



HHS Public Access

Author manuscript

Curr Opin Biotechnol. Author manuscript; available in PMC 2019 October 05.

Published in final edited form as:

Curr Opin Biotechnol. 2019 October ; 59: 24–30. doi:10.1016/j.copbio.2019.02.005.

Metabolic kinetic modeling provides insight into complex biological questions, but hurdles remain

Jonathan Strutz^{1,2}, Jacob Martin^{1,2}, Jennifer Greene^{1,2}, Linda Broadbelt¹, Keith Tyo^{1,2}

¹Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA

²Center for Synthetic Biology, Northwestern University, Evanston, IL, USA

Abstract

Metabolic models containing kinetic information can answer unique questions about cellular metabolism that are useful to metabolic engineering. Several kinetic modeling frameworks have recently been developed or improved. In addition, techniques for systematic identification of model structure, including regulatory interactions, have been reported. Each framework has advantages and limitations, which can make it difficult to choose the most appropriate framework. Common limitations are data availability and computational time, especially in large-scale modeling efforts. However, recently developed experimental techniques, parameter identification algorithms, as well as model reduction techniques help alleviate these computational bottlenecks. Opportunities for additional improvements may come from the rich literature in catalysis and chemical networks. In all, kinetic models are positioned to make significant impact in cellular engineering.

Introduction

Metabolic networks are highly integrated networks that are regulated in sophisticated ways. Because of this complexity, strategies for metabolic engineering are often non-intuitive, and engineering questions are often best addressed by metabolic models. Constraint-based models (CBMs), based on stoichiometry, have been widely successful over the last 15 years at guiding engineering efforts without the need for mechanistic detail [1]. However, CBMs cannot capture the relationship between flux, enzyme expression, metabolite levels, and regulation that is possible with kinetic models (Box 1) [2*]. Although computationally costly, kinetic models are more predictive and are especially appropriate when there is not an obvious objective function for optimization or when exploring dynamic effects [3]. However, within kinetic modeling, it can be difficult to determine where to start due to the great wealth of published frameworks. Here we highlight the questions that are well suited for kinetic models and the various hurdles to their use.

Corresponding author: Tyo, Keith (k-tyo@northwestern.edu).

Conflict of interest statement

Nothing declared.

Questions addressed by recent kinetic modeling frameworks

Recent kinetic modeling frameworks primarily seek to answer four types of questions: those involving (1) metabolic state prediction and engineering strategies, (2) identification of unmodeled phenomena, (3) metabolic stability, and (4) kinetic variation. The relative strengths of each framework are shown in Table 1.

Metabolic state prediction and engineering strategies

Most kinetic modeling frameworks are designed to address questions involving prediction of metabolic states or rate-limiting steps, as these questions are central to metabolic engineering efforts. As kinetic parameters of individual enzymes are rarely known or are uncertain, many frameworks rely on exploring a range of parameters and selecting a subset that are consistent with experimental observations. Using the parameters derived from model training, network kinetics can be analyzed and predictions made for potential engineering strategies. The Ensemble Modeling (EM) paradigm is widely used because it requires minimal data and can accommodate large uncertainties in kinetic parameters [4]. EM explores parameters that are thermodynamically consistent with models that collectively describe dynamic and steady-state behavior. Because many parameters might be equally consistent with experimental data, the range of parameters effectively captures prediction uncertainty (i.e. a small range means low uncertainty, and a large range means high uncertainty). ABC-GRASP is similar in concept to EM but uses a wider range of kinetics and uses probability distributions, rather than discrete parameter sets [5]. However, these features can add a large computational cost relative to EM.

The ORACLE framework incorporates Metabolic Control Analysis (MCA) to succinctly characterize steady-state behavior and predict rate-limiting steps [6]. Recently, two extensions to the ORACLE framework were developed that significantly expand its utility. iSCHRUNK uses machine learning tools on the final set of kinetic parameters to further reduce uncertainty [7]. In another work, inverse MCA (IMCA) was integrated into ORACLE to incorporate the effects of transcriptional regulation on enzyme expression, which affects steady-state behavior [8]. However, validating predicted enzyme expression rate against measured mRNA concentration changes was challenging, potentially due to the low correlation between mRNA and enzyme expression levels [9].

