support from European Research Council, the Novo Nordisk Foundation and the US Department of Energy.

References

Abner K, Aaviksaar T, Adamberg K & Vilu R (2013) Single-cell model of prokaryotic cell cycle. *J. Theor. Biol.* 341: 78–87.

Achcar F, Kerkhoven EJ, Bakker BM, Barrett MP & Breitling R (2012) Dynamic modelling under uncertainty: the case of Trypanosoma brucei energy metabolism. *PLoS Comput. Biol.* 8: e1002352.

Adamczyk M & Westerhoff H V (2012) Engineering of self-sustaining systems: substituting the yeast glucose transporter plus hexokinase for the Lactococcus lactis phosphotransferase system in a Lactococcus lactis network in silico. *Biotechnol. J.* 7: 877–883.

Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I & Nielsen J (2013) The RAVEN toolbox and its use for generating a genome-scale metabolic model for Penicillium chrysogenum. *PLoS Comput. Biol.* 9: e1002980.

Agren R, Otero JM & Nielsen J (2013) Genome-scale modeling enables metabolic engineering of Saccharomyces cerevisiae for succinic acid production. *J. Ind. Microbiol. Biotechnol.* 40: 735–747.

Almario MP, Reyes LH & Kao KC (2013) Evolutionary engineering of Saccharomyces cerevisiae for enhanced tolerance to hydrolysates of lignocellulosic biomass. *Biotechnol. Bioeng.* 110: 2616–2623.

Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J & Jirstrand M (2014) Kinetic models in industrial biotechnology - Improving cell factory performance. *Metab. Eng.* 24C: 38–60.

Alvarez-Vasquez F, Riezman H, Hannun YA & Voit EO (2011) Mathematical modeling and validation of the ergosterol pathway in Saccharomyces cerevisiae. *PLoS One* 6: e28344.

Alvarez-Vasquez F, Sims KJ, Hannun YA & Voit EO (2004) Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models. *J. Theor. Biol.* 226: 265–291.

Annaluru N et al. (2014) Total synthesis of a functional designer eukaryotic chromosome. *Science* 344: 55–58.

Asadollahi MA, Maury J, Patil KR, Schalk M, Clark A & Nielsen J (2009) Enhancing sesquiterpene production in Saccharomyces cerevisiae through in silico driven metabolic engineering. *Metab. Eng.* 11: 328–334.

Avrahami-Moyal L, Engelberg D, Wenger JW, Sherlock G & Braun S (2012) Turbidostat culture of Saccharomyces cerevisiae W303-1A under selective pressure elicited by ethanol selects for mutations in SSD1 and UTH1. *FEMS Yeast Res.* 12: 521–533.

Bailey J (1991) Toward a science of metabolic engineering. *Science (80-.).* 252: 1668–1675.

Blank LM, Kuepfer L & Sauer U (2005) Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* 6: R49.

Bordel S, Agren R & Nielsen J (2010) Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. *PLoS Comput. Biol.* 6: e1000859.

Bro C, Regenberg B, Förster J & Nielsen J (2006) In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production. *Metab. Eng.* 8: 102–111.

Brochado AR, Matos C, Møller BL, Hansen J, Mortensen UH & Patil KR (2010) Improved vanillin production in baker's yeast through in silico design. *Microb. Cell Fact.* 9: 84.

Bruggeman FJ & Westerhoff H V (2007) The nature of systems biology. *Trends Microbiol.* 15: 45–50.

Buescher JM et al. (2012) Global network reorganization during dynamic adaptations of Bacillus subtilis metabolism. *Science* 335: 1099–1103.

Burgard AP & Maranas CD (2003) Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol. Bioeng.* 82: 670–677.

Cakar ZP, Seker UOS, Tamerler C, Sonderegger M & Sauer U (2005) Evolutionary engineering of multiple-stress resistant Saccharomyces cerevisiae. *FEMS Yeast Res.* 5: 569–578.

Chakrabarti A, Miskovic L, Soh KC & Hatzimanikatis V (2013) Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiology constraints. *Biotechnol. J.* 8: 1043–1057.

