Graphical Abstract

Several modeling frameworks for describing and redirecting cellular metabolism have been developed keeping pace with the rapid development in high-throughput data generation and advances in metabolic engineering techniques. The incorporation of kinetic information within stoichiometry-only modeling techniques offers potential advantages for improved phenotype prediction and consequently more precise computational strain design. In addition to substrate-level kinetic regulatory information, the integration of a number of additional layers of regulation at the transcription, translation, and post-translation levels is sought after by many research groups. However, the practical integration of these complex biological processes into a unified framework amenable to design remains an ongoing challenge.

Figure 1. Integration of different layers of regulation to improve metabolic phenotype prediction in kinetic models. (a) Signal transduction, (b) metabolic feedback inhibition, (c) metabolic feedback activation, (d) transcription factor regulation, (e) Riboswitch binding to small metabolite, (f) sRNA basepairing, (g) Phosphorylation modulation, (h) Allosteric modification, (i) Substrate-level feedforward activation, and (j) Substrate-level feedback inhibition.

Improving Prediction Fidelity of Cellular Metabolism with Kinetic Descriptions

Anupam Chowdhury¹, Ali Khodayari¹ and Costas D. Maranas

¹Joint first authors

Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802

*Corresponding author

E-mail: costas@psu.edu

Phone: 814-863-9958

Abstract

Several modeling frameworks for describing and redirecting cellular metabolism have been developed keeping pace with the rapid development in high-throughput data generation and advances in metabolic engineering techniques. The incorporation of kinetic information within stoichiometry-only modeling techniques offers potential advantages for improved phenotype prediction and consequently more precise computational strain design. In addition to substrate-level kinetic regulatory information, the integration of a number of additional layers of regulation at the transcription, translation, and post-translation levels is sought after by many research groups. However, the practical integration of these complex biological processes into a unified framework amenable to design remains an ongoing challenge.

Introduction

The central goal of metabolic engineering is to harness the cellular machinery of microorganisms and eukaryotic cell lines for converting low-value feedstock to economically viable biochemical products and biofuels. While in some cases productivity or yield limitations arise for reasons outside the scope of metabolic models [1] (e.g., toxicity, strain stability, etc.), in many cases the cause is within the range of metabolic modeling. For example, product yield may be limited due to enzyme promiscuity, product inhibition, enzyme de-activation, depleted precursor metabolite pools or insufficient enzyme levels. Metabolic models are increasingly becoming more complex expanding beyond simple stoichiometric descriptions of metabolic networks and gene-protein-reaction (GPR) relationships. Kinetic information is being used to postulate a mechanistic description of cellular function from gene transcription to translation and post-translational events [2]. In addition to providing a dynamic view of metabolism, kinetic descriptors offer the opportunity of linking together several layers of regulatory events into an integrated framework. Kinetic model development, however, is still plagued by a number of challenges chief among which are paucity of kinetic parameters, non-universality of rate laws, and implementation of regulatory events. These challenges become even more pressing when kinetic models are tasked with predictions at a genome-scale.

In this review, we concentrate on the recent advances in the construction of kinetic models and

kinetics-aided procedures for improved understanding and redirection of cellular metabolism. Incorporation of regulatory events at transcription, translation, post-translational, and substrate-levels with existing formalisms are described along with their associated challenges. Finally, we discuss state-of-the-art strain design protocols for metabolic engineering that incorporate available kinetic description for improving their prediction accuracy.

Kinetic model development

Kinetic model development requires knowledge of stoichiometry, kinetic rate laws, and fluxomic/metabolomics data to support parameter estimation. Earlier efforts relying on *in vitro* derived lumped kinetics led to inconsistencies of model predictions for *in vivo* phenotypes [3]. A number of linear approximate kinetic laws (e.g., log-lin [4], lin-log [5], Taylor series approximation [6], thermokinetic [7], and convenience rate laws [8]) were put forth instead to standardize the kinetic expressions and provide a tractable parameterization workflow. However, the prediction accuracy of these models was inherently limited to the proximity of the reference phenotype as parameterization did not, in general, include data for multiple phenotypes under genetic/environmental perturbations and only a limited number of substrate-level regulatory loops were included [9]. A promising way forward has been to deconstruct complex kinetic expressions into their elementary kinetic steps [10] while retaining all known regulation. First, metabolic reactions and their substrate-level regulatory interactions are decomposed into their elementary steps followed by the derivation of mass action kinetics for each step. This reduces the nonlinearity of the kinetic description to only bilinear terms without losing any mechanistic detail at the expense of a larger model size. Next, model variables (e.g., elementary kinetic parameters [10] or elasticity [11]) are expressed in terms of elementary reaction parameters (i.e., enzyme fractions and reactions reversibility), both of which are bounded between zero and one. This representation provides a natural scaling for parameter sampling from the entire kinetic space consistent with thermodynamic principles [10]. For example, the Ensemble Modeling (EM) paradigm was developed by uniformly sampling multi-dimensional space of elementary kinetics to construct an ensemble of kinetic models using a reference flux dataset [10]. In order to reduce the number of model variables during parameter sampling, metabolite concentrations are generally scaled by their reference strain values.

