Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Advancing metabolic models with kinetic information

Hannes Link[1], Dimitris Christodoulou[1,2] and Uwe Sauer[1]

Kinetic models are crucial to quantitatively understand and predict how functional behavior emerges from dynamic concentration changes of cellular components. The current challenge is on resolving uncertainties about parameter values of reaction kinetics. Additionally, there are also major structural uncertainties due to unknown molecular interactions and only putatively assigned regulatory functions. What if one or few key regulators of biochemical reactions are missing in a metabolic model? By reviewing current advances in building kinetic models of metabolism, we found that such models experience a paradigm shift away from fitting parameters towards identifying key regulatory interactions.

Addresses
[1] Institute of Molecular Systems Biology, ETH Zurich, Auguste-Piccard-Hof 1, 8093 Zurich, Switzerland
[2] Life Science Zurich PhD Program on Systems Biology, Zurich, Switzerland

Corresponding author: Sauer, Uwe (sauer@imsb.biol.ethz.ch)

Current Opinion in Biotechnology 2014, **29**:8–14

This review comes from a themed issue on **Cell and pathway engineering**

Edited by **Tina Lütke-Eversloh** and **Keith EJ Tyo**

0958-1669/$ – see front matter,

http://dx.doi.org/10.1016/j.copbio.2014.01.015

Introduction

Computational models provide a framework for integrating existing knowledge about metabolic networks with experimental data, thus enabling quantitative understanding of metabolism and its regulation. Their ultimate goal is to predict phenotypes that result from the reorganization of metabolism after genetic or environmental perturbations. Currently this is achieved by constraint-based approaches that rely exclusively on reaction stoichiometries. These models employ metabolic networks based on genome sequences, resulting in predictions that advanced many biotechnological applications [1] and brought fundamental biological insights [2,3]. Yet, these models are static by predicting what metabolic fluxes are but not how they are achieved. What is missing are reaction kinetics and their regulators, whose implementation through ordinary differential equation (ODE) models allows to simulate dynamics of metabolism.

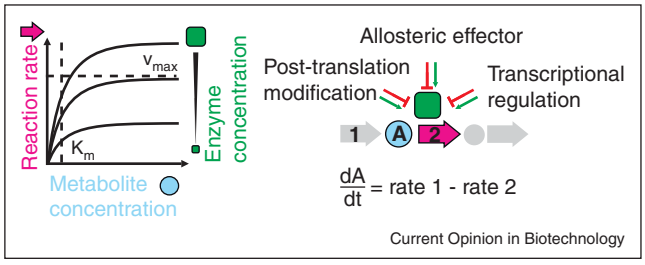
ODE models are based on mathematical equations describing reaction fluxes as a function of metabolite and enzyme concentration (Figure 1) that can describe metabolism mechanistically or coarse-grained. In coarse-grained models, metabolism is broken down to its essential features and such models successfully addressed key biological questions about cellular economics [4] or allocation of proteomic resources [5]. However, the high abstraction level precludes identification of molecular targets for applications like metabolic engineering or drug design. Conversely, mechanistic models contain many molecular details and aim to identify molecular targets that control metabolic fluxes. Although formal control theory approaches were developed for this purpose [6,7], mechanistic models are still an exception in biotechnological applications today.

The prevailing problem is our incomplete knowledge about kinetic parameters and regulatory interactions, complicating enlargement of coarse-grained models and limiting predictions of mechanistic models. The difficulty is to decide *a priori* which of the many known interactions to include in a model, and key regulators are often unknown altogether. Nevertheless, several successful metabolic models with kinetic information were published over the past three years. What has enabled their predictions? First, we reviewed approaches that resolve parameter uncertainties by integrating experimental data with kinetic models. Since many of the recent models couple regulatory mechanisms with metabolism we discuss advances based on including allosteric enzyme–metabolite interactions, transcriptional regulation and signaling pathways. Finally we introduce the nascent approach of ensemble modeling that systematically addresses structural uncertainties about regulatory mechanisms.