Regulatory structure inference

While kinetic modeling can be used to test the presence/absence of structures in the network and regulation [5,8], only a small fraction of the possibilities can typically be explored due to computational limitations. Two notable exceptions include a method where dynamic metabolite data were used to systematically test a large set of putative allosteric interactions [10] and another method where Michaelis–Menten rate laws with variable allosteric terms were fit to an -omics dataset on a reaction-by-reaction basis [11].

Recently, however, numerous untargeted experimental strategies which do not require kinetic modeling have also been developed [12^{**},13,14^{*}]. For example, LiP-SMap detects allosteric interactions by analyzing the difference in the protease cleavage sites of proteins

when bound to a given metabolite [12**]. Another study used high-throughput metabolomics data from a single-gene knockout library to infer gene-metabolite relationships [13]. Finally, a library of fluorescent transcriptional reporters was used to quantify the activity of metabolic promoters in *Escherichia coli*, and their activity was correlated with metabolome response to identify metabolites involved in transcriptional regulation [14*]. These experimental methods can be used to identify regulatory mechanisms that can be incorporated into kinetic models.

Metabolic stability

Metabolic pathways are stabilized during evolution, so cells are typically robust to perturbations to native enzyme concentrations. However, the addition of heterologous pathways may not result in similar steady-state stability. While frameworks predicting metabolic states can assess steady-state stability indirectly, Ensemble Modeling for Robustness Analysis (EMRA) directly calculates the likelihood of a perturbation causing metabolic instability [15]. Recently, EMRA was used to assess the stability of cell-free systems, which are not optimized via evolution [16]. EMRA was also used to show how incorporation of kinetics and stability constraints further constrains the Maximum Theoretical Yield as predicted by FBA [17*].

Kinetic variation

While there is variability in the metabolome and proteome among cell strains, it is thought that variation in the kinetome – the space of kinetic rate constants in metabolism – correlates better with genetic variation [18]. The MASS (Mass Action Stoichiometric Simulation) framework can perform large-scale analysis of kinetic variation [19]. Because MASS assumes mass action kinetics (e.g. $\text{Rate} = k_1[A][B]$), a population of thermodynamically consistent rate constant sets can be efficiently computed and studied. While mass action kinetics is an oversimplification of many metabolic reactions, comparison of the fitted rate constants can reveal loci of important variation between systems, for example, in red blood cells from different patients [18] as well as metabolic response to specific drugs [20].

Hurdles to large-scale quantitative predictions

While kinetic models can address complex biological questions, data limitations can result in large uncertainties in predictions, and large models can result in computational intractability. Here, we discuss these hurdles, focusing on recent developments to overcome them.

Challenges in experimental measurements used to train kinetic models

The most important type of data for kinetic models is structure. At a minimum, structure includes all reactions relevant to the system but may also include regulatory interactions and cell growth. While most reaction networks are well-defined, promiscuous enzyme activity and heterologous pathways can have unforeseen effects. Tools such as BNICE [21], Pickaxe [22], Retropath2.0 [23], and novoStoic [24] can elucidate these reactions through comparison to known reaction chemistry [25].

Because kinetic models ultimately depict system properties as a function of kinetic parameters, there have been several efforts in constructing ‘bottom-up’ models of metabolism where one uses directly measured or predicted kinetic properties of enzymes in model parameterization. While much progress has been made, using *in vitro* enzyme properties can cause unrealistic model behavior without the manual curation of regulatory effects [26,27]. Recent efforts [28] have shown that Michaelis–Menten rate law approximations using kinetic data can replace detailed rate laws; however, as enzyme kinetic information remains sparse, this approach is only valid for a small number of well-characterized reactions.

More commonly, -omics measurements are used to train kinetic models. However, different kinetic modeling frameworks utilize different types of -omics data to varying degrees (Figure 1); therefore, data types readily available for a given project should be taken into consideration when deciding on a kinetic modeling framework.

Flux data are arguably the most difficult to obtain. Nevertheless, fluxes are generally the most important property to measure, as most frameworks, including ORACLE, EM, ABC-GRASP, and MASS, require an accurate intracellular reference flux distribution. ¹³C tracer studies elucidate aspects of intracellular metabolism, while uptake and secretion fluxes are obtained by measuring changes in extracellular metabolite concentrations.

Measurements of intracellular metabolomics are also useful for parameter inference, but they are difficult to obtain. Metabolomics can refer to either relative or absolute concentrations of metabolites. While relative measurements – which do not require internal standards and are higher throughput – can typically be incorporated in rate laws, absolute metabolomics are required to accurately implement thermodynamic constraints [29]. Reaction thermodynamics strongly control fluxes for most reversible reactions and thus help to ensure feasibility in kinetic models [11]. For a thorough review of metabolomics methods, see Ref. [30].