Given a kinetic model scaffold and sampling procedure, the remaining challenge is the availability of metabolomic and fluxomic data for a variety of phenotypes subject to environmental (e.g., aerobic vs. anaerobic), carbon substrate, and genetic perturbations. In an earlier effort, an optimization-based parameter estimation technique was proposed to parameterize EM formalism by integrating multiple mutant strain data under aerobic **conditions** in *E. coli* [12]. The developed model led to accurate predictions for aerobic succinate overproduction, but prediction quality was poor when the model was queried under anaerobic **conditions** [13]. This was because model parameterization was carried out in the absence of anaerobic flux data. This confirms that kinetic model prediction reliability can only be expected if the kinetic model **is** parameterized using information from perturbations (either genetic or environmental) that approximate the queried metabolic phenotype. At the same time, all relevant substrate level regulatory interactions must be *a priori* included in the model to allow for their parameterization.

Metabolic regulation

Microbial organisms leverage intertwined layers of regulation to make optimal use of metabolic resources in response to genetic/environmental stimuli. These regulatory layers of transcription, translation, post-translational modification, and substrate-level regulatory interaction act by altering the capacity of metabolic enzymes. Transcriptional and translational regulations mostly control allocation of enzyme resources, while post-translational regulation mainly modulates

enzyme kinetics. Substrate-level regulation also provides rapid responses through passive flux regulation by enzyme saturation and thermodynamics. Despite many advances, a gap still exists in effectively linking transcription and translation levels with cellular metabolism. The challenge is twofold: (a) identification of the structure of the underlying regulatory network [14] and (b) integration of the identified network with the rest of metabolic events through detailed mechanistic descriptions [15].

a. Regulatory network reconstruction

Environmental and/or metabolic changes modulate gene expression through signal transduction and/or feedback/feedforward regulation (e.g., flux sensing [16]) (Figure 1a-d). There have been several system biology efforts aimed at elucidating transcriptional and translational regulatory networks using forward-engineering [17] and reverse-engineering [14] approaches. These approaches, however, generally recover the totality of all potential interactions of which only a small subset becomes active under any given condition [18].

A comprehensive assessment of transcriptional regulation in yeast [19] and *E. coli* [20] was carried out by metabolic flux elucidation upon deletion of transcription factors (one at a time). The results revealed a sparse network of active transcriptional factors that control metabolic flux distribution for any given condition. In another effort, the Environmental and Gene Regulatory Influence Network (EGRIN2.0) software tool was developed to systematically identify condition-specific gene regulatory network by combining a biclustering technique (cMonkey) with a regression-based approach (Inferelator) [18]. Unlike transcriptional regulation, the global assessment of translational regulation remains largely uncharted. Riboswitches [21] (Figure 1e) and small noncoding RNAs (sRNAs) [22] (Figure 1f) are two major classes of translational regulators that act by affecting mRNAs translation through binding to small metabolites and basepairing to the target mRNA, respectively. While the events underpinning the control of translation initiation have been identified, many efforts are still underway to elucidate their interplay with the translation-initiation factors and ribosomal subunits [23]. Developing global techniques to systematically identify condition-specific translation regulatory networks at a genome-scale remains an open challenge.

