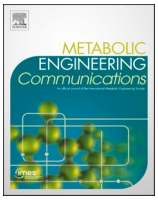
Metabolic Engineering Communications 18 (2024) e00234

Contents lists available at ScienceDirect

Metabolic Engineering Communications

journal homepage: www.elsevier.com/locate/mec



Bottom-up parameterization of enzyme rate constants: Reconciling inconsistent data

Daniel C. Zielinski [a,1], Marta R.A. Matos [b,1], James E. de Bree [a,1], Kevin Glass [a], Nikolaus Sonnenschein [b], Bernhard O. Palsson [a,b,c,*]

[a] Department of Bioengineering, University of California, San Diego, CA, 92093, USA
[b] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[c] Department of Pediatrics, University of California, San Diego, CA, 92093, USA

ARTICLE INFO

Handling Editor: Mattheos Koffas

ABSTRACT

Kinetic models of metabolism are promising platforms for studying complex metabolic systems and designing production strains. Given the availability of enzyme kinetic data from historical experiments and machine learning estimation tools, a straightforward modeling approach is to assemble kinetic data enzyme by enzyme until a desired scale is reached. However, this type of 'bottom up' parameterization of kinetic models has been difficult due to a number of issues including gaps in kinetic parameters, the complexity of enzyme mechanisms, inconsistencies between parameters obtained from different sources, and *in vitro-in vivo* differences. Here, we present a computational workflow for the robust estimation of kinetic parameters for detailed mass action enzyme models while taking into account parameter uncertainty. The resulting software package, termed MASSef (the Mass Action Stoichiometry Simulation Enzyme Fitting package), can handle standard 'macroscopic' kinetic parameters, including K_m, k_{cat}, K_i, K_{eq}, and n_h, as well as diverse reaction mechanisms defined in terms of mass action reactions and 'microscopic' rate constants. We provide three enzyme case studies demonstrating that this approach can identify and reconcile inconsistent data either within *in vitro* experiments or between *in vitro* and *in vivo* enzyme function. We further demonstrate how parameterized enzyme modules can be used to assemble pathway-scale kinetic models consistent with *in vivo* behavior. This work builds on the legacy of knowledge on kinetic behavior of enzymes by enabling robust parameterization of enzyme kinetic models at scale utilizing the abundance of historical literature data and machine learning parameter estimates.

1. Introduction

There has been a resurgence of interest in the construction of large-scale kinetic models of metabolism for model organisms in recent years (Millard et al. 2017; Srinivasan et al. 2015; Foster et al., 2019). These models hold promise in a number of applications that constraint-based models have difficulty addressing, such as understanding quantitatively how metabolite levels control metabolic flux across experimental conditions(Link et al. 2013; Andreozzi et al., 2016; Savoglidis et al., 2016; Chowdhury et al. 2015). However, the primary issue impeding the development of practical large-scale kinetic models of metabolism is the need for a large number of kinetic parameters, the vast majority of which have not been experimentally measured(Heijnen and Verheijen 2013).

To address this parameterization challenge, a number of approaches have been developed(P. A. Saa and Nielsen 2017; Shepelin et al., 2020). Sampling methods randomly select parameters in particular expected ranges, sometimes with constraints such as thermodynamic consistency enforced in the sampling procedure(P. Saa and Nielsen 2015; Miskovic and Hatzimanikatis 2010; Tran et al. 2008). Top-down parameterization methods use data on the kinetic or steady-state behavior of the entire system and parameterize the entire model simultaneously to match this data(Chassagnole et al., 2002; Khodayari and Maranas 2016; Jamshidi and Palsson 2010). Bayesian methods have been proposed as a flexible and powerful tool for parameterization (Linden et al. 2022; P. A. Saa and Nielsen 2016). A number of software tools are available for parameterization using these methods(Adams et al., 2013; Gábor et al. 2017; Hoops et al., 2006; Gábor and Banga 2015). However, performance of

* Corresponding author. Department of Bioengineering, University of California, San Diego, CA, 92093, USA.
E-mail address: bpalsson@ucsd.edu (B.O. Palsson).
[1] equal contribution.

https://doi.org/10.1016/j.mec.2024.e00234
Received 7 February 2024; Received in revised form 16 April 2024; Accepted 18 April 2024
Available online 23 April 2024

such methods can suffer due to sampling too wide a parameter space or having too many parameters for the limited amount of data.

Bottom-up methods on the other hand use data related to the individual components of the network to construct a model piece by piece (Kuzmic 1996; Olp et al. 2020; Choi et al. 2017; Ishii et al., 2007; Teusink et al., 2000). Bottom-up methods have the advantage of utilizing the extensive amount of historical enzyme data(Placzek et al., 2017; Costa et al. 2014) as well as machine learning estimates(Heckmann et al., 2018; Li et al., 2022). Further, recent high-throughput studies have shown substantial correlations between *in vivo* and *in vitro* kinetic parameters, supporting the use of available enzyme kinetic data(Davidi et al., 2016; Bennett et al., 2009). Our group and others have demonstrated that detailed enzyme kinetics can be inserted into a simplified kinetic background to construct large-scale kinetic models of metabolism(Jamshidi and Palsson 2010; Haiman et al., 2021; Du et al., 2016). However, there are a number of difficulties with bottom-up construction of kinetic models that have impeded their development. First, the majority of enzymes do not have detailed kinetic assays performed, requiring additional parameter estimation(Kotte and Heinemann 2009). Second, parameters from disparate sources may be inconsistent, requiring a framework to integrate this data(Choudhury et al., 2022; P. Saa and Nielsen 2015). If these challenges can be met, bottom-up methods may be complementary alternatives for construction of large-scale kinetic models of metabolism.

Here, we present a computational workflow that attempts to address a number of issues with bottom-up construction of kinetic models of metabolism. This workflow parameterizes individual enzyme kinetic models to fit a variety of measured kinetic data types (e.g. k_{cat}, K_M), attempts to reconcile inconsistent data, and accounts for parameter uncertainty. To aid this, we utilize flexible user-defined microscopic mass action reaction mechanisms that can account for the vast majority of observed kinetic behavior and can be extended to the necessary resolution on an enzyme-by-enzyme basis. This workflow is implemented as a software package in Mathematica termed MASSef (Mass Action Stoichiometric Simulation Enzyme Fitting). There are three core features of this software package: 1) a symbolic algebra system that generates equations for comparison with measured data based upon the enzyme mechanism, 2) robust nonlinear optimization to fit the model to data, and 3) a workflow that perturbs the fitting problem to characterize uncertainty in the parameters. We first present an overview of the computational workflow for parameter fitting before discussing the details of individual components, and then we present a series of case studies demonstrating the workflow for enzymes with different available kinetic data. Finally, we demonstrate a workflow to assemble pathway-scale kinetic models by inserting these parameterized enzyme modules in an approximate mass action kinetic model background.

2. Results

2.1. Overview of enzyme kinetic parameter fitting pipeline

The parameterization workflow is divided into six steps (Fig. 1). 1) Gather available enzyme kinetic data and associated experimental conditions. 2) Construct a mass action enzyme mechanism, consisting of known individual reaction steps involved in catalysis and regulation of the enzyme. 3) Based upon the enzyme mechanism, define kinetic