Proteomics, while a low-throughput measurement, can be directly incorporated into most kinetic modeling frameworks. This is especially true in the case of highly irreversible reactions in which fluxes are often sensitive to enzyme concentration [11]. Relative proteomics are currently higher throughput than absolute measurements, although label-free absolute proteomics methods have made recent advances [31].

Parameter identifiability

Usually, experimental datasets measure too few conditions or not enough fluxes/metabolites/proteins at a given condition to resolve all parts of a network, that is, those parameters lack identifiability with the data provided. Methods for determining parameter and structural identifiability are complex and previously reviewed [32]. However, model solutions containing many feasible model parameterizations, including aspects of the network that are non-identifiable, are typically sufficient to answer design hypotheses, even if individual members of the solution contain loosely fit parameters [33]. We especially want to highlight recent modeling efforts that result in a distribution of parameters, rather than just one optimal set, and provide impactful insight into system behavior despite lacking a sufficient

amount of data for direct parameter identification [15,16,34,35]. In general, kinetic models are best positioned to identify overall changes in system behavior rather than individual parameter values, as different sets of kinetic parameters can give rise to the same overall behavior.

Most kinetic models are written as large systems of stiff ordinary differential equations (ODEs) that are computationally expensive to solve. However, optimization tools such as Markov Chain Monte Carlo (MCMC) [36], genetic [37], and particle swarm [38,39] algorithms can decrease computational burden. For example, a parallelizable scatter search method was developed [38]. This method initiates multiple threads of a search algorithm, and the combination of these results is then communicated between parallel threads to better search the solution space. Additionally, a genetic algorithm (GA) step was added to the traditional EM framework to identify a single parameter set rather than an ensemble [37]. However, the lack of an ensemble makes it challenging to gauge uncertainty as there is no dispersion in one model, so several GA searches should be deployed in parallel. Furthermore, while the GA step more efficiently searches the parameter space, it often results in lower overall computational efficiency due to parameter tuning and overhead costs. Thus, these tools may not always reduce computational burden. However, other techniques that reduce the parameter sampling space directly by, for example, imposing reference flux states, reaction thermodynamic limits, and local stability constraints often reduce computational burden [4,16,29,34,40,41*]. While these strategies efficiently search and constrain the kinetic parameter space, directly reducing the number of kinetic parameters can also improve computational tractability.

Model reduction

The goal of model reduction is to reduce the number of kinetic parameters while maintaining predictive power. Although reduction techniques are often applied after parameter estimation [42–45], here we highlight *a priori* model reduction methods applicable to most kinetic modeling efforts.

1. Model scope - selecting the appropriate amount of detail to include before using post-parameter estimation reduction techniques [42–45] is an ill-defined process. Too much detail leads to an impractical amount of time spent on model reduction (due to large computational costs), while too little detail may not capture observed effects at all.
2. Rate expression - the type of rate expression determines the number of parameters required. Approximate rate laws, such as lin-log kinetics, are model-reducing as they use fewer kinetic parameters [46]. However, they typically invoke assumptions whose validity should be checked [2*,47]. This natural trade-off should be considered during initial model development [28,48].
3. Conservation analysis - biological networks inherently contain conserved moieties, such as the total pool of ATP, ADP, and AMP. These conserved groups are considered independent variables in many kinetic modeling approaches and increase the stiffness of ODE systems, increasing solve time. Because it is difficult to manually identify all conserved moieties, a tool that performs

conservation analysis on biological networks was developed [49]. We have found that using this method decreases computation time.

4. Lumping - lumping techniques, such as a condensing a linear pathway into a single reaction, can greatly shrink the parameter space [50,51]. Often, optimization methods are used to identify potential lumping schemes throughout the network [26,52].

Use of appropriate model reduction techniques can greatly reduce the resources required to interrogate large-scale biological networks [43]. Conservation analysis, in particular, is easy to implement [49] and would provide value to most kinetic modeling efforts.