Once the protein is translated, post-translational modification may alter protein catalytic capability through covalent (e.g., (de)phosphorylation) (Figure 1g) and noncovalent modifications (e.g., allostery) (Figure 1h) [24]. Advances in mass-spectrometry based proteomics have greatly facilitated a comprehensive discovery and quantification of covalent protein modifications [25]. Several efforts have also been made towards the systematic inference of allosteric interactions by combining dynamic metabolomics with ensemble modeling [26]. The substrate-level protein regulation (e.g., (un)competitive inhibition) (Figure 1i-j), however, is still limited to *in vitro* activity assays.

b. Integration of regulatory events with metabolism

Several efforts have been made towards integration of transcriptional regulation into stoichiometry-based models driven by advances in regulatory network reconstructions and availability of DNA microarray and RNAseq datasets. Early studies used Boolean operators by assuming a positive correlation between metabolic flux and gene expression level [27]. However, experimental studies revealed that negative correlations often arise [16]. Alternatively, while *ad hoc* constraints could be proposed to adjust metabolic flux ranges with transcriptomic [28] and/or proteomic [29] datasets under any given condition, however, the predictive accuracy of these approaches is condition dependent [30]. This is due to the fact that detailed regulatory events at the transcription, translation, and substrate-levels are not integrated into the models. Successful implementation of allosteric modification [31] and substrate-level regulation [12] with elementary

kinetic mechanism showed that kinetic approach can potentially address such limitations. In the same spirit, kinetic models can be utilized to incorporate regulatory events beyond those at substrate-level. This can be achieved by posing the rate of transcription and translation in terms of regulatory factor activity, mRNA level, and protein abundance using Hill-derived equations, mass action kinetics [32], or thermodynamic-based models [33].

Integration of regulatory and metabolic events under a unified cell-wide model is a long-term goal. Among the first attempts was the whole-cell modeling of *M. genitalium* that combined a stoichiometric representation of metabolism with an ODE representation of cytokinesis [34]. Alternatively, integration of Metabolism with macromolecules Expression (ME-Models) linked cellular physiology with gene expression machinery [35]. Transcription and translation rates as well as enzyme efficiency are expressed as a function of the organism's doubling time, as opposed to stoichiometry-only models that include them as demand constraints. In another effort, signal transduction was integrated with a stoichiometric representation of metabolism and gene expression machinery [36]. However, the systematic integration of all layers of regulation remains unfulfilled. Advances in kinetic modeling and comprehensive databases of omics information to support model parameterization provide hope for reaching this goal.

Improving strain design accuracy with kinetic information

Over the last decade, many algorithmic techniques for strain design relying on primarily genome-scale stoichiometry-based metabolic models were developed [37]. This was motivated by the relative tractability of Linear Programming models, availability of genome-scale reconstructions and their global coverage of metabolism. Similar developments for kinetic models were hampered due to the paucity of metabolomics/fluxomic datasets to support robust parameterization, relative narrow scope of earlier kinetic models, and associated computational difficulties. However, the increased availability of kinetic parameters [38] and workflows for kinetic model construction [39] has led to a renewed interest in deploying kinetic models for microbial strain design. While stoichiometry-only strain design methods can provide an initial roadmap of how to construct an improved strain, questions associated with enzyme level sufficiency, concentration bottlenecks and unwanted regulation remain unanswered [40]. Several recent efforts attempt to leverage both descriptions of metabolism as guidelines for improved strain design.

Principles derived from Metabolic Control Analysis (MCA) have been used in multiple studies to derive suggestions for strain construction. One at a time, enzymes in a kinetic model are perturbed by a small amount from their steady-state values to identify the reaction(s) that have the largest impact on the target flux. Variations of this method have been used for constructing strains for ethanol [41] and curcumin [42] production in *E. coli*, 2,3-butanediol in *L. lactis* [43] and for exploring metabolic bottlenecks in engineering the Entner Doudoroff pathway in *Z. mobilis* [44]. A related work explores the effect of periodic oscillations of select enzyme activities for improving the cumulative production of phosphoenolpyruvate in *E. coli* [45]. Kinetic models have also been used to optimize **enzyme expression** of the terminal (linear) pathways for the heterologous production of amorphadiene [46] and neurosporene [47] in *E. coli*. In particular, the latter study [47] developed a gene Sequence-Expression-Activity mapping (SEAMAP) procedure for optimizing the translation rates of the terminal pathway enzymes using the RBS Library Calculator [48] while the remaining fluxes were predicted with Flux Balance Analysis (FBA). Kinetic models are also well suited for metabolic engineering in cell-free systems due to the relatively small size of the reaction networks and higher reliability of the estimated kinetic parameters for *in vitro* systems. In a recent study, a kinetic model was used to optimize the **enzyme expression** of an *in vitro* reaction network for stable hydrogen production from co-

utilization of C₅ and C₆ carbon substrates [49]. Despite several success stories these approaches are generally restricted in their scope and wide applicability.