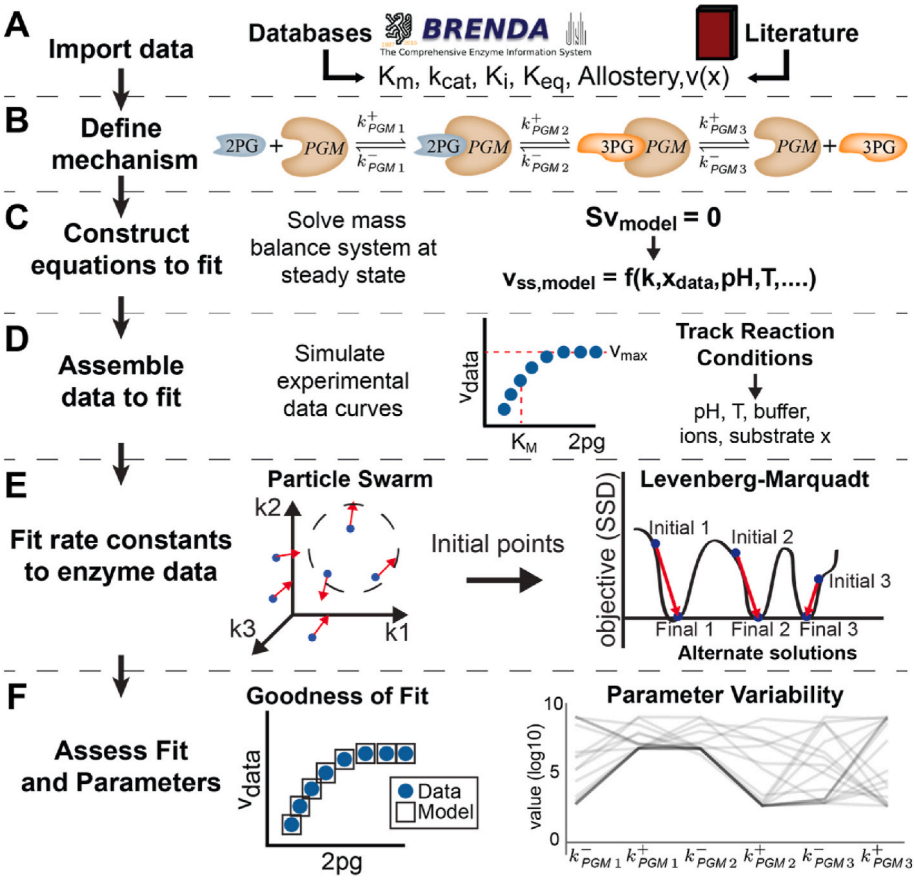


Fig. 1. Workflow for parameterization of rate constants using enzyme kinetic data. This workflow consists of six steps: A) Gathering and curating kinetic data, B) Defining an enzyme mechanism, C) Defining equations that relate model behavior to each data type, D) Processing kinetic data and correcting it to *in vivo*-like conditions, E) A nonlinear least-squares optimization to identify rate constant sets that fit available kinetic data, F) Assessing the fit for goodness of fit and parameter uncertainty, as well as calculating clusters of similar rate constant set results.

equations corresponding to the macroscopic kinetic data types (e.g. K_M, k_{cat}) to be fit. 4) Pre-process the kinetic data to allow direct comparison to the equations defined in Step 3.5) Fit the equations defined in step 3 to the processed data in Step 4 using a nonlinear least squares optimization approach to obtain a locally optimal set of microscopic rate constants that reproduces the experimental data. Data priorities specified by the user determine the weight of each data point in the fitting procedure. This optimization problem is solved N times with pseudo-random start points to account for the under-determined nature of problem. 6) Cluster the resulting rate constant sets to identify a reduced number of characteristic rate constant sets for the enzyme, each of which reproduces the experimentally-determined kinetic behavior of the enzyme. Finally, fit performance is summarized, effective macroscopic kinetic parameters are recalculated from the fitted microscopic rate constants, and the parameterized model can be exported as text files for modeling in desired kinetic simulation software. We discuss the details of each step below.

2.2. Step 1: Preparation of enzyme kinetic data

First, kinetic data on enzymes is gathered, curated, and placed into a table in a standard format (Fig. 1A). Data types currently handled include enzyme structure, reaction stoichiometry, kinetic reaction mechanism, reaction equilibrium constant (K_{eq}), dissociation constants (K_d), and standard initial rate kinetic assay constants such as the Michaelis-Menten constant, K_M, the turnover rate, k_{cat}, and inhibition constants, K_i. Additional data types, such as *in vivo* flux data, can be utilized as well, provided that a comparison equation is specified by the user in step 3. Experimental condition data such as co-substrate concentrations, pH and temperature are also extracted for use in data adjustments to *in vivo*-like conditions, as discussed later. Weights for each data point are determined by the user. This allows the user to decide how heavily to consider data points that may be conflicting with other measurements or have reliability concerns.

2.3. Step 2: Specify the enzyme reaction mechanism

A mechanism for the enzymatic reaction, consisting of individual reaction steps that describe the binding of the enzyme to substrates, catalysis, and product release, is then specified by the user (Fig. 1B). Each of these steps is modeled using mass action rate laws. Since the reaction steps do not proceed through a single transition state, these are not true elementary reactions. Instead, we refer to them using Cleland's nomenclature(Cleland 1963) as microscopic reaction steps, with microscopic kinetic rate constants, contrasted with macroscopic kinetic parameters such as Michaelis constants K_M.

The method is flexible to various reaction schemes, such as sequential versus random binding orders, ping pong mechanisms, and slow enzyme transitions. Both reversible and irreversible reactions can be specified, but fully reversible mechanisms are recommended for later use of thermodynamic Haldane relationships(Alberty 1953). Generally, protons and water are assumed constants and excluded from reaction mechanisms by convention. At this stage, individual catalytic tracks and thermodynamic cycles can be defined by the user to serve as thermodynamic constraints, such as Haldane constraints arising from the First Law of Thermodynamics. A catalytic track consists of a particular set of catalytic microscopic reaction steps that convert substrates into products. Haldane constraints, which consist of a multiplication of equilibrium constants along a particular catalytic track, are later fit to be equal to the overall reaction K_{eq}. In addition, small molecule enzyme inhibition and activation can be modeled by adding the corresponding microscopic binding reaction step(s). For enzyme competitive, uncompetitive, and mixed inhibition mechanisms, these microscopic reactions are added automatically if the inhibitor and respective affected metabolites are specified.

We note that although the mechanism is specified in terms of mass action reactions, an overall rate equation is derived based on a quasi-steady state assumption, and either the original mass action equations or this overall rate equation can be used for downstream kinetic modeling, as the user desires.

2.4. Step 3: Constructing symbolic comparison equations

We then set up the equations that are used to fit the kinetic data (Fig. 1C). These equations are functions of microscopic rate constants and in some cases of metabolite concentrations as well. For K_{eq} values, Haldane relationships are defined in terms of individual rate constants for each catalytic track through the enzyme mechanism. For k_{cat}, K_M, and K_i values, equations are derived from the overall steady-state rate equation of the reaction, v_{ss}, as described in the methods. The equation for the steady-state flux, v_{ss}, is found by solving the system of mass balance equations at steady-state along with a total enzyme sum equation representing total enzyme conservation (see **Methods** for details and **Supplementary Information** for a case study). Other data types such as dissociation constants can be used as well, if the user specifies a corresponding equation relating the data type to the enzyme microscopic rate constants.

2.5. Step 4a: Preparing kinetic data for fitting

The macroscopic kinetic data, such as K_{eq}, k_{cat}, K_M, and K_i, are then processed into a form that enables direct comparison to model behavior (Fig. 1D). For reaction K_{eq} values, the processed form is simply the K_{eq} value calculated under experimental conditions (pH, IS, and T). For k_{cat}, the processed form is the value along with the measured experimental conditions, including substrate concentrations. If substrate concentrations are not available, a concentration that is likely to be saturating is assumed (i.e. 1 M), in order to represent the excess concentrations typically used in the measurement of turnover rates. In these cases, it is recommended that K_M values are specified as well to ensure this constraint is satisfied by the final parameter set. For K_M and K_i, an initial rate curve is generated using the classical Michaelis-Menten equation with the K_M and K_i value substituted when applicable. This curve is simulated at substrate concentrations an order of magnitude above and below the measured K_M value. Other experimental conditions such as co-substrate concentration and media conditions are reported when available. Co-substrate concentration values in particular are substituted in the equations used to fit the K_M and K_i data. This procedure attempts to simulate the original experiment. However, in lieu of this experimental plot simulation procedure, raw data could be used in principle as well.

To deal with uncertainty in macroscopic kinetic data, upper and lower bounds can be specified and data sampled from a normal distribution with a given mean and standard deviation. In the case of unknown macroscopic kinetic data, a uniform distribution spanning a characteristic range for the data type may be sampled. For example, K_m values tend to fall in the range of 10^{-7} to 10^{-2} M, while k_{cat} values may fall in the range of 10 to 10^5 s^{-1}(Bar-Even et al., 2011). When sampling is performed, a number of samples is defined, and a data set is created for each sample point.

2.6. Step 4b: Correcting data for in vitro to in vivo differences

To correct these values for *in vitro* to *in vivo* differences, several adjustments to the data are implemented. First, k_{cat} can be adjusted to *in vivo* temperature from *in vitro* conditions using a user defined Q10 value for the enzyme, which is specified by the user but has a default value of 2.5, typical for metabolic enzymes(Hochachka 1991; Yurkovich et al., 2017). Effects of temperature on K_m are thought to be less substantial and thus are currently ignored(Scopes 1995). Further, for each data type, concentrations can be corrected to the more accurate chemical activities using a Debye-Huckel model for the activity coefficient given the ionic strength under experimental conditions, though currently this

must be performed by the user(de Jiménez et al., 1964). Reaction equilibrium constants can be calculated at a specific pH and ionic strength as well using published tools(Du et al., 2018). Finally, inhibitory effects of pH changes on enzyme behavior can be modeled by adding proton binding and dissociation reactions to the enzyme mechanism with associated dissociation constants, as has been done in the enzymology literature(Tipton and Dixon 1979; Moxley et al. 2014).