Chen N, Koumpouras GC, Polizzi KM & Kontoravdi C (2012) Genome-based kinetic modeling of cytosolic glucose metabolism in industrially relevant cell lines: Saccharomyces cerevisiae and Chinese hamster ovary cells. *Bioprocess Biosyst. Eng.* 35: 1023–1033.

Chen WW, Niepel M & Sorger PK (2010) Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev.* 24: 1861–1875.

Chen Y, Bao J, Kim I-K, Siewers V & Nielsen J (2014) Coupled incremental precursor and co-factor supply improves 3-hydroxypropionic acid production in Saccharomyces cerevisiae. *Metab. Eng.* 22: 104–109.

Choi J & Antoniewicz MR (2011) Tandem mass spectrometry: a novel approach for metabolic flux analysis. *Metab. Eng.* 13: 225–233.

Christensen TS, Oliveira AP & Nielsen J (2009) Reconstruction and logical modeling of glucose repression signaling pathways in Saccharomyces cerevisiae. *BMC Syst. Biol.* 3: 7.

Chubukov V, Uhr M, Le Chat L, Kleijn RJ, Jules M, Link H, Aymerich S, Stelling J & Sauer U (2013) Transcriptional regulation is insufficient to explain substrate-induced flux changes in Bacillus subtilis. *Mol. Syst. Biol.* 9: 709.

Covert MW, Schilling CH & Palsson B (2001) Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* 213: 73–88.

Covert MW, Xiao N, Chen TJ & Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. *Bioinformatics* 24: 2044–2050.

Covert, Vasser M, Polakis P & Palsson BØ (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429: 92–96.

Cronwright GR, Rohwer JM & Prior BA (2002) Metabolic Control Analysis of Glycerol Synthesis in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* 68: 4448–4456.

Dantas G et al. (2006) Mis-translation of a computationally designed protein yields an exceptionally stable homodimer: implications for protein engineering and evolution. *J. Mol. Biol.* 362: 1004–1024.

Daran-Lapujade P et al. (2007) The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels. *Proc. Natl. Acad. Sci. U. S. A.* 104: 15753–15758.

Daran-Lapujade P, Jansen ML a, Daran J-M, van Gulik W, de Winde JH & Pronk JT (2004) Role of transcriptional regulation in controlling fluxes in central carbon metabolism of Saccharomyces cerevisiae. A chemostat culture study. *J. Biol. Chem.* 279: 9125–9138.

Demeke MM et al. (2013) Development of a D-xylose fermenting and inhibitor tolerant industrial Saccharomyces cerevisiae strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol. Biofuels* 6: 89.

Dhar R, Sägesser R, Weikert C, Yuan J & Wagner a (2011) Adaptation of Saccharomyces cerevisiae to saline stress through laboratory evolution. *J. Evol. Biol.* 24: 1135–1153.

Dikicioglu D, Pir P, Onsan ZI, Ulgen KO, Kirdar B & Oliver SG (2008) Integration of metabolic modeling and phenotypic data in evaluation and improvement of ethanol production using respiration-deficient mutants of Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* 74: 5809–5816.

Duarte NC, Herrgård MJ & Palsson BØ (2004) Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* 14: 1298–1309.

Dymond JS et al. (2011) Synthetic chromosome arms function in yeast and generate phenotypic diversity by design. *Nature* 477: 471–476.

Elowitz MB & Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403: 335–338.

Van Eunen K, Kiewiet JAL, Westerhoff H V & Bakker BM (2012) Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics. *PLoS Comput. Biol.* 8: e1002483.

Feder ME & Walser J-C (2005) The biological limitations of transcriptomics in elucidating stress and stress responses. *J. Evol. Biol.* 18: 901–910.

Feizi A, Österlund T, Petranovic D, Bordel S & Nielsen J (2013) Genome-scale modeling of the protein secretory machinery in yeast. *PLoS One* 8: e63284.

Fendt S-M, Oliveira AP, Christen S, Picotti P, Dechant RC & Sauer U (2010) Unraveling condition-dependent networks of transcription factors that control metabolic pathway activity in yeast. *Mol. Syst. Biol.* 6: 432.