Conclusions

While we have reviewed the types of questions addressed by recent kinetic modeling frameworks, comparative studies need to be done in order to further characterize similar kinetic frameworks, similar to comparisons done for CBM frameworks [53]. In addition, it is important to survey modeling efforts in other fields, such as heterogeneous catalysis [54–56], and especially efforts that take advantage of the increasing availability of -omics data, particularly fluxomics, intracellular metabolomics, and proteomics [57**]. Improving parameter estimation and model reduction techniques will allow tractable simulations for large-scale kinetic models while the discussed structural inference techniques can best inform model structure. As metabolic networks are large, complicated, and highly-coupled, sophisticated modeling frameworks will be essential in predicting system-level behavior.

Acknowledgements

This work was supported by the National Science Foundation (J.S., L.B., K.T. MCB-1614953, J.G. DGE-1324585), the National Institute of Health (J.S. T32-GM008449–23), and the Department of Energy (J.M., J.G., L.B., K.T. DEEE0007728 and DESC0018249).

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. Kim WJ, Kim HU, Lee SY: Current state and applications of microbial genome-scale metabolic models. *Curr Opin Syst Biol* 2017, 2:10–18.
 - 2 • Saa PA, Nielsen LK: Formulation, construction and analysis of kinetic models of metabolism: a review of modeling frameworks. *Biotechnol Adv* 2017, 35:981–1003 10.1016/j.biotechadv.2017.09.005. [PubMed: 28916392] This very comprehensive review of advanced kinetic modeling frameworks looks at the numerous rate law forms commonly used, discusses recent kinetic modeling frameworks in great detail, and then does a great in-depth comparative analysis of these frameworks.
 3. Tummler K, Klipp E: The discrepancy between data for and expectations on metabolic models: how to match experiments and computational efforts to arrive at quantitative predictions? *Curr Opin Syst Biol* 2018, 8:1–6.
 4. Tran LM, Rizk ML, Liao JC: Ensemble modeling of metabolic networks. *Biophys J* 2008, 95:5606–5617. [PubMed: 18820235]

5. Saa PA, Nielsen LK: Construction of feasible and accurate kinetic models of metabolism: a Bayesian approach. *Sci Rep* 2016, 6.
6. Miskovic L, Hatzimanikatis V: Production of biofuels and biochemicals: in need of an ORACLE. *Trends Biotechnol* 2010, 28:391–397. [PubMed: 20646768]
7. Andreozzi S, Miskovic L, Hatzimanikatis V: iSCHRUNK—in silico approach to characterization and reduction of uncertainty in the kinetic models of genome-scale metabolic networks. *Metab Eng* 2016, 33:158–168. [PubMed: 26474788]
8. Savoglidis G, da Silveira dos Santos AX, Riezman I, Angelino P, Riezman H, Hatzimanikatis V: A method for analysis and design of metabolism using metabolomics data and kinetic models: application on lipidomics using a novel kinetic model of sphingolipid metabolism. *Metab Eng* 2016, 37:46–62. [PubMed: 27113440]
9. Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, Goodlett DR, Kruglyak L: Genetic basis of proteome variation in yeast. *Nat Genet* 2007, 39:1369–1375. [PubMed: 17952072]
10. Link H, Kochanowski K, Sauer U: Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nat Biotechnol* 2013, 31:357–361. [PubMed: 23455438]
11. Hackett SR, Zanotelli VRT, Xu W, Goya J, Park JO, Perlman DH, Gibney PA, Botstein D, Storey JD, Rabinowitz JD: Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* (80-) 2016, 354.
12. •• Piazza I, Kochanowski K, Cappelletti V, Fuhrer T, Noor E, Sauer U, Picotti P: A map of protein-metabolite interactions reveals principles of chemical communication. *Cell* 2018, 172:358–372. e23. [PubMed: 29307493] Here, the authors present LiP-SMap, a high-throughput chemoproteomic workflow to systematically identify metabolite-level regulatory interactions on a proteome-wide scale. It does not require metabolomics measurements and is the first method of its kind that can be performed directly in the native cellular matrix.
13. Fuhrer T, Zampieri M, Sévin DC, Sauer U, Zamboni N: Genomewide landscape of gene–metabolome associations in *Escherichia coli*. *Mol Syst Biol* 2017, 13:907. [PubMed: 28093455]
14. • Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev YV, Sauer U: Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli*. *Mol Syst Biol* 2017, 13:903. [PubMed: 28049137] The authors developed a data-driven, untargeted technique to unravel transcriptional regulatory networks using a library of fluorescent transcriptional reporters and high-throughput metabolomics data. Interestingly, they show that only a few metabolites, along with global transcriptional regulation, caused 90% of enzyme expression changes in *E. coli* central metabolism.
15. Lee Y, Lafontaine Rivera JG, Liao JC: Ensemble modeling for robustness analysis in engineering non-native metabolic pathways. *Metab Eng* 2014, 25:63–71. [PubMed: 24972370]
16. Theisen MK, Lafontaine Rivera JG, Liao JC: Stability of ensemble models predicts productivity of enzymatic systems. *PLOS Comput Biol* 2016, 12:e1004800. [PubMed: 26963521]
17. • Lafontaine Rivera JG, Theisen MK, Chen PW, Liao JC: Kinetically accessible yield (KAY) for redirection of metabolism to produce exo-metabolites. *Metab Eng* 2017, 41:144–151. [PubMed: 28389394] The maximum theoretical yield (MTY) is commonly used to evaluate the potential of heterologous pathways. However, this work shows that kinetically accessible yield (KAY), which takes into account metabolic stability, is a more meaningful indicator of pathway performance than MTY.
18. Bordbar A, McCloskey D, Zielinski DC, Sonnenschein N, Jamshidi N, Palsson BO: Personalized whole-cell kinetic models of metabolism for discovery in genomics and pharmacodynamics. *Cell Syst* 2015, 1:283–292. [PubMed: 27136057]
19. Jamshidi N, Palsson BØ: Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models. *Biophys J* 2010, 98:175–185. [PubMed: 20338839]
20. Mih N, Brunk E, Bordbar A, Palsson BO: A multi-scale computational platform to mechanistically assess the effect of genetic variation on drug responses in human erythrocyte metabolism. *PLOS Comput Biol* 2016, 12:e1005039. [PubMed: 27467583]