2.7. Step 5: Two-stage randomized fitting of microscopic rate constants to macroscopic kinetic data

Once both the data and equations have been prepared, these are passed to a two-stage nonlinear least squares optimization procedure (Fig. 1E). The target values are given by the data, while the model values are given by the equations with the experimental conditions substituted into the equations, leaving them functions of the microscopic rate constants alone. The fitting procedure then yields microscopic rate constant values that cause the enzyme model to reproduce the measured data. As an objective, we use the absolute difference between the logarithm of the model-predicted value and the logarithm of the data, multiplied by the user-defined weights on each data value. We use bounds on the microscopic rate constants of 10^{-6} and 10^{9} s^{-1} based on typically assumed limits of diffusion and an arbitrarily slow lower bound(Alberty and Hammes 1958). These rate constant bounds may not be relevant for reactions not involving association or dissociation. However, we have not found these bounds to affect the fitting error for the cases we have examined thus far, and the lower bound in particular is rarely hit in practice. Additionally, to address the broad scale of potential rate constants, we log transform the rate constants during this optimization. Furthermore, weights can be assigned to each data point for balancing the contribution of each data point to the fitting objective function or to reflect confidence in particular measurements.

However, the equations are highly nonlinear, causing many optimization algorithms to fail to converge to an acceptable fit. To address this challenge, we first run a non-derivative based particle swarm optimization (PSO) to find rate constant sets that fit the data well enough to serve as initial points for a more precise derivative-based optimization. The second optimization is a Levenberg-Marquadt (LMA) derivative-based optimization with the same log-parameter scaling and objective. Resulting fits are examined for total residual, and points that fit satisfactorily are kept. These rate constants are then once more substituted into the equations for the kinetic data to verify that this kinetic data used to fit the model can be reproduced.

Importantly, as PSO is a randomized algorithm, initial points passed to LMA are different each time that the algorithm is run. For typical underdetermined systems, different rate constant sets are returned by the optimization every time, each of which fits the kinetic data equivalently well. This effectively samples the rate constant space in the prevalent cases where rate constants are not uniquely specified by the available kinetic data. Thus, this procedure inherently addresses the underdetermined system issue in parameterization of enzyme reaction mechanisms.

2.8. Step 6: Cluster parameters to extract characteristic rate constant sets

As initial points to the LMA optimization are defined pseudo-randomly based on the PSO optimization results, the resulting rate constants are generally different in each optimization. The fitting procedure is then repeated a number of times until qualitatively new alternative rate constant sets are no longer found (Fig. 1F). To determine convergence, clustering is used to identify whether new clusters have been formed by addition of new fitted points, or whether these points fall within existing clusters. Then, once convergence has been reached, characteristic rate constants sets for the enzyme are selected based on proximity to the centroids of these clusters. These centroids then serve as candidate rate constant sets for the enzyme, each of which reproduces measured kinetic data for the enzyme equivalently well.

After the parameters are fit, the best fit parameter set is evaluated. The macroscopic parameters corresponding to the fit microscopic rate constants are recalculated and compared to the data used in training to assess final error on each parameter. Centroid cluster parameter sets can be exported to text files for import into desired kinetic modeling packages, such as the MASSpy kinetic modeling package in Python(Haiman et al., 2021).

2.9. Case studies demonstrating the ability of MASSef to fit specific data types

Having described the workflow for parameterizing microscopic mass action reaction mechanisms using enzyme kinetic data, we now present a series of case studies demonstrating the workflow in action when different types of data are available.

2.10. Case study 1: demonstrating the fitting procedure on k_{cat}, K_m, and k_{eq} data for glyceraldehyde-3-phosphate dehydrogenase

We first demonstrate the basic capabilities of the fitting procedure, using the case of the glycolytic reaction Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) from *E. coli* (Fig. 2). This reaction catalyzes the substrate level phosphorylation of glyceraldehyde-3-phosphate (g3p) to produce 1,3-diphosphoglycerate (13dpg) and NADH. The enzyme has reported K_m values of 0.89 mM, 0.045 mM, and 0.53 mM for glycerol-3-phosphate, NAD, and inorganic phosphate (P_i), respectively (Eyschen et al., 1999) (Fig. 2B). Also reported in the study was a forward k_{cat} of 268 s^{-1} at 295K and a dissociation constant K_d for NAD of 0.00032 mM. A K_{eq} value of 0.452 was extract from eQuilibrator at a pH of 7 and ionic strength of 0.25M. While the substrate binding order in *E. coli* has not been experimentally determined, it has been reported in humans to be an ordered Bi Bi mechanism with a binding order NAD, then g3p, then P_i, and a release order of 13dpg followed by NADH(Wang and Alaupovic 1980). This mechanism was constructed and fit to the data using the MASSef workflow, and rate constants enabling the model to perfectly fit the data were found (Fig. 2C). Rate constants were not uniquely identified, but instead were constrained to certain ranges (Fig. 2D). Due to the presence of more data on the forward direction than the reverse, the microscopic rate constants associated to the substrate binding steps are constrained to a greater degree. Examination of good fitting rate constant sets showed nonlinear dependencies between the rate constants, a reflection of the alternate possible rate constant sets that can equivalently satisfy the constraints on enzyme function places by measured turnover rates and Michaelis-Menten constants (Fig. 2E). Clustering of the identified rate constants then extracts sets of characteristic rate constants with particular nonlinear dependencies (Fig. 2F).

2.11. Case study 2: reconciling inconsistent data for Phosphoglycerate Mutase

We then demonstrate the ability of MASSef to reconcile data that is inherently inconsistent, using Phosphoglycerate Mutase (PGM) as a case study. This enzyme catalyzes the conversion of 2-phosphoglycerate to 3-phosphoglycerate and has well specified kinetics, with both forward and reverse k_{cat} and K_M defined by data (Fig. 3A and B). The k_{cat} data was corrected to 37°C using a Q_{10} of 2.5. The equilibrium constant for the reaction was obtained from the eQuilibrator web server(Flamholz et al., 2012) for pH of 7.5 and IS of 0.25M. The fitting procedure reveals that the kinetic data and K_{eq} are not kinetically consistent (Fig. 3C). The reason for this inconsistency is apparent once the Haldane relationship is calculated using the kinetic data and compared to the reaction equilibrium constant. The K_{eq} for the reaction is 5.3, while the K_{eq} calculated from kinetic parameters is only 1.4, indicating significant inconsistency. This type of inconsistency suggests that some data may not be trustworthy. While distinguishing good data from bad data is the task of the