A number of computational strain design tools that allow for the incorporation of kinetic expressions rely on a hybrid description of metabolism, where genome-scale stoichiometry models are supplemented with kinetic expressions for a subset of reactions (Figure 2a). The **Dynamic Strain Scanning Optimization (DySScO)** procedure [50] builds upon **dynamic FBA algorithm** [51] whereby genome-scale models are augmented with uptake (export) kinetic expressions for the carbon substrate (and desired products). Intervention strategies are identified [50] that directly maximize not just product yield but also their titer and productivity, however, internal reactions are restricted to a stoichiometry-only description. Alternatively, a core kinetic model parameterized using the EM approach [52] was used to prioritize reaction deletion suggestions by stoichiometry-only OptKnock [53] and **Genetic Design through Local Search (GDLS)** [54] algorithms. Results showed that only a subset of stoichiometry-only constrained deletion strategies (three out of seven) provided high product yields when implemented on the kinetic model, while others did not meet the imposed cutoffs of product yield or biomass flux. The k-OptForce procedure [55] directly accounts for substrate-level regulatory bottlenecks and allowable ranges for intracellular metabolite concentration while suggesting flux-redirection strategies. In this bilevel Mixed Integer NonLinear Program (MINLP) optimization formulation, the outer problem optimizes the fluxes of the kinetic part of the model by suggesting enzymatic modifications (i.e., v_{max} changes) while the inner problem identifies flux redirections for the stoichiometry-only part of metabolism (Figure 2c) using a worst-case approach. The inclusion of kinetic restrictions was shown to prevent flux redirections that violate concentration bounds (e.g., serine overproduction in *E. coli* [55]) by often using additional but less drastic interventions. In other instances, kinetic expressions support desired flux distributions without the introduction of extra genetic interventions (e.g., triacetic acid lactone production in *S. cerevisiae* [55]). The nonlinearity of the kinetic expressions significantly increases the computational burden. A recent effort addresses this challenge by replacing kinetic descriptions with representative power-law approximations [56].

Generally, hybrid models provide tight flux ranges for reactions with kinetic information (Figure 2b) but are unable to do the same for most stoichiometry-only reactions decoupled from the ones with kinetic information. Substrate-level regulation is often unaccounted for, especially in secondary metabolism, thus affecting model predictions [13]. This requires re-parameterization of models upon introduction of missing regulatory loops revealed by discrepancies between model predictions and experimental data. Genome-scale models with a full kinetic description for all reactions are also beginning to emerge [57]. However, robust parameterization remains a challenge as experimental datasets are generally insufficient. It is important to note that kinetic model predictions are context specific and are only informative under the conditions for which the underlying kinetic model has been parameterized. Unlike FBA analysis where the maximization of the drain of all biomass components coupled through cofactor balancing propagates through the network often leading to non-intuitive insight, kinetic models generally require that each regulatory event, concentration bound and enzyme saturation level to be *a priori* provided. As kinetic models become more comprehensive in their description, it becomes much more difficult to analyze the underlying factors that lead to the obtained results. Generally, the sensitivity of the results to concentration bounds, allowable fold-change in enzyme levels, and strength of regulation needs to be carefully assessed. Deterministic optimization methods approach their limits of tractability for large kinetic models both at the parameterization step and for strain design. Careful recasting of the original optimization problem (e.g., decomposing the bilevel MINLP problem in k-OptForce into two steps [55]), simplification of the kinetic expressions (e.g., decomposing overall reaction mechanisms to their elementary kinetics having

only bilinear terms [13]), reduction of model size [11], and implementation of stochastic methods (e.g., Genetic Algorithm [12]) offer avenues for alleviating the computational challenge.

Conclusions

Harnessing the advantages afforded by kinetic descriptions for improving the prediction of cellular metabolism remains an ongoing area of research. Advances in generating fluxomic [58] and metabolomic [46] information and gene-editing techniques [59] offer the opportunity to generate more comprehensive datasets to support the construction of kinetic models beyond model organisms. Construction of standardized databases for fluxomic and metabolomic information for wild-type and mutant strains, similar to repositories for gene and protein expression [60], can improve the accessibility of such information for kinetic model construction. Likewise, a dedicated metabolic engineering database for storing negative results (e.g., BioNOT [61], Negatome [62]) can be a useful resource for mining experimental datasets. The remaining tasks include expansion of models' scope towards genome-scale along with a tight integration of regulatory events and alleviation of the computational burden of model parameterization. Cell-wide modeling approaches [34-36] are providing the framework for integrating regulatory information at different levels for improving overall phenotype prediction and identifying optimal (non-)metabolic intervention strategies. Key considerations such as the context-dependent nature of each type of regulation must be properly accounted for meaningful results. The elucidation of transcription and translation regulatory networks has benefitted from the development of techniques probing gene and protein expression on a massive scale in parallel. Substrate-level regulation has so far not benefitted from a similar throughput expansion in probing techniques. Existing approaches based on developing hypothesis and testing them with *in vitro* experiments [63] are generally slow and sequential in nature. Reliable computational prediction of potential metabolic regulators for an enzyme based on structural (and protein sequence) information, similar to transcriptomic regulation prediction approaches using protein structure information [64], would greatly accelerate this process. At the same time, the development of efficient methods for reducing complex and large kinetic models into more manageable ones without any information loss would increase usability and community acceptance.