21. Hatzimanikatis V, Li CH, Ionita JA, Henry CS, Jankowski MD, Broadbelt LJ: Exploring the diversity of complex metabolic networks. *Bioinformatics* 2005, 21:1603–1609. [PubMed: 15613400]
22. Jeffries JG, Colastani RL, Elbadawi-Sidhu M, Kind T, Niehaus TD, Broadbelt LJ, Hanson AD, Fiehn O, Tyo KEJ, Henry CS: MINEs: open access databases of computationally predicted enzyme promiscuity products for untargeted metabolomics. *J Cheminform* 2015, 7:44. [PubMed: 26322134]
23. Delépine B, Duigou T, Carbonell P, Faulon J-L: RetroPath2.0: a retrosynthesis workflow for metabolic engineers. *Metab Eng* 2018, 45:158–170. [PubMed: 29233745]
24. Kumar A, Wang L, Ng CY, Maranas CD: Pathway design using de novo steps through uncharted biochemical spaces. *Nat Commun* 2018, 9.
25. Wang L, Dash S, Ng CY, Maranas CD: A review of computational tools for design and reconstruction of metabolic pathways. *Synth Syst Biotechnol* 2017, 2:243–252. [PubMed: 29552648]
26. Rao S, van der Schaft A, van Eunen K, Bakker BM, Jayawardhana B: A model reduction method for biochemical reaction networks. *BMC Syst Biol* 2014, 8:52 10.1186/1752-0509-8-52. [PubMed: 24885656]
27. Heijnen JJ, Verheijen PJT: Parameter identification of in vivo kinetic models: limitations and challenges. *Biotechnol J* 2013, 8:768–775. [PubMed: 23813763]
28. Du B, Zielinski DC, Kavvas ES, Dräger A, Tan J, Zhang Z, Ruggiero KE, Arzumanyan GA, Palsson BO: Evaluation of rate law approximations in bottom-up kinetic models of metabolism. *BMC Syst Biol* 2016, 10:40. [PubMed: 27266508]
29. Henry CS, Jankowski MD, Broadbelt LJ, Hatzimanikatis V: Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys J* 2006, 90:1453–1461. [PubMed: 16299075]
30. Zampieri M, Sekar K, Zamboni N, Sauer U: Frontiers of highthroughput metabolomics. *Curr Opin Chem Biol* 2017, 36:15–23. [PubMed: 28064089]
31. Maaß S, Becher D: Methods and applications of absolute protein quantification in microbial systems. *J Proteomics* 2016, 136:222–233. [PubMed: 26825536]
32. Villaverde AF, Barreiro A: Identifiability of large nonlinear biochemical networks. *Commun Math Comput Chem* 2016, 76:259–296.
33. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP: Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol* 2007, 3:e189.
34. Lafontaine Rivera JG, Lee Y, Liao JC: An entropy-like index of bifurcational robustness for metabolic systems. *Integr Biol* 2015, 7:895–903 10.1039/C4IB00257A.
35. Zomorodi AR, Lafontaine Rivera JG, Liao JC, Maranas CD: Optimization-driven identification of genetic perturbations accelerates the convergence of model parameters in ensemble modeling of metabolic networks. *Biotechnol J* 2013, 8:1090–1104. [PubMed: 23450699]
36. Kramer A, Stathopoulos V, Girolami M, Radde N: MCMC_CLIB-an advanced MCMC sampling package for ODE models. *Bioinformatics* 2014, 30:2991–2992. [PubMed: 25005749]
37. Khodayari A, Maranas CD: A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat Commun* 2016, 7:13806. [PubMed: 27996047]
38. Penas DR, González P, Egea JA, Doallo R, Banga JR: Parameter estimation in large-scale systems biology models: a parallel and self-adaptive cooperative strategy. *BMC Bioinform* 2017, 18:52.
39. Villaverde AF, Egea JA, Banga JR: A cooperative strategy for parameter estimation in large scale systems biology models. *BMC Syst Biol* 2012, 6:1 10.1186/1752-0509-6-75. [PubMed: 22222070]
40. Tan Y, Rivera JGL, Contador CA, Asenjo JA, Liao JC: Reducing the allowable kinetic space by constructing ensemble of dynamic models with the same steady-state flux. *Metab Eng* 2011, 13:60–75. [PubMed: 21075211]
- 41 •. Greene JL, Wächter A, Tyo KEJ, Broadbelt LJ: Acceleration strategies to enhance metabolic ensemble modeling performance. *Biophys J* 2017, 113:1150–1162. [PubMed: 28877496] The authors show how strategies such as performing conservation analysis, presorting training data, and sampling local stability provide large benefits to computational efficiency and that these benefits scale favorably with increasing model size. Because kinetic modeling frameworks on