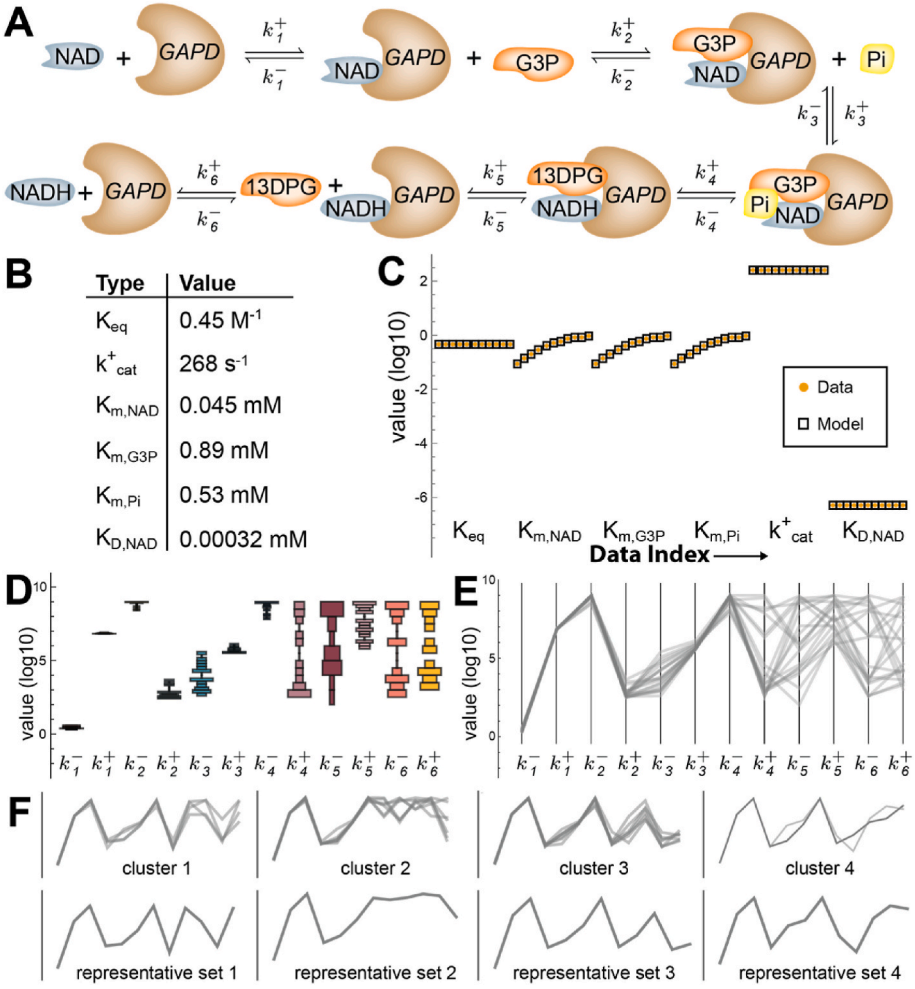


Fig. 2. Demonstrating the fitting procedure for Glyceraldehyde- 3-phosphate Dehydrogenase (GAPDH) from *E. coli*. A) Enzyme mechanism for GAPDH. B) Available kinetic data for GAPDH. C) Best fit results for GAPDH and resulting kinetic parameters calculated using symbolic equations for these parameters. Target data is shown as circles, while the model results are shown as boxes for clarity. Data is ordered by index as it is generated from the data simulating equations by type of data (e.g. K_{eq}, K_m, k_{cat}). Data is replicated for certain data types (K_{eq}, k_{cat}, K_D) to balance the contribution of error on these data points to the objective function with that of K_M data. For this enzyme, the fit between model and data was tight, as no discrepancies are apparent. D) Distribution plot of rate constant sets obtained from 100 optimizations showing parameter ranges E) Rate constant sets showing connections between rate constants F) Clustering of rate constant sets reveals nonlinear dependencies between rate constants. The four identified clusters and the rate constant sets that are nearest to the cluster means are shown.

modeler, the workflow results in rate constants that attempt to reconcile the inconsistent data as well as possible. The resulting kinetic constants are not far from the measured values. Furthermore, different weight values can be used to weigh particular data points more heavily based on user confidence. We demonstrate several possible data point weights for this enzyme, based on individually deprioritizing K_ms, k_{cat}s, or K_{eq} data (Fig. 3C,D,E).

2.12. *Case study 3: incorporating in vivo flux data for Triose Phosphate Isomerase*

A common issue with the use of *in vitro* data is an incompatibility with the *in vivo* requirements of pathway flux. For example, a k_{cat} may be measured inaccurately or in a non-physiological condition, resulting in less activity than required to sustain a physiological flux given a measured protein concentration. Here, we demonstrate how MASSef can incorporate *in vivo* data to obtain a set of microscopic kinetic parameters that balance *in vitro* measurements with *in vivo* requirements. We show this for the glycolytic enzyme triose-phosphate isomerase, TPI, in *E. coli*, which converts dihydroxyacetone phosphate (DHAP) to glyceraldehyde-3-phosphate (G3P) (Fig. 4A). The equilibrium of this reaction is slightly favored toward DHAP with a Keq of 0.11 under standard conditions, and extremely fast kinetics in the direction toward DHAP production with a catalytic efficiency near the diffusion limit, while an experimental K_M for DHAP was not available in the literature (Fig. 4B). We gathered *in vivo* data on fluxes estimated from flux balance analysis based on growth rate and metabolic exchanges during aerobic growth on glucose and acetate(Gerosa et al., 2015), metabolite concentration data(Gerosa et al., 2015), and proteomics data(Schmidt et al., 2016) (Fig. 4C). We first fit the *in vitro* data alone using MASSef and found that the model was able to perfectly capture the data (Fig. 4D). We then added the *in vivo* data as additional target values, by asking the model to fit the *in vivo* flux given the measured metabolite and enzyme concentrations (Fig. 4E). The model was unable to perfectly fit both *in vitro* and *in vivo* data, indicating an inconsistency. As lower glycolysis reactions are known to be near equilibrium, which affects the thermodynamic efficiency of the enzyme(Noor et al., 2013), we hypothesized that changing the metabolite concentrations may enable a consistent fit. Indeed, by adjusting the metabolite concentrations only slightly to affect the distance of the reaction from equilibrium, by changing G3P from 0.08 mM to 0.1 mM on the glucose condition, we were able to find rate constants that match both *in vitro* and *in vivo* behavior (Fig. 4F). We recalculated the effective macroscopic parameters and found that the unadjusted *in vivo* data resulted in less efficient kinetic parameters than predicted by *in*

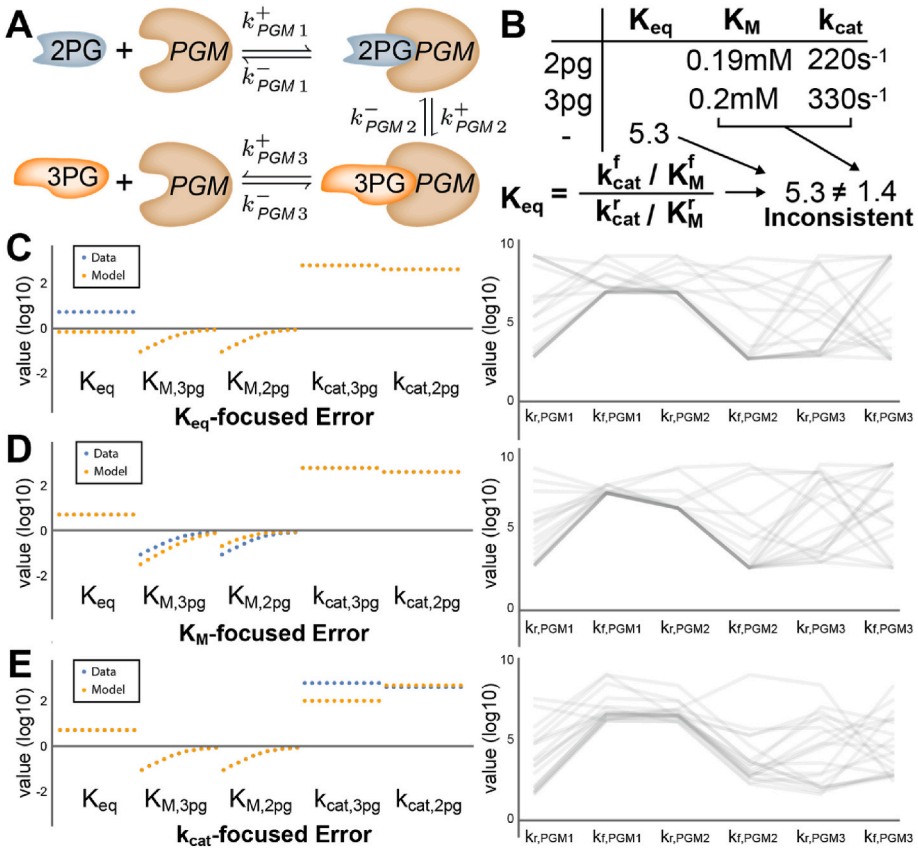


Fig. 3. Fitting data for Phosphoglycerate Mutase (PGM) from *E. coli* shows customizable handling of inconsistent data. A) Enzyme mechanism for PGM. B) Available kinetic data for PGM. C) Best fit results for PGM and resulting kinetic parameters calculated using symbolic equations for these parameters. Error in the K_{eq} values are due to data conflicts between K_m, k_{cat}, and K_{eq} due to the Haldane constraint that relates these parameters. On the right, distribution plot of rate constant sets obtained from 100 optimizations showing parameter ranges D) Best fit results for PGM and resulting kinetic parameters calculated with data priorities scores adjusted to de-prioritize K_m data, which has the effect of placing the fitting error on these values. On the right, resulting rate constants show small changes as a result of the adjusted fit. E) Best fit results for PGM and resulting kinetic parameters calculated with data priorities scores adjusted to de-prioritize k_{cat} data, which has the effect of placing the fitting error on these values. On the right, resulting rate constants show small changes as a result of the adjusted fit.

vitro experiments, while the adjusted *in vivo* data was consistent with the *in vitro* experiments (Fig. 4G). The adjusted concentrations were well within apparent experimental error, especially given that condition-specific concentrations for DHAP were not available. Thus, we interpret this result to indicate that neither *in vitro* nor *in vivo* data are inherently problematic or erroneous, but rather the system is inherently highly sensitive to small changes in all parameters due to the reaction being near equilibrium where sensitivity to metabolite concentration differences is high. This work highlights the need to systematically reconcile *in vitro* and *in vivo* data to obtain consistent kinetic models, and specifically suggests the sensitivity of reactions near equilibrium to fine adjustments in metabolite concentrations as a source of potential discrepancies.