Frick O & Wittmann C (2005) Characterization of the metabolic shift between oxidative and fermentative growth in Saccharomyces cerevisiae by comparative 13C flux analysis. *Microb. Cell Fact.* 4: 30.

Förster J, Famili I, Fu P, Palsson BØ & Nielsen J (2003) Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res.* 13: 244–253.

Galazzo J & Bailey J (1990) Fermentation pathway kinetics and metabolic flux control in suspended and immobilized Saccharomyces cerevisiae. *Enzyme Microb. Technol.* 12.

Gardner TS, Cantor CR & Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature* 403: 339–342.

Gerosa L & Sauer U (2011) Regulation and control of metabolic fluxes in microbes. *Curr. Opin. Biotechnol.* 22: 566–575.

Gianchandani EP, Oberhardt MA, Burgard AP, Maranas CD & Papin JA (2008) Predicting biological system objectives de novo from internal state measurements. *BMC Bioinformatics* 9: 43.

Gombert AK, Moreira M, Christensen B, Nielsen J & Acteriol JB (2001) Network Identification and Flux Quantification in the Central Metabolism of Saccharomyces cerevisiae under Different Conditions of Glucose Repression. 183: 1441–1451.

Van Gulik WM & Heijnen JJ (1995) A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol. Bioeng.* 48: 681–698.

Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR & Sethna JP (2007) Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3: 1871–1878.

Gutiérrez-Lomelí M, Torres-Guzmán JC, González-Hernández GA, Cira-Chávez LA, Pelayo-Ortiz C & Ramírez-Córdova JDJ (2008) Overexpression of ADH1 and HXT1 genes in the yeast Saccharomyces cerevisiae improves the fermentative efficiency during tequila elaboration. *Antonie Van Leeuwenhoek* 93: 363–371.

Hanly TJ & Henson MA (2014) Dynamic model-based analysis of furfural and HMF detoxification by pure and mixed batch cultures of S. cerevisiae and S. stipitis. *Biotechnol. Bioeng.* 111: 272–284.

Hatzimanikatis V & Bailey JE (1996) MCA has more to say. *J. Theor. Biol.* 182: 233–242.

Hauf J, Zimmermann FK & Müller S (2000) Simultaneous genomic overexpression of seven glycolytic enzymes in the yeast Saccharomyces cerevisiae . *Enzyme Microb. Technol.* 26: 688–698.

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

Henriksen P et al. (2012) Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. *Mol. Cell. Proteomics* 11: 1510–1522.

Henry CS, DeJongh M, Best A a, Frybarger PM, Linsay B & Stevens RL (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28: 977–982.

Herrgård MJ et al. (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* 26: 1155–1160.

Hjersted JLL & Henson MA a (2009) Steady-state and dynamic flux balance analysis of ethanol production by Saccharomyces cerevisiae. *IET Syst. Biol.* 3: 167–179.

Hong K-K, Vongsangnak W, Vemuri GN & Nielsen J (2011) Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc. Natl. Acad. Sci. U. S. A.* 108: 12179–12184.

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

Ishii N et al. (2007) Multiple high-throughput analyses monitor the response of E. coli to perturbations. *Science (80-.).* 316: 593–597.

Jol SJ, Kümmel A, Terzer M, Stelling J & Heinemann M (2012) System-level insights into yeast metabolism by thermodynamic analysis of elementary flux modes. *PLoS Comput. Biol.* 8: e1002415.

Kacser H & Burns J (1979) Molecular democracy: who shares the controls? *Biochem. Soc. Trans.* 7: 1149–1160.

Kacser H & Burns J (1973) The control of flux. *Symp. Soc. Exp. Biol.* 27: 65–104.

Kacser H & Burns J (1981) The molecular basis of dominance. *Genetics* 97: 639–666.

Karr JR, Sanghvi JC, Macklin DN, Gutschow M V., Jacobs JM, Bolival B, Assad-Garcia N, Glass JI & Covert MW (2012) A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell* 150: 389–401.

Klipp E, Nordlander B, Krüger R, Gennemark P & Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nat. Biotechnol.* 23: 975–982.