Acknowledgement

The authors gratefully acknowledge funding from the DOE (<http://www.energy.gov/>) grant numbers DE-SC0008091 and DE-SC0012377, and the NSF (<http://www.nsf.gov/>) award number EEC-0813570. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Figures:

Figure 1. Metabolic and regulatory interactions. (a) Signal transduction, (b) metabolic feedback inhibition, (c) metabolic feedback activation, (d) transcription factor regulation, (e) Riboswitch binding to small metabolite, (f) sRNA basepairing, (g) Phosphorylation modulation, (h) Allosteric modification, (i) Substrate-level feedforward activation, and (j) Substrate-level feedback inhibition. A list of on-line databases is shown for regulatory information [65].

Figure 2. Integration of kinetic information in stoichiometric frameworks of metabolism for improving strain design accuracy. (a) Hybrid description of metabolic network, (b) Increased flux resolution due to additional kinetic constraints, and (c) Bilevel optimization formulation for identifying metabolic intervention strategies consistent with stoichiometric and kinetic limitations.

1
2
3
4

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. Royce LA, Yoon JM, Chen Y, Rickenbach E, Shanks JV, Jarboe LR: **Evolution for exogenous octanoic acid tolerance improves carboxylic acid production and membrane integrity.** *Metabolic engineering* 2015, **29**:180-188.
2. Link H, Christodoulou D, Sauer U: **Advancing metabolic models with kinetic information.** *Current opinion in biotechnology* 2014, **29**:8-14.
3. van Eunen K, Kiewiet JA, Westerhoff HV, Bakker BM: **Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics.** *PLoS computational biology* 2012, **8**(4):e1002483.
4. Wang L, Birol I, Hatzimanikatis V: **Metabolic control analysis under uncertainty: framework development and case studies.** *Biophysical journal* 2004, **87**(6):3750-3763.
5. Visser D, Heijnen JJ: **Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics.** *Metabolic engineering* 2003, **5**(3):164-176.
6. Sorribas A, Hernandez-Bermejo B, Vilaprinyo E, Alves R: **Cooperativity and saturation in biochemical networks: a saturable formalism using Taylor series approximations.** *Biotechnol Bioeng* 2007, **97**(5):1259-1277.
7. Nielsen J: **Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates.** *The Biochemical journal* 1997, **321** (Pt 1):133-138.
8. Liebermeister W, Klipp E: **Bringing metabolic networks to life: convenience rate law and thermodynamic constraints.** *Theor Biol Med Model* 2006, **3**:41.
9. Almquist J, Cvijovic M, Hatzimanikatis V, Nielsen J, Jirstrand M: **Kinetic models in industrial biotechnology - Improving cell factory performance.** *Metabolic engineering* 2014, **24**:38-60.
10. Tan Y, Liao JC: **Metabolic ensemble modeling for strain engineers.** *Biotechnology journal* 2012, **7**(3):343-353.
- ** A parameter sampling technique is proposed to construct an ensemble of elementary kinetic models all of which converge to the same steady-state flux distribution in the reference mutant strain. Subsequently, various quantitative and/or qualitative experimental data was iteratively utilized to filter out inconsistent models.
11. Chakrabarti A, Miskovic L, Soh KC, Hatzimanikatis V: **Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints.** *Biotechnology journal* 2013, **8**(9):1043-1057.
12. Khodayari A, Zomorodi AR, Liao JC, Maranas CD: **A kinetic model of Escherichia coli core metabolism satisfying multiple sets of mutant flux data.** *Metabolic engineering* 2014, **25**:50-62.

** An optimization-based parameter estimation method is integrated with EM formalism to construct a large-scale kinetic model of *E. coli* central metabolism (including 138 reactions) that satisfies flux data for the wild-type as well as seven mutant strains.