2.13. Constructing large-scale models with parameterized enzyme modules

Once a desired set of enzyme modules have been parameterized, these can be integrated to form pathway or network scale kinetic models of metabolism(Haiman et al., 2021; Du et al., 2016). We provide a case study integrating enzyme modules within a mass action model of glycolysis as a Jupyter notebook workflow implemented in Python (scripts available on github.com/opencobra/massef). Code is provided to export text files of enzyme modules from MASSef in a format that can be directly imported using the MASSpy Python package(Haiman et al., 2021). In this workflow, we demonstrate that a steady state solution can be obtained through simulation of the system before and after enzyme modules are introduced, supporting the viability of the complete bottom-up workflow for model construction.

3. Discussion

In this work, we present a workflow and software tools for the parameterization of detailed kinetic models of enzyme reactions using enzyme kinetic data. This software tool has a number of powerful features. First, the workflow allows flexible user-defined reaction mechanisms and can handle the majority of common reaction schemes, including different binding orders, reaction mechanisms such as ping-pong mechanisms, inhibition schemes, protonation reactions, and allostery. Second, the workflow enables the fitting of standard kinetic data types, including thermodynamic and initial rate data, and can correct these data, to a degree, to *in vivo*-like conditions. Third, the optimization procedure can handle the inherently highly nonlinear least-squares optimization and sample rate constant sets to deal with underdetermined systems.

The modeling procedure here depends on user-specified reaction mechanisms. These manual curation efforts can be seen as analogous to metabolic reconstructions in the constraint-based modeling framework (Monk et al., 2017). Reconstructions require an up-front time investment from model curators to assess literature information and incorporate known enzyme kinetic mechanisms. However, once established these kinetic mechanisms represent a community resource that can be iteratively improved and applied for diverse applications. Currently, kinetic modules can be built for tens of enzymes per day, primarily gated

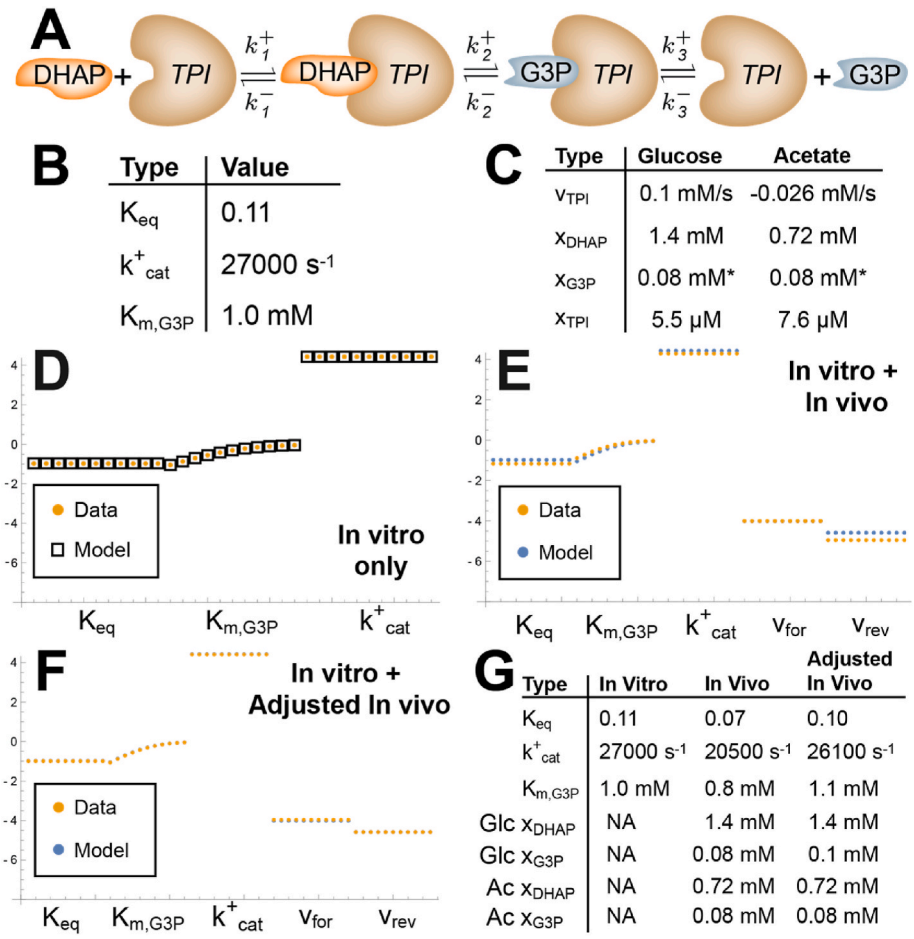


Fig. 4. Fitting in vivo data for Triose Phosphate Isomerase (TPI) from E. coli enforces consistency of *in vitro* data with physiological function. A) Enzyme mechanism for TPI. B) Thermodynamic and In vitro kinetic data for TPI. C) *In vivo* data for TPI, including flux, metabolite concentration, and enzyme concentration data. *indicates manually adjusted value to ensure thermodynamic feasibility. D) Data fit using only thermodynamic and *in vitro* data. E) Data fit using thermodynamic, *in vitro*, and *in vivo* data, with *in vivo* data heavily weighted to force a tight fit. F) Data fit using thermodynamic, *in vitro*, and adjusted *in vivo* data. G) Recalculated kinetic parameters for each fit, as well as *in vivo* metabolite concentrations used in the fit.

by determining the enzyme mechanism and reconciling any discrepancies between model behavior and kinetic data. Further workflow scaling is likely possible through semi-automated application of established mechanisms and quality control assessment. This approach to bottom up reconstruction of enzyme kinetics contrasts with other recent approaches(Khodayari and Maranas 2016; P. Saa and Nielsen 2015; Miskovic and Hatzimanikatis 2010; Choudhury et al., 2022) that utilize somewhat approximated Michaelis-Menten or other kinetic rate laws that are simultaneously parameterized to match systems-level data such as measured steady state fluxes. One strength of the bottom up approach is inconsistencies between model behavior and systems-level data, due for example to factors such as unknown regulators, can be assessed on an enzyme-by-enzyme behavior, simplifying diagnosis of issues. Meanwhile, a strength of systems level fitting approaches is the ability to take into account system level model behavior during parameterization. Currently, although multiple *in vivo* steady state flux values can be fit in the bottom up workflow as demonstrated for TPI, the stability and dynamic behavior of resulting network models is not guaranteed when assembling large-scale models in a bottom up fashion, requiring further workflow extensions.

The reaction schemes that we have tested thus far are those based on standards set by Cleland and other enzymologists(Cleland 1963). These reaction mechanisms have certain assumptions but have seen practical success in representing enzyme behavior in initial rate and progress curve experiments(Teusink et al., 2000). Beyond the relatively simple kinetic mechanisms demonstrated in this work, there are additional levels of detail that potentially could be represented that we have not yet tested. These include additional breakdown of catalysis into individual interactions between the substrate and catalytic residues as well as breakdown of common deterministic and well-mixed assumptions due to restricted geometries(Ma and Nussinov 2013), low copy numbers (Tzafriri 2003), channeling(Spivey and Ovádi 1999), stochastic behavior(Sanft et al. 2011), proton tunneling (Klinman and Amnon, 2013), and other complex kinetic phenomena. These more complex situations have been handled by others in various ways but not yet integrated into our workflow.