Kohlstedt M et al. (2014) Adaptation of Bacillus subtilis carbon core metabolism to simultaneous nutrient limitation and osmotic challenge: a multi-omics perspective. *Environ. Microbiol.* 16: 1898–1917.

Kotte O, Zaugg JB & Heinemann M (2010) Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol. Syst. Biol.* 6: 1–9.

Kuepfer L, Sauer U & Blank LM (2005) Metabolic functions of duplicate genes in Saccharomyces cerevisiae. *Genome Res.* 15: 1421–1430.

Lee KH, Park JH, Kim TY, Kim HU & Lee SY (2007) Systems metabolic engineering of Escherichia coli for L-threonine production. *Mol. Syst. Biol.* 3: 149.

Lewis NE, Nagarajan H & Palsson BO (2012) Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nat. Rev. Microbiol.* 10: 291–305.

Liebermeister W & Klipp E (2006) Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Model.* 3: 41.

Liebermeister W, Uhlendorf J & Klipp E (2010) Modular rate laws for enzymatic reactions: thermodynamics, elasticities and implementation. *Bioinformatics* 26: 1528–1534.

Link H, Kochanowski K & Sauer U (2013) Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nat. Biotechnol.* 31: 357–361.

Machado D & Herrgård M (2014) Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* 10: e1003580.

Mendes P & Kell D (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* 14: 869–883.

Messiha H, Kent E, Malys N & Carroll K (2014) Enzyme characterisation and kinetic modelling of the pentose phosphate pathway in yeast. *PeerJ Prepr.* 2.

Moisset P, Vaisman D, Cintolesi A, Urrutia J, Rapaport I, Andrews BA & Asenjo JA (2012) Continuous modeling of metabolic networks with gene regulation in yeast and in vivo determination of rate parameters. *Biotechnol. Bioeng.* 109: 2325–2339.

Moreno-Sánchez R, Saavedra E, Rodríguez-Enríquez S & Olín-Sandoval V (2008) Metabolic control analysis: a tool for designing strategies to manipulate metabolic pathways. *J. Biomed. Biotechnol.* 2008: 597913.

Nissen TL, Schulze U, Nielsen J & Villadsen J (1997) Flux Distributions in Anaerobic, Glucose-Limited Continuous Cultures of Saccharomyces Cerevisiae. *Microbiology* 143: 203–218.

Nookaew I, Jewett MC, Meechai A, Thammarongtham C, Laoteng K, Cheevadhanarak S, Nielsen J & Bhumiratana S (2008) The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism. *BMC Syst. Biol.* 2: 71.

Oliveira AP, Ludwig C, Picotti P, Kogadeeva M, Aebersold R & Sauer U (2012) Regulation of yeast central metabolism by enzyme phosphorylation. *Mol. Syst. Biol.* 8: 623.

Oliveira AP & Sauer U (2012) The importance of post-translational modifications in regulating Saccharomyces cerevisiae metabolism. *FEMS Yeast Res.* 12: 104–117.

Osterlund T, Nookaew I & Nielsen J (2012) Fifteen years of large scale metabolic modeling of yeast: developments and impacts. *Biotechnol. Adv.* 30: 979–988.

Otero JM, Cimini D, Patil KR, Poulsen SG, Olsson L & Nielsen J (2013) Industrial systems biology of Saccharomyces cerevisiae enables novel succinic acid cell factory. *PLoS One* 8: e54144.

Paddon CJ et al. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496: 528–532.

Parachin NS, Bergdahl B, van Niel EWJ & Gorwa-Grauslund MF (2011) Kinetic modelling reveals current limitations in the production of ethanol from xylose by recombinant Saccharomyces cerevisiae. *Metab. Eng.* 13: 508–517.

Patil KR, Rocha I, Förster J & Nielsen J (2005) Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics* 6: 308.

Peebo K, Valgepea K, Nahku R, Riis G, Oun M, Adamberg K & Vilu R (2014) Coordinated activation of PTA-ACS and TCA cycles strongly reduces overflow metabolism of acetate in Escherichia coli. *Appl. Microbiol. Biotechnol.* 98: 5131–5143.