13. Khodayari A, Chowdhury A, Maranas CD: **Succinate Overproduction: A Case Study of Computational Strain Design Using a Comprehensive Escherichia coli Kinetic Model.** *Frontiers in bioengineering and biotechnology* 2014, **2**:76.
14. Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, Consortium D, Kellis M, Collins JJ *et al*: **Wisdom of crowds for robust gene network inference.** *Nature methods* 2012, **9**(8):796-804.
- ** A collaborative effort by computational and experimental biologists to evaluate network inference methods that predict genome-scale transcriptional regulatory networks from gene-expression microarray datasets.
15. Hyduke DR, Lewis NE, Palsson BO: **Analysis of omics data with genome-scale models of metabolism.** *Molecular bioSystems* 2013, **9**(2):167-174.
16. Kotte O, Zaugg JB, Heinemann M: **Bacterial adaptation through distributed sensing of metabolic fluxes.** *Molecular systems biology* 2010, **6**:355.
17. De Smet R, Marchal K: **Advantages and limitations of current network inference methods.** *Nature reviews Microbiology* 2010, **8**(10):717-729.
18. Brooks AN, Reiss DJ, Allard A, Wu WJ, Salvanha DM, Plaisier CL, Chandrasekaran S, Pan M, Kaur A, Baliga NS: **A system-level model for the microbial regulatory genome.** *Molecular systems biology* 2014, **10**:740.
19. Fendt SM, Oliveira AP, Christen S, Picotti P, Dechant RC, Sauer U: **Unraveling condition-dependent networks of transcription factors that control metabolic pathway activity in yeast.** *Molecular systems biology* 2010, **6**:432.
20. Haverkorn van Rijsewijk BR, Nanchen A, Nallet S, Kleijn RJ, Sauer U: **Large-scale ¹³C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in Escherichia coli.** *Molecular systems biology* 2011, **7**:477.
21. Batey RT: **Riboswitches: still a lot of undiscovered country.** *Rna* 2015, **21**(4):560-563.
22. Updegrove TB, Shabalina SA, Storz G: **How do base-pairing small RNAs evolve?** *FEMS microbiology reviews* 2015, **39**(3):379-391.
23. Morris KV, Mattick JS: **The rise of regulatory RNA.** *Nature reviews Genetics* 2014, **15**(6):423-437.
24. Kochanowski K, Sauer U, Noor E: **Posttranslational regulation of microbial metabolism.** *Current opinion in microbiology* 2015, **27**:10-17.
25. Gersch M, Hackl MW, Dubiella C, Dobrinevski A, Groll M, Sieber SA: **A mass spectrometry platform for a streamlined investigation of proteasome integrity, posttranslational modifications, and inhibitor binding.** *Chemistry & biology* 2015, **22**(3):404-411.
26. Strumillo M, Beltrao P: **Towards the computational design of protein post-translational regulation.** *Bioorganic & medicinal chemistry* 2015, **23**(12):2877-2882.
27. Kerkhofs J, Geris L: **A Semiquantitative Framework for Gene Regulatory Networks: Increasing the Time and Quantitative Resolution of Boolean Networks.** *PloS one* 2015, **10**(6):e0130033.
28. Blazier AS, Papin JA: **Integration of expression data in genome-scale metabolic network reconstructions.** *Frontiers in physiology* 2012, **3**:299.
29. Cotten C, Reed JL: **Mechanistic analysis of multi-omics datasets to generate kinetic parameters for constraint-based metabolic models.** *BMC bioinformatics* 2013, **14**:32.