The current workflow has been developed to handle reaction K_{eq} values as well as enzyme initial rate data such as k_{cat}, K_M, and K_i, due to the dominance of these data types in the enzyme kinetic literature (Chang et al., 2021). However, in principle this workflow could be extended to additional data types such as progress curve data(Choi et al. 2017; Eicher et al. 2012) and stop-flow data(Hartwell and Grudpan 2012). The requisite for utilizing these data types is the construction of equations to be used in the least-squares optimization. For example, to compare to progress rate data, the mass balance equations for the enzyme would be integrated over time and compared to the experimental time course data. Additionally, in the current procedure we use simulated experimental curves that are effectively back-calculated for the parameters in the case of K_M and K_i. This procedure was implemented because the original data curves are often not available or are inconvenient to extract from the literature. However, in principle the original data plots could be fit directly and should have mostly

equivalent results to the current workflow.

Microscopic mass action enzyme systems typically have many more rate constants than kinetic data, and thus are highly underdetermined. One of the most powerful aspects of this workflow is the ability to fit rate constants for these underdetermined systems, due to the inherent parameter sampling built into the optimization procedure. This capability is enabled by the pseudo-random start points for the LMA optimization that are provided by the initial non-derivative based PSO optimization. One potential issue is to determine how completely the parameter space is being sampled. We attempt to address this by clustering the rate constants through successive optimizations, and considering the parameter space to be fully specified when new clusters are no longer found. However, this procedure cannot discount inherent bias in the optimization procedure that would lead parameters with particular biases to be oversampled. Thus, more powerful sampling methods, such as Monte Carlo sampling, would still be desirable to be implemented.

4. Conclusions

Kinetic modeling of metabolic networks has seen a resurgence in recent years, driven largely by parameterization strategies involving sampling and system-level parameter fitting. The work here is intended to increase the accessibility of a 'bottom-up' parameterization approach that makes use of the plethora of historical enzyme kinetic data available in the literature. As these models are parameterized enzyme-by-enzyme, this parameterization approach is expected to scale well to the network scale. Additionally, there have been recent efforts to fill gaps in critical parameters such as turnover rates at the genome-scale that should further enable these efforts(Davidi et al., 2016; Heckmann et al., 2018). As parameterization becomes increasingly computationally feasible and biochemically accurate, practical kinetic models will likely soon become accessible and powerful tools for the systems biology community to study metabolism.

5. Materials and methods

5.1. Setting up comparison equations

To construct comparison equations for each data type, the overall steady-state equation of the system is first solved by solving the system of equations consisting of mass balances on each species along with an expression of the total sum of enzyme:

$$\frac{dx}{dt} = Sv(k,x)$$

$$\sum E_i = E_{tot}$$

Where **S** is the reaction stoichiometry. This yields an expression for the steady state flux v_{ss} that is a function of the microscopic rate constants **k**, the metabolite concentrations **x**, and the total enzyme $\mathbf{E_{tot}}$. This expression can be used to derive comparison equations for specific data types as described below.

k_{cat}: The enzyme turnover rate $\mathbf{k_{cat}}$ is a proportionality constant between the maximum rate of the enzyme $\mathbf{v_{max}}$ and the total enzyme concentration $\mathbf{E_{tot}}$. To obtain an expression that approximates $\mathbf{v_{max}}$, we insert concentrations of the metabolite that are assumed to be saturating, creating an approximating function $\mathbf{v_{sat}}$. A standard value of 1 M is used for this saturating concentration as a default, but the user can specify this value as needed. Then, the comparison equation is found by:

$$k_{cat} = \frac{v_{sat}}{E_{tot}}$$

In this expression, the value of $\mathbf{E_{tot}}$ does not need to be specified, as $\mathbf{v_{sat}}$ is a function of $\mathbf{E_{tot}}$ as well and these terms cancel from the equation. Product and inhibitor concentrations are set to zero in these expressions, unless otherwise defined, which matches the standard conditions for initial rate experiments.

K_m: The Michaelis constant $\mathbf{K_m}$ for a metabolite is defined as the concentration of that metabolite at which the enzyme rate is half of its maximal value.

$$v_{ss}\ (x = K_m) = \frac{1}{2} v_{max}$$

While this equation can be directly solved to generate an expression for K_m, this can yield multiple solutions for certain reaction mechanisms. As a simplifying approach, we instead define a relative rate $\mathbf{v_{rel}}$

$$v_{rel} = \frac{v_{ss}}{v_{sat}}$$

Where v_{sat} is the overall rate law with an assumed saturating concentration of the metabolites substituted in. The validity of this saturating concentration must be reassessed at the end by the user to ensure that the effective K_M of the enzyme is sufficiently below the saturating concentration to ensure effective saturation. In generating this expression, the concentration of any co-substrates is set to their experimental values. This expression can then be directly fit to a simulated Michaelis-Menten curve generated with the experimental $\mathbf{K_m}$ value.

$$v_{rel} = \frac{x}{K_m + x}$$

K_i: Inhibition constants are fit analogously to $\mathbf{K_m}$ values, but additional terms are present in the simulated Michaelis-Menten curve, dependent upon the types of inhibition present. Competitive (K_{ic}), uncompetitive (K_{iu}), and noncompetitive schemes are currently supported.

$$\text{Competitive } v_{rel} = \frac{x}{K_m\left(1 + \frac{I}{K_{ic}}\right) + x}$$

$$\text{Uncompetitive } v_{rel} = \frac{x}{K_m + x\left(1 + \frac{I}{K_{iu}}\right)}$$

$$\text{Noncompetitive } v_{rel} = \frac{x}{K_m\left(1 + \frac{I}{K_{ic}}\right) + x\left(1 + \frac{I}{K_{iu}}\right)}$$

K_{eq}: Reaction equilibrium constant data is fit by generating a Haldane relationship for each catalytic path through the enzyme.

$$K_{eq} = \frac{\prod k^+}{\prod k^-}$$

These catalytic paths are currently manually specified by the user.

5.2. Optimization

To find microscopic rate constants that measured macroscopic enzyme kinetic data, we utilized a two-stage non-linear regression approach. Rate constants were constrained to be between 10^{-6} s^{-1} and 10^{9} s^{-1}, where the upper bound is set based on the diffusion limit and the lower bound is an arbitrarily slow reaction rate. In practice, we did not observe rate constants within several orders of magnitude of the lower bound in any of the rate constant set solutions. Rate constant variables were transformed to logarithmic variables to help span the large range in possible rate constants.

Each optimization step minimizes the sum of squared residuals between the comparison equations and the data values. Priorities on each data point are specified by the user and used as a weighting on errors on those data points in the objective function. As certain data types, such as $\mathbf{K_m}$ values, consist of curves rather than single points and thus may artificially weigh more heavily in the optimization, the data points are balanced to contribute equally to the objective.

In the first optimization step, a non-derivative-based particle swarm

optimization was run. Parameters for this optimization that were found to be successful across diverse data values and enzyme mechanisms are included as defaults in the fitting code, but can be changed by the user. The resulting rate constants from this initial optimization problem were then used as initial values to a second optimization problem. The second optimization was a derivative-based Levenberg-Marquadt method that refines the rate constant sets to a low sum of squared deviations to the measured data. Once again, robust options for this algorithm, such as tolerances, were identified based on performance across multiple enzyme mechanisms and are included as default options.

This optimization procedure did not always return a good fit, due to the pseudo-randomness of initial points and the local optimality properties of the derivative-based optimization. Repeated runs of the fit were executed until a user-controlled number of good fits were found. Additional termination criteria around convergence of rate constants into a consistent clustering set were also tested.

The optimization procedure is implemented in Python.

5.3. Clustering

Once a series of rate constant sets was calculated, these rate constant sets were clustered using the Mathematica FindClusters function, with the "Optimize" method, and 2 iterations. The median rate constant set for each cluster was then calculated. The rate constant set closest to the median value was selected as the characteristic rate constant set for that cluster.

5.4. Software and requirements

The MASSef package is available at https://github.com/opencobra/MASSef. The MASS Toolbox is available at http://opencobra.github.io/MASS-Toolbox/. The following Python packages and versions are required for the optimization:

numpy 1.12.0
scipy 0.18.1
ecspy 1.1 (deprecated, package included internally to MASSef)
lmfit 0.9.5.
Python 3.7+
Mathematica 10+

Computational run time varies by machine, enzyme complexity, and extent of exploring alternate parameter sets. Tested on a machine with a Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz processor with 32.0 GB of memory, the GAPD notebook from Fig. 2 took under 10 min to run 10 fits (parameter sets computed with different randomized initializations) on a single core. The number of cores used directly improves run time as the bottleneck step is the parameter optimization, which can be split across cores. For more complex enzymes, such as allosteric enzymes with several effectors, the algebraic solution of the steady-state equations can also take a substantial amount of time (on the order of an hour), but results can be saved so that they only need to be run once per enzyme. Complex enzymes also take more time to fit, scaling with the number of parameters, but in our experience individual fits have not exceeded ~30 min per fit per core.