their own generally do not scale well to large networks, developing acceleration strategies for all frameworks (especially those that parameterize models as systems of differential equations) will be invaluable for large-scale kinetic modeling efforts.

42. Snowden TJ, van der Graaf PH, Tindall MJ: Methods of model reduction for large-scale biological systems: a survey of current methods and trends. *Bull Math Biol* 2017, 79:1449–1486 10.1007/s11538-017-0277-2. [PubMed: 28656491]
43. Snowden TJ, van der Graaf PH, Tindall MJ: A combined model reduction algorithm for controlled biochemical systems. *BMC Syst Biol* 2017, 11:1 10.1186/s12918-017-0397-1. [PubMed: 28061857]
44. Liepe J, Filippi S, Komorowski M, Stumpf MPH: Maximizing the information content of experiments in systems biology. *PLoS Comput Biol* 2013, 9:e1002888 10.1371/journal.pcbi.1002888. [PubMed: 23382663]
45. Zomorodi AR, Suthers PF, Ranganathan S, Maranas CD: Mathematical optimization applications in metabolic networks. *Metab Eng* 2012, 14:672–686. [PubMed: 23026121]
46. Wu L, Wang W, Van Winden WA, Van Gulik WM, Heijnen JJ: A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 2004, 271:3348–3359. [PubMed: 15291812]
47. Nikerel IE, van Winden WA, Verheijen PJT, Heijnen JJ: Model reduction and a priori kinetic parameter identifiability analysis using metabolome time series for metabolic reaction networks with linlog kinetics. *Metab Eng* 2009, 11:20–30 10.1016/j.ymben.2008.07.004. [PubMed: 18718548]
48. Schmidt H, Madsen MF, Danø S, Cedersund G: Complexity reduction of biochemical rate expressions. *Bioinformatics* 2008, 24:848–854 10.1093/bioinformatics/btn035. [PubMed: 18267948]
49. Bedaso Y, Bergmann FT, Choi K, Sauro HM: A portable structural analysis library for reaction networks. *bioRxiv* 2018 10.1101/245068.
50. Röhl A, Bockmayr A: A mixed-integer linear programming approach to the reduction of genome-scale metabolic networks. *BMC Bioinformatics* 2017, 18:1 10.1186/s12859-016-1412-z. [PubMed: 28049414]
51. Ataman M, Hernandez Gardiol DF, Fengos G, Hatzimanikatis V: RedGEM: systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models. *PLoS Comput Biol* 2017, 13:e1005444 10.1371/journal.pcbi.1005444. [PubMed: 28727725]
52. Erdrich P, Steuer R, Klamt S: An algorithm for the reduction of genome-scale metabolic network models to meaningful core models. *BMC Syst Biol* 2015, 9:1 10.1186/s12918-015-0191-x. [PubMed: 25582171]
53. Machado D, Herrgård M: Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput Biol* 2014, 10:e1003989.
54. Villaverde AF, Banga JR: Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *J R Soc Interface* 2014, 11:20130505. [PubMed: 24307566]
55. Han K, Green WH, West RH: On-the-fly pruning for rate-based reaction mechanism generation. *Comput Chem Eng* 2017, 100:1–8.
56. Hough BR, Beck DAC, Schwartz DT, Pfaendtner J: Application of machine learning to pyrolysis reaction networks: reducing model solution time to enable process optimization. *Comput Chem Eng* 2017, 104:56–63.
- 57 •• Costello Z, Martin HG: A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *npj Syst Biol Appl* 2018, 4:19. [PubMed: 29872542] Their generalizable machine learning method requires no network structure, only time-course -omics data, yet it outperformed Michaelis–Menten models of two metabolic pathways. Although the authors use one of the largest time-course proteomics and metabolomics datasets available, this work demonstrates that purely data-driven approaches can successfully predict biological dynamics.