- 1 30. Machado D, Herrgard M: **Systematic evaluation of methods for integration of**
2 **transcriptomic data into constraint-based models of metabolism.** *PLoS*
3 *computational biology* 2014, **10**(4):e1003580.
- 4 31. Saa P, Nielsen LK: **A general framework for thermodynamically consistent**
5 **parameterization and efficient sampling of enzymatic reactions.** *PLoS*
6 *computational biology* 2015, **11**(4):e1004195.
- 7 * A parameter sampling technique is proposed to describe kinetics of oligomeric enzymes using
8 elementary reaction formalism.
- 9
- 10 32. Rybakova KN, Tomaszewska A, van Mourik S, Blom J, Westerhoff HV, Carlberg C,
11 Bruggeman FJ: **Tracing the molecular basis of transcriptional dynamics in noisy**
12 **data by using an experiment-based mathematical model.** *Nucleic acids research*
13 2015, **43**(1):153-161.
- 14 33. Espah Borujeni A, Channarasappa AS, Salis HM: **Translation rate is controlled by**
15 **coupled trade-offs between site accessibility, selective RNA unfolding and**
16 **sliding at upstream standby sites.** *Nucleic acids research* 2014, **42**(4):2646-2659.
- 17 * A biophysiscal model is developed to elucidate how RNA structures in the 5' UTR of
18 bacterial operons control their translation initiation rate.
- 19
- 20 34. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Jr., Assad-Garcia
21 N, Glass JI, Covert MW: **A whole-cell computational model predicts phenotype**
22 **from genotype.** *Cell* 2012, **150**(2):389-401.
- 23 35. O'Brien EJ, Palsson BO: **Computing the functional proteome: recent progress**
24 **and future prospects for genome-scale models.** *Current opinion in biotechnology*
25 2015, **34C**:125-134.
- 26 36. Carrera J, Estrela R, Luo J, Rai N, Tsoukalas A, Tagkopoulos I: **An integrative, multi-**
27 **scale, genome-wide model reveals the phenotypic landscape of Escherichia**
28 **coli.** *Molecular systems biology* 2014, **10**:735.
- 29 37. Zomorodi AR, Suthers PF, Ranganathan S, Maranas CD: **Mathematical**
30 **optimization applications in metabolic networks.** *Metabolic engineering* 2012,
31 **14**(6):672-686.
- 32 38. Schomburg I, Chang A, Placzek S, Sohngen C, Rother M, Lang M, Munaretto C, Ulas S,
33 Stelzer M, Grote A *et al*: **BRENDA in 2013: integrated reactions, kinetic data,**
34 **enzyme function data, improved disease classification: new options and**
35 **contents in BRENDA.** *Nucleic acids research* 2013, **41**(Database issue):D764-772.
- 36 39. Zomorodi AR, Lafontaine Rivera JG, Liao JC, Maranas CD: **Optimization-driven**
37 **identification of genetic perturbations accelerates the convergence of model**
38 **parameters in ensemble modeling of metabolic networks.** *Biotechnology journal*
39 2013, **8**(9):1090-1104.
- 40 40. King ZA, Lloyd CJ, Feist AM, Palsson BO: **Next-generation genome-scale models**
41 **for metabolic engineering.** *Current opinion in biotechnology* 2015, **35C**:23-29.
- 42 41. Cintolesi A, Clomburg JM, Rigou V, Zygorakis K, Gonzalez R: **Quantitative analysis**
43 **of the fermentative metabolism of glycerol in Escherichia coli.** *Biotechnol*
44 *Bioeng* 2012, **109**(1):187-198.
- 45 42. Machado D, Rodrigues LR, Rocha I: **A kinetic model for curcumin production in**
46 **Escherichia coli.** *Bio Systems* 2014, **125**:16-21.
- 47 43. Costa RS, Hartmann A, Gaspar P, Neves AR, Vinga S: **An extended dynamic model**
48 **of Lactococcus lactis metabolism for mannitol and 2,3-butanediol production.**
49 *Molecular bioSystems* 2014, **10**(3):628-639.