CRediT authorship contribution statement

Daniel C. Zielinski: Writing – original draft, Supervision, Software, Methodology, Formal analysis, Conceptualization. **Marta R.A. Matos:** Writing – original draft, Software, Methodology, Formal analysis. **James E. de Bree:** Writing – original draft, Software, Methodology. **Kevin Glass:** Writing – original draft, Software, Methodology. **Nikolaus Sonnenschein:** Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Bernhard O. Palsson:** Writing – original draft, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Novo Nordisk Foundation [grant numbers NNF10CC1016517, NNF20CC0035580]. We would like to thank Zachary Haiman and Zafrin Dhali for useful discussions and testing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2024.e00234.

References

Adams, Richard, Clark, Allan, Yamaguchi, Azusa, Hanlon, Neil, Tsorman, Nikos, Ali, Shakir, Lebedeva, Galina, et al., 2013. SBSI: an extensible distributed software Infrastructure for parameter estimation in systems biology. Bioinformatics 29 (5), 664–665.

Alberty, Robert A., 1953. The relationship between Michaelis constants, maximum Velocities and the equilibrium constant for an enzyme-Catalyzed reaction. J. Am. Chem. Soc. 75 (8), 1928–1932.

Alberty, Robert A., Hammes, Gordon G., 1958. Application of the theory of diffusion-controlled reactions to enzyme kinetics. The Journal of Physical Chemistry 62 (2), 154–159.

Andreozzi, Stefano, Chakrabarti, Anirikh, Cher Soh, Keng, Burgard, Anthony, Yang, Tae Hoon, Van Dien, Stephen, Miskovic, Ljubisa, Hatzimanikatis, Vassily, 2016. Identification of metabolic Engineering targets for the Enhancement of 1,4-Butanediol production in Recombinant E. Coli using large-scale kinetic models. Metab. Eng. 35 (May), 148–159.

Bar-Even, Arren, Noor, Elad, Savir, Yonatan, Liebermeister, Wolfram, Davidi, Dan, Tawfik, Dan S., Milo, Ron, 2011. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50 (21), 4402–4410.

Bennett, Bryson D., Kimball, Elizabeth H., Gao, Melissa, Osterhout, Robin, Van Dien, Stephen J., Rabinowitz, Joshua D., 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat. Chem. Biol. 5 (8), 593–599.

Chang, Antje, Jeske, Lisa, Ulbrich, Sandra, Hofmann, Julia, Koblitz, Julia, Schomburg, Ida, Neumann-Schaal, Meina, Jahn, Dieter, Schomburg, Dietmar, 2021. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. Nucleic Acids Res. 49 (D1), D498–D508.

Chassagnole, Christophe, Noisommit-Rizzi, Naruemol, Schmid, Joachim W., Mauch, Klaus, Reuss, Matthias, 2002. Dynamic modeling of the central carbon metabolism of Escherichia coli. Biotechnol. Bioeng. 79 (1), 53–73.

Choi, Boseung, Rempala, Grzegorz A., Kim, Jae Kyoung, 2017. Beyond the michaelis-menten equation: accurate and efficient estimation of enzyme kinetic parameters. Sci. Rep. 7 (1), 1–11.

Choudhury, Subham, Moret, Michael, Salvy, Pierre, Weilandt, Daniel, Hatzimanikatis, Vassily, Miskovic, Ljubisa, 2022. Reconstructing kinetic models for dynamical studies of metabolism using generative adversarial networks. Nat. Mach. Intell. 4 (8), 710–719.

Chowdhury, Anupam, Ali, Khodayari, Maranas, Costas D., 2015. Improving prediction fidelity of cellular metabolism with kinetic descriptions. Curr. Opin. Biotechnol. 36 (December), 57–64.

Cleland, W.W., 1963. The kinetics of enzyme-catalyzed reactions with two or more substrates or products: I. Nomenclature and rate equations. Biochim. Biophys. Acta (BBA) - Spec. Sect. Enzymol. Subj. 67 (Suppl. C), 104–137.

Costa, Rafael S., Veríssimo, André, Vinga, Susana, 2014. KiMoSys: a web-based repository of experimental data for KInetic MOdels of biological SYStems. BMC Syst. Biol. 8 (August), 85.

Davidi, Dan, Noor, Elad, Liebermeister, Wolfram, Bar-Even, Arren, Flamholz, Avi, Tummler, Katja, Barenholz, Uri, Goldenfeld, Miki, Shlomi, Tomer, Milo, Ron, 2016. Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro kcat measurements. Proc. Natl. Acad. Sci. USA 113 (12), 3401–3406.

Du, Bin, Zhang, Zhen, Grubner, Sharon, Yurkovich, James T., Palsson, Bernhard O., Zielinski, Daniel C., 2018. Temperature-dependent estimation of gibbs Energies using an Updated group-contribution method. Biophys. J. 114 (11), 2691–2702.

Du, Bin, Zielinski, Daniel C., Kavvas, Erol S., Dräger, Andreas, Tan, Justin, Zhang, Zhen, Ruggiero, Kayla E., Arzumanyan, Garri A., Palsson, Bernhard O., 2016. Evaluation of