Box 1**Type of question informs the type of model used****CBMs are more appropriate for some types of questions:****Flux distribution (during growth):**

What does the intracellular flux distribution look like?

Growth rate:

How might the ratio of media components Y1 and Y2 affect growth?

Knockouts (during growth):

Which enzyme(s) should be knocked out to increase flux through pathway P?

Maximum theoretical yield (MTY):

How does the MTY of product X change if I change media composition?

Kinetic models are better suited for others:**State prediction:**

Which enzyme(s) should I overexpress to increase production of metabolite X?

Knockouts (during non-growth):

Which enzyme(s) should be knocked out to increase flux through pathway P during non-growth conditions?

Metabolic stability:

Will incorporating heterologous pathway P limit productivity due to metabolic instability? How much can I overexpress enzyme E without losing stability?

Regulatory interactions:

Is there an allosteric interaction between enzyme E and metabolite X?

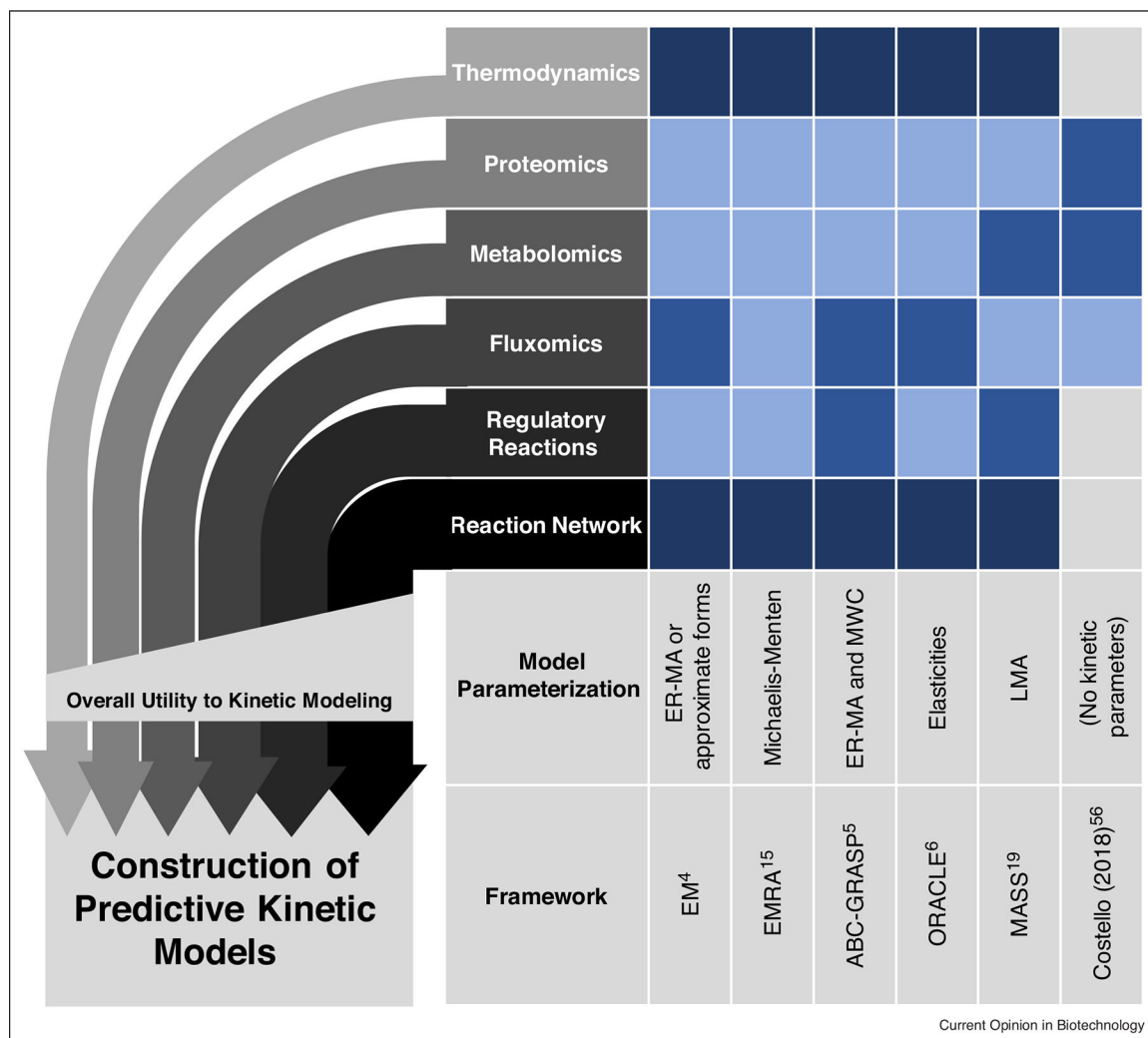


Figure 1.

Different data types have varying value to kinetic modeling frameworks.

-omics types are required to varying levels by different model frameworks (darkest boxes indicate data type is required, less dark boxes indicate data type is used to a high degree in practice, lighter blue boxes indicate data type can be used, light gray boxes indicate data type is not used) [4,5,6,15,19,57**]. While all the -omics data types shown have utility in kinetic modeling, modeling results are usually most sensitive to variation in those near the bottom (e.g. variation in network structure). Thus, those data types generally provide more utility to kinetic modeling efforts and should be