Parameter fitting

Increasingly comprehensive multi-omics data sets become available for stationary [8–10] and dynamic conditions [11]. The current challenge is to use such data to increase model information content and predictive power. Usually this is achieved through parameter identification in metabolic models using computational methods that drew much attention in the recent years [12,13]. Non-linear regression is a method to fit parameters to a series of steady state experiments (Figure 2a). For instance, Reed and colleagues revised a benchmark model of *Escherichia coli* central metabolism using non-linear regression to a multi-omics dataset [14•]. They exploited information about parameter confidence intervals to select targets for model simplification, enabling predictions about putative regulatory enzyme–metabolite interactions and

Figure 1



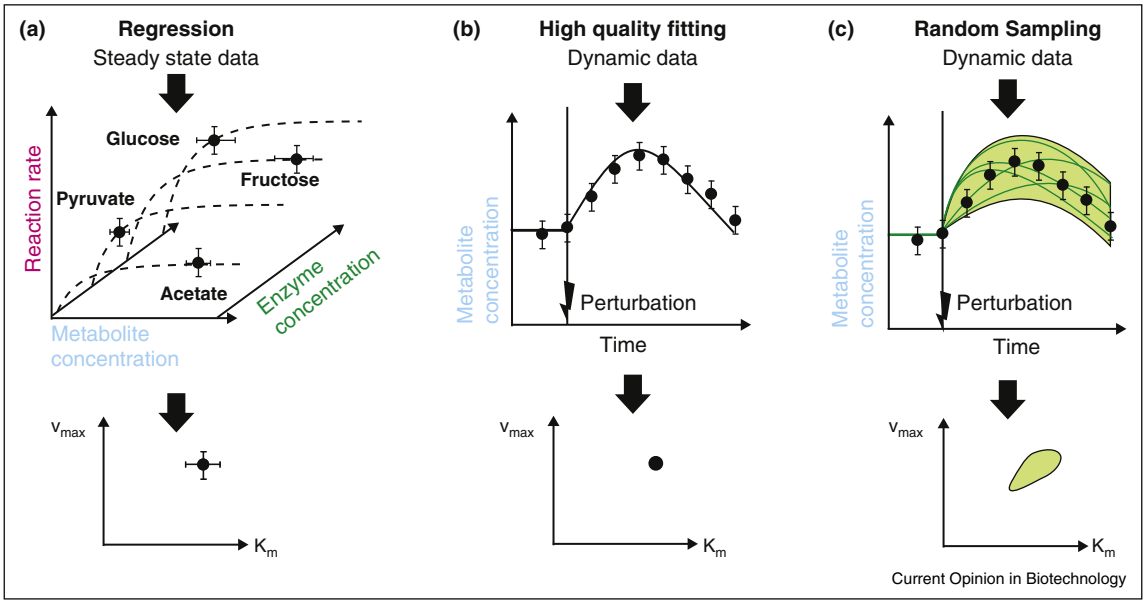
ODE-based kinetic models. Reaction kinetics and their parameters determine how concentrations of metabolites and enzymes influence the reaction rate. The reaction rate depends also on allosteric effectors, transcriptional regulation and post-translational modifications, which change enzymatic activity. Reaction kinetics are summed up in ordinary differential equations that describe the mass balances around a metabolite.

flux ranges. The same data set was used to develop a parameter estimation method that compensates for missing data [15]. Both studies present systematic rigorous parameter fitting to steady state data. Integrating the more informative dynamic data with kinetic models requires global optimization methods that are computationally costly (Figure 2b). Consequently large efforts focused on developing efficient optimization algorithms, exceeding a hundred publications already some years ago [16]. New algorithms focus on the confidence of estimated parameters [17–19] and on reducing computational costs by slope-estimation methods [20,21].

The above methods aim at high quality fits to existing experimental data through identification of a single, optimal parameter set. A major problem is that this optimum is often local, meaning that other parameter sets can lead to a model that performs equally well or better. Multi-start optimization and random sampling methods provide a solution to this problem and in particular the latter method has become increasingly popular. Random sampling populates the model with random kinetic parameter sets and averages predictions from the best parameters sets (Figure 2c). To improve random exploration of a high-dimensional parameter space, Liao and colleagues constrained sampling to parameter sets that achieved the experimentally determined steady state [22•]. This approach was applied to obtain 24 out of 5000 parameter sets for a model of *E. coli* central metabolism that matched mutant physiology and enabled prediction of validated metabolic engineering targets. Related approaches integrate Monte Carlo sampling and Metabolic Control Analysis to search for regulatory sites [23] or to identify drug targets [24,25]. The success of these random sampling methods shows that good predictions do not necessarily require precise parameters, as already realized a couple of years ago [26].

How important are precise parameters for a model's ability to predict unobserved data? First, only few parameters might limit the parameter space and a divide-and-conquer approach was developed to infer this space from multi-omics data [27]. Second, Steuer et al.

Figure 2



Parameter fitting. **(a)** Regression analysis using multi-omics data from steady state experiments performed under different conditions results in parameter values and their confidence intervals. **(b)** Minimization of residuals between dynamic data and simulation results in optimized parameter values. Information about parameter confidence intervals is not straight-forward. **(c)** Random sampling identifies regions in the parameter space that match experimental data equally well.