rate law approximations in bottom-up kinetic models of metabolism. BMC Syst. Biol. 10 (1), 40.
Eicher, Johann J., Snoep, Jacky L., Rohwer, Johann M., 2012. Determining enzyme kinetics for systems biology with nuclear magnetic resonance spectroscopy. Metabolites 2 (4), 818–843.
Eyschen, J., Vitoux, B., Marraud, M., Cung, M.T., Branlant, G., 1999. Engineered glycolytic glyceraldehyde-3-phosphate Dehydrogenase Binds the Anti Conformation of NAD+ Nicotinamide but does not experience A-specific Hydride transfer. Arch. Biochem. Biophys. 364 (2), 219–227.
Flamholz, Avi, Noor, Elad, Bar-Even, Arren, Milo, Ron, 2012. eQuilibrator–the biochemical thermodynamics calculator. Nucleic Acids Res. 40 (Database issue), D770–D775.
Foster, Charles J., Gopalakrishnan, Saratram, Antoniewicz, Maciek R., Maranas, Costas D., 2019. From Escherichia coli Mutant 13C Labeling data to a core kinetic model: a kinetic model parameterization Pipeline. PLoS Comput. Biol. 15 (9), e1007319.
Gábor, Attila, Banga, Julio R., 2015. Robust and efficient parameter estimation in dynamic models of Biological systems. BMC Syst. Biol. 9 (October), 74.
Gábor, Attila, Villaverde, Alejandro F., Banga, Julio R., 2017. Parameter Identifiability analysis and Visualization in large-scale kinetic models of Biosystems. BMC Syst. Biol. 11 (1), 54.
Gerosa, Luca, Bart, R.B., van Rijsewijk, Haverkorn, Christodoulou, Dimitris, Kochanowski, Karl, Schmidt, Thomas S.B., Noor, Elad, Sauer, Uwe, 2015. Pseudo-transition analysis Identifies the Key regulators of dynamic metabolic Adaptations from steady-state data. Cell Systems 1 (4), 270–282.
Haiman, Zachary B., Zielinski, Daniel C., Koike, Yuko, Yurkovich, James T., Palsson, Bernhard O., 2021. MASSpy: Building, simulating, and Visualizing dynamic Biological models in Python using mass action kinetics. PLoS Comput. Biol. 17 (1), e1008208.
Hartwell, Supaporn Kradtap, Grudpan, Kate, 2012. Flow-based systems for Rapid and high-Precision enzyme kinetics studies. Journal of Analytical Methods in Chemistry 2012 (April), 450716.
Heckmann, David, Lloyd, Colton J., Mih, Nathan, Ha, Yuanchi, Zielinski, Daniel C., Haiman, Zachary B., Amer Desouki, Abdelmoneim, Lercher, Martin J., Palsson, Bernhard O., 2018. Machine learning applied to enzyme turnover numbers reveals protein structural Correlates and improves metabolic models. Nat. Commun. 9 (1), 5252.
Heijnen, Joseph J., Verheijen, Peter J.T., 2013. Parameter Identification of in vivo kinetic models: Limitations and challenges. Biotechnol. J. 8 (7), 768–775.
Hochachka, P.W., 1991. Chapter 12 - temperature: the Ectothermy option. In: Hochachka, P.W., Mommsen, T.P. (Eds.), Biochemistry and Molecular Biology of Fishes, vol. 1. Elsevier, pp. 313–322.
Hoops, Stefan, Sahle, Sven, Gauges, Ralph, Lee, Christine, Pahle, Jürgen, Simus, Natalia, Singhal, Mudita, Xu, Liang, Mendes, Pedro, Kummer, Ursula, 2006. COPASI—a COmplex PAthway SImulator. Bioinformatics 22 (24), 3067–3074.
Ishii, Nobuyoshi, Suga, Yoshihiro, Hagiya, Akiko, Watanabe, Hisami, Mori, Hirotada, Yoshino, Masataka, Tomita, Masaru, 2007. Dynamic simulation of an in vitro Multi-enzyme system. FEBS (Fed. Eur. Biochem. Soc.) Lett. 581 (3), 413–420.
Jamshidi, Neema, Palsson, Bernhard Ø., 2010. Mass action stoichiometric simulation models: Incorporating kinetics and regulation into stoichiometric models. Biophys. J. 98 (2), 175–185.
Jiménez, Estela Sánchez de, Lee, Evangelina, Torres, Jesus, Soberón, Guillermo, 1964. On the mechanism of the effect of ionic strength on Crystalline Aldolase activity. J. Biol. Chem. 239 (12), 4154–4158.
Khodayari, Ali, Maranas, Costas D., 2016. A genome-scale Escherichia coli kinetic metabolic model K-ecoli457 satisfying flux data for multiple Mutant Strains. Nat. Commun. 7 (December), 13806.
Klinman, Judith P., Amnon, Kohen, 2013. Hydrogen tunneling Links protein dynamics to enzyme catalysis. Annu. Rev. Biochem. 82, 471–496.
Kotte, Oliver, Heinemann, Matthias, 2009. A Divide-and-Conquer approach to Analyze underdetermined biochemical models. Bioinformatics 25 (4), 519–525.
Kuzmic, P., 1996. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV Proteinase. Anal. Biochem. 237 (2), 260–273.
Li, Feiran, Yuan, Le, Lu, Hongzhong, Li, Gang, Chen, Yu, Engqvist, Martin K.M., Kerkhoven, Eduard J., Nielsen, Jens, 2022. Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. Nat. Catal. 5 (8), 662–672.
Linden, Nathaniel J., Kramer, Boris, Rangamani, Padmini, 2022. Bayesian parameter estimation for dynamical models in systems biology. PLoS Comput. Biol. 18 (10), e1010651.
Link, Hannes, Kochanowski, Karl, Sauer, Uwe, 2013. Systematic Identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. Nat. Biotechnol. 31 (4), 357–361.
Ma, Buyong, Nussinov, Ruth, 2013. Structured Crowding and its effects on enzyme catalysis. Top. Curr. Chem. 337, 123–137.
Millard, Pierre, Smallbone, Kieran, Mendes, Pedro, 2017. Metabolic regulation is sufficient for Global and robust Coordination of glucose Uptake, Catabolism, Energy production and growth in Escherichia coli. PLoS Comput. Biol. 13 (2), e1005396.
Miskovic, Ljubisa, Hatzimanikatis, Vassily, 2010. Production of Biofuels and Biochemicals: in need of an ORACLE. Trends Biotechnol. 28 (8), 391–397.
Monk, Jonathan M., Lloyd, Colton J., Brunk, Elizabeth, Mih, Nathan, Sastry, Anand, King, Zachary, Takeuchi, Rikiya, et al., 2017. iML1515, a Knowledgebase that Computes Escherichia coli Traits. Nat. Biotechnol. 35 (10), 904–908.
Moxley, Michael A., Beard, Daniel A., Bazil, Jason N., 2014. A pH-dependent kinetic model of Dihydrolipoamide Dehydrogenase from multiple organisms. Biophys. J. 107 (12), 2993–3007.
Noor, Elad, Flamholz, Avi, Liebermeister, Wolfram, Bar-Even, Arren, Milo, Ron, 2013. A note on the kinetics of enzyme action: a Decomposition that highlights thermodynamic effects. FEBS (Fed. Eur. Biochem. Soc.) Lett. 587 (17), 2772–2777.
Olp, Michael D., Kalous, Kelsey S., Smith, Brian C., 2020. ICEKAT: an interactive online tool for calculating initial rates from Continuous enzyme kinetic Traces. BMC Bioinf. 21 (1), 186.
Placzek, Sandra, Schomburg, Ida, Chang, Antje, Jeske, Lisa, Ulbrich, Marcus, Tillack, Jana, Schomburg, Dietmar, 2017. BRENDA in 2017: new Perspectives and new tools in BRENDA. Nucleic Acids Res. 45 (D1), D380–D388.
Saa, Pedro A., Nielsen, Lars K., 2016. Construction of feasible and accurate kinetic models of metabolism: a Bayesian approach. Sci. Rep. 6 (1), 1–13.
Saa, Pedro A., Nielsen, Lars K., 2017. Formulation, construction and analysis of kinetic models of metabolism: a Review of Modelling frameworks. Biotechnol. Adv. 35 (8), 981–1003.
Saa, Pedro, Nielsen, Lars K., 2015. A general framework for Thermodynamically consistent parameterization and efficient sampling of enzymatic reactions. PLoS Comput. Biol. 11 (4), e1004195.
Sanft, K.R., Gillespie, D.T., Petzold, L.R., 2011. Legitimacy of the stochastic Michaelis-Menten approximation. IET Syst. Biol. 5 (1), 58.
Savoglidis, Georgios, Xavier da Silveira dos Santos, Aline, Riezman, Isabelle, Angelino, Paolo, Howard, Riezman, Hatzimanikatis, Vassily, 2016. 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. 37 (Suppl. C), 46–62.
Schmidt, Alexander, Kochanowski, Karl, Vedelaar, Silke, Ahrné, Erik, Volkmer, Benjamin, Callipo, Luciano, Knoops, Kèvin, Bauer, Manuel, Aebersold, Ruedi, Heinemann, Matthias, 2016. The quantitative and condition-dependent Escherichia coli Proteome. Nat. Biotechnol. 34 (1), 104–110.
Scopes, R.K., 1995. The effect of temperature on enzymes used in Diagnostics. Clinica Chimica Acta; International Journal of Clinical Chemistry 237 (1–2), 17–23.
Shepelin, Denis, Machado, Daniel, Nielsen, Lars K., Herrgård, Markus J., 2020. Benchmarking Kinetic Models of Escherichia Coli Metabolism. https://doi.org/10.1101/2020.01.16.908921 bioRxiv.
Spivey, H.O., Ovádi, J., 1999. Substrate channeling. Methods 19 (2), 306–321.
Srinivasan, Shyam, Cluett, William R., Mahadevan, Radhakrishnan, 2015. Constructing kinetic models of metabolism at genome-Scales: a Review. Biotechnol. J. 10 (9), 1345–1359.
Teusink, Bas, Passarge, Jutta, Reijenga, Corinne A., Esgalhado, Eugenia, van der Weijden, Coen C., Schepper, Mike, Walsh, Michael C., et al., 2000. Can Yeast glycolysis Be Understood in terms of in vitro kinetics of the Constituent enzymes? Testing Biochemistry. Eur. J. Biochem./FEBS 267 (17), 5313–5329.
Tipton, K.F., Dixon, H.B., 1979. Effects of pH on enzymes. Methods Enzymol. 63, 183–234.
Tran, Linh M., Rizk, Matthew L., Liao, James C., 2008. Ensemble modeling of metabolic networks. Biophys. J. 95 (12), 5606–5617.
Tzafriri, A.R., 2003. Michaelis-menten kinetics at high enzyme concentrations. Bull. Math. Biol. 65 (6), 1111–1129.
Wang, C.S., Alaupovic, P., 1980. Glyceraldehyde-3-Phosphate Dehydrogenase from human Erythrocyte Membranes. Kinetic mechanism and competitive substrate inhibition by glyceraldehyde 3-phosphate. Arch. Biochem. Biophys. 205 (1), 136–145.
Yurkovich, James T., Zielinski, Daniel C., Yang, Laurence, Paglia, Giuseppe, Rolfsson, Ottar, Sigurjónsson, Ólafur E., Broddrick, Jared T., et al., 2017. Quantitative time-course Metabolomics in human red Blood Cells reveal the temperature dependence of human metabolic networks. J. Biol. Chem. 292 (48), 19556–19564.