Pfleger BF, Pitera DJ, Newman JD, Martin VJJ & Keasling JD (2007) Microbial sensors for small molecules: development of a mevalonate biosensor. *Metab. Eng.* 9: 30–38.

Pitkänen E et al. (2014) Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species. *PLoS Comput. Biol.* 10: e1003465.

Polisetty PK, Gatzke EP & Voit EO (2008) Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods. *Biotechnol. Bioeng.* 99: 1154–1169.

Savageau MA (1969) Biochemical systems analysis: I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.* 25: 365–369.

Schellenberger J et al. (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* 6: 1290–1307.

Schomburg I et al. (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res.* 41: D764–72.

Schuetz R, Kuepfer L & Sauer U (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol. Syst. Biol.* 3: 119.

Schuetz R, Zamboni N, Zampieri M, Heinemann M & Sauer U (2012) Multidimensional optimality of microbial metabolism. *Science* 336: 601–604.

Segre D, Vitkup D & Church GM (2002) Analysis of optimality in natural and perturbed. *Proc. Natl. Acad. Sci.* 99: 15112–15117.

Smallbone K et al. (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: 6.

Snoep JL, Bruggeman F, Olivier BG & Westerhoff H V (2006) Towards building the silicon cell: a modular approach. *Biosystems.* 83: 207–216.

Stanford NJ, Lubitz T, Smallbone K, Klipp E, Mendes P & Liebermeister W (2013) Systematic construction of kinetic models from genome-scale metabolic networks. *PLoS One* 8: e79195.

Teusink B et al. (2000) Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* 267: 5313–5329.

Tran LM, Rizk ML & Liao JC (2008) Ensemble modeling of metabolic networks. *Biophys. J.* 95: 5606–5617.

Valgepea K, Adamberg K, Seiman A & Vilu R (2013) Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. Biosyst.* 9: 2344–2358.

Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, Dijken JP van & Heijnen S (1996) Validation of a Metabolic Network for Saccharomyces cerevisiae Using Mixed Substrate Studies. *Biotechnol. Prog.* 12: 434–448.

Vaseghi S, Baumeister A, Rizzi M & Reuss M (1999) In Vivo Dynamics of the Pentose Phosphate Pathway in Saccharomyces cerevisiae. *Metab. Eng.* 140.

Westerhoff H V (2001) The silicon cell, not dead but live! *Metab. Eng.* 3: 207–210.

Visser D & Heijnen JJ (2003) Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab. Eng.* 5: 164–176.

Wittig U et al. (2012) SABIO-RK--database for biochemical reaction kinetics. *Nucleic Acids Res.* 40: D790–6.

Voit EO (2013) Biochemical Systems Theory: A Review. *ISRN Biomath.* 2013: 1–53.

Yeh BJ & Lim W a (2007) Synthetic biology: lessons from the history of synthetic organic chemistry. *Nat. Chem. Biol.* 3: 521–525.

Yuan Y & Zhao H (2013) Directed evolution of a highly efficient cellobiose utilizing pathway in an industrial Saccharomyces cerevisiae strain. *Biotechnol. Bioeng.* 110: 2874–2881.

Zelezniak A, Sheridan S & Patil KR (2014) Contribution of network connectivity in determining the relationship between gene expression and metabolite concentration changes. *PLoS Comput. Biol.* 10: e1003572.

Zhang F, Carothers JM & Keasling JD (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30: 354–359.

Zhang M, Galdieri L & Vancura A (2013) The yeast AMPK homolog SNF1 regulates acetyl coenzyme A homeostasis and histone acetylation. *Mol. Cell. Biol.* 33: 4701–4717.

Legends

Figure 1: Reconstruction and uses of genome-scale models. (A) The genome annotation, but also other omics data can be used to reconstruct the metabolic network. (B) Gene-protein-reaction relationships are defined for the metabolic model. While this whole map is used for contextualizing high-throughput data, only the reaction network is used to construct the stoichiometry matrix (C). A solution space is defined from the constraints applied to the model and subsequent analysis can, among other things, provide instructions for metabolic engineering strategies, or contextualizing high-throughput data.