prioritized. Note that while regulatory reactions provide much value to kinetic modeling, they are not always incorporated, either because they are unknown or because they cannot be incorporated easily using a given framework. Data-driven models, while requiring very large amounts of data, may not require knowledge of the reaction network or regulatory interactions at all [57**].

ABC-GRASP, Approximate Bayesian Computation – General Reaction Assembly and Sampling Platform; EM, Ensemble Modeling; EMRA, Ensemble Modeling for Robustness Analysis; ER-MA, Elementary Reaction Mass Action; LMA, Law of Mass Action; MASS,

Mass Action Stoichiometric Simulation; MWC, Monod-Wyman-Changeux; ORACLE, Optimization and Risk Analysis of Complex Living Entities.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Applicability of mechanistic and data-driven metabolic modeling and inference frameworks

Framework name	Metabolic engineering strategies				Applicability of framework in providing insights regarding:				Resource availability			
	Rate-limiting steps	State prediction	Metabolite-level	Transcriptional	Metabolic stability	Kinetic variation	Experimental	Computational	Scalability	Recent applications		
Mechanistic, targeted	**	***	*	***	***	**	***	**	***	[16,17]		
EM-RA	***	***	***	***	***	**	***	*	*	[5]		
ABC-GRASP	***	*		*	*	**	**	***	**	[8]		
ORACLE+ IMCA	***	*			*	**	***	***	**	[7]		
ORACLE+ iSCHRUNK	***	*			*	**	***	***	***	[18,20]		
MASS	*	*			*	***	**	***	***	[12**]		
Data-driven, untargeted			***				**	***	***	[13]		
LiP-SMap			***				*	***	***	[14*]		
Fuhrer [13]			***				*	***	***	[11]		
Kochanowski [14*]			***				*	***	***	[57**]		
SIMMER			***				*	***	***			
Costello [57**]	***	***	*	*	*		*	**	***			

Key:

[] = N/A

[*] = Low

[**] = Medium

[***] = High