- 1 44. Rutkis R, Kalnenieks U, Stalidzans E, Fell DA: **Kinetic modelling of the *Zymomonas***
2 ***mobilis* Entner-Doudoroff pathway: insights into control and functionality.**
3 *Microbiology* 2013, **159**(Pt 12):2674-2689.
- 4 45. Sowa SW, Baldea M, Contreras LM: **Optimizing metabolite production using**
5 **periodic oscillations.** *PLoS computational biology* 2014, **10**(6):e1003658.
- 6 46. Weaver LJ, Sousa MM, Wang G, Baidoo E, Petzold CJ, Keasling JD: **A kinetic-based**
7 **approach to understanding heterologous mevalonate pathway function in *E.***
8 ***coli*.** *Biotechnol Bioeng* 2015, **112**(1):111-119.
- 9 47. Farasat I, Kushwaha M, Collens J, Easterbrook M, Guido M, Salis HM: **Efficient**
10 **search, mapping, and optimization of multi-protein genetic systems in diverse**
11 **bacteria.** *Molecular systems biology* 2014, **10**:731.
- 12 ** The RBS Library Calculator is used to tune the enzyme translation rates of a metabolic
13 pathway that is subsequently used to parameterize and optimize the kinetic model of
14 neurosporene production in *E. coli*.
- 15
- 16 48. Salis HM: **The ribosome binding site calculator.** *Methods in enzymology* 2011,
17 **498**:19-42.
- 18 49. Rollin JA, Martin del Campo J, Myung S, Sun F, You C, Bakovic A, Castro R,
19 Chandrayan SK, Wu CH, Adams MW *et al*: **High-yield hydrogen production from**
20 **biomass by in vitro metabolic engineering: Mixed sugars coutilization and**
21 **kinetic modeling.** *Proceedings of the National Academy of Sciences of the United*
22 *States of America* 2015, **112**(16):4964-4969.
- 23 * A kinetic model is constructed and parameterized (using Genetic Algorithm) to optimize the
24 production of hydrogen from C₅ and C₆ co-substrates in a cell-free system.
- 25
- 26 50. Zhuang K, Yang L, Cluett WR, Mahadevan R: **Dynamic strain scanning**
27 **optimization: an efficient strain design strategy for balanced yield, titer, and**
28 **productivity. DySScO strategy for strain design.** *BMC biotechnology* 2013, **13**:8.
- 29 * An iterative procedure is described where reaction removal suggestions from a stoichiometry-
30 only strain design protocol are implemented on a kinetic model of metabolism to select those
31 combinations that are consistent with both stoichiometric and kinetic limitations.
- 32
- 33 51. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, Lovley DR:
34 **Genome-scale dynamic modeling of the competition between *Rhodoferrax* and**
35 ***Geobacter* in anoxic subsurface environments.** *The ISME journal* 2011, **5**(2):305-
36 316.
- 37 52. Cautha SC, Gowen CM, Lussier F, Gold ND, Martin VJ, Mahadevan R: **Model-Driven**
38 **Design of a *Saccharomyces Cerevisiae* Platform Strain with Improved Tyrosine**
39 **Production Capabilities.** In: *Computer Applications in Biotechnology: 2013*. 221-
40 226.
- 41 53. Chowdhury A, Zomorodi AR, Maranas CD: **Bilevel optimization techniques in**
42 **computational strain design.** *Computers & Chemical Engineering* 2015, **72**:363-
43 372.
- 44 54. Lun DS, Rockwell G, Guido NJ, Baym M, Kelner JA, Berger B, Galagan JE, Church GM:
45 **Large-scale identification of genetic design strategies using local search.**
46 *Molecular systems biology* 2009, **5**:296.
- 47 55. Chowdhury A, Zomorodi AR, Maranas CD: **k-OptForce: integrating kinetics with**
48 **flux balance analysis for strain design.** *PLoS computational biology* 2014,
49 **10**(2):e1003487.

** An optimization strain design protocol is described that directly incorporates enzyme activity and other available substrate-level limitations in a stoichiometry-only genome-scale model to identify improved genetic intervention strategies.

56. Pozo C, Miró A, Guillén-Gosálbez G, Sorribas A, Alves R, Jiménez L: **Gobal optimization of hybrid kinetic/FBA models via outer-approximation.** *Computers & Chemical Engineering* 2015, **72**:325-333.
57. Stanford NJ, Lubitz T, Smallbone K, Klipp E, Mendes P, Liebermeister W: **Systematic construction of kinetic models from genome-scale metabolic networks.** *PloS one* 2013, **8**(11):e79195.
58. Crown SB, Long CP, Antoniewicz MR: **Integrated ¹³C-metabolic flux analysis of 14 parallel labeling experiments in Escherichia coli.** *Metabolic engineering* 2015, **28**:151-158.
59. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.** *Cell* 2013, **152**(5):1173-1183.
60. Petryszak R, Burdett T, Fiorelli B, Fonseca NA, Gonzalez-Porta M, Hastings E, Huber W, Jupp S, Keays M, Kryvych N *et al*: **Expression Atlas update--a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments.** *Nucleic acids research* 2014, **42**(Database issue):D926-932.
61. Agarwal S, Yu H, Kohane I: **BioNOT: a searchable database of biomedical negated sentences.** *BMC bioinformatics* 2011, **12**:420.
62. Blohm P, Frishman G, Smialowski P, Goebels F, Wachinger B, Ruepp A, Frishman D: **Negatome 2.0: a database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis.** *Nucleic acids research* 2014, **42**(Database issue):D396-400.
63. Yu X, Liu T, Zhu F, Khosla C: **In vitro reconstitution and steady-state analysis of the fatty acid synthase from Escherichia coli.** *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(46):18643-18648.
64. Morozov AV, Siggia ED: **Connecting protein structure with predictions of regulatory sites.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(17):7068-7073.
65. Bader GD, Cary MP, Sander C: **Pathguide: a pathway resource list.** *Nucleic acids research* 2006, **34**(Database issue):D504-506.