Figure 2: Overview of kinetic modeling of metabolism. Kinetic parameters need to be defined through experimentation, estimation or literature data (A), to populate the rate expressions (B) that describe the reactions of the metabolic network that is being modeled (C). The resulting system of ordinary differential equations can be used to evaluate engineering strategies (here the effect of doubling the enzyme activity on the metabolic flux is shown) or identify flux controlling reactions through MCA (the control coefficient is related to the slope of the flux over enzyme activity).

Figure 3: Perspective of the combined use of genome-scale models and kinetic models: GEMs provide the large context of the whole metabolic network and analysis of this system can indicate where more detailed description of the kinetics can be of interest. A number of examples of investigations that are possible with kinetic modeling are indicated with the colored boxes.

Table 1: A selection of applications of mathematical models of metabolism in metabolic engineering of *S. cerevisiae*.

Table 1

Reference	**Aim**	**Model**	**Outcome**	**Experimental confirmation**
Bro *et al.*, 2006	Production of ethanol	GEM – iFF708 (Förster et al. 2003)	The resulting strain had a 40% lower glycerol yield on glucose while the ethanol yield increased with 3% without affecting the maximum specific growth rate	Yes
Dikicioglu *et al.*, 2008	Production of ethanol	GEM – iFF708	Several single mutants that effected in partial or complete deficiency of respiration were cultivated in chemostat experiments and the outcome was compared with the FBA results.	Yes
Hjersted & Henson, 2009	Production of ethanol	GEM – iJH732 (same paper)	A combined deletion/overexpression/insertion mutant with improved ethanol productivity capabilities was computationally identified by dynamically screening multiple combinations of the ten metabolic engineering strategies.	No
Asadollahi *et al.*, 2009	Production of sesquiterpenes	GEM – iFF708; OptGene (Patil et al. 2005) and MOMA (Segre et al. 2002)	Deletion of GDH1 led to an approximately 85% increase in the final cubebol titer.	Yes
Brochado *et al.*, 2010	Production of vanillin	GEM – iFF708	Three of the mutants showed up to 1.5 fold higher vanillin β-D-glucoside yield in batch mode, while continuous culture of the Δ*pdc1* mutant showed a 2-fold productivity improvement. This	Yes
Otero *et*	Production	GEM – OptGene	The resulting strain represents a 30-fold	Yes

al., 2013	of succinic acid		improvement in succinate titer, and a 43-fold improvement in succinate yield on biomass	
Agren *et al.*, 2013b	Production of succinic acid	GEM – iFF708	Metabolic modeling suggested several potential mutations for improved succinic acid productions, where Δ*dic1* produced the product with a yield of 0.02 C-mol/C-mol glucose	Yes
Hanly & Henson, 2014	Production of ethanol in co-culture with furfural present	GEM – iMM904 and iBB814	A co-culture of *S. cerevisiae* and *S. stipitis* was simulated by combining two GEMS and including extracellular mass-balances. Reduced synthesis of acetate by *S. cerevisiae* yielded the highest ethanol production.	No
Cronwright *et al.*, 2002	Decrease glycerol production	Kinetic model – MCA	MCA of a small model of glycerol synthesis indicated that G3PDH has a control of 0.85 over the glycerol flux, but that also DHAP and ATP levels are controlling the flux.	No
Polisetty *et al.*, 2008	Production of ethanol	Kinetic model – BST	Overexpression of glucose uptake would lead to the highest increase in ethanol yield. This was not experimentally validated by the authors.	Not by same authors (Gutierriez-Lomeli *et al.*, 2008)
Parachin *et al.*, 2011	Xylose as carbon source	Kinetic model – simulate different enzyme activities	Increasing xylulokinase activity improved the xylose consumption *in silico*, while experimental validation showed a 27-37% increase in xylose consumption.	Yes
Chen *et al.*, 2012	Production of DHAP	Kinetic model – simulate knockout	*In silico* knockout of the triose phosphate isomerase (TPI) resulted in higher DHAP accumulation (28% more than the measured experimental flux), depending on the glucose uptake rate.	No