5

Kinetic Models of Metabolism

Hongzhong Lu[1], Yu Chen[1], Jens Nielsen[1,2], and Eduard J. Kerkhoven[1]

[1] Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden
[2] BioInnovation Institute, Copenhagen N, Denmark

5.1 Introduction

Metabolism is the sum of all biochemical reactions that take place in living cells. Chapters 2 and 4 show that metabolism can be mathematically converted into equations by connecting all the reactions and metabolites based on reaction stoichiometry. The rate of each metabolic reaction can accordingly be estimated using constrained-based approaches, e.g. flux balance analysis (Chapter 2). While such approaches rely mainly on the stoichiometry of all reactions, rates of single reactions could also be calculated based on their own characteristics including enzyme concentrations, enzyme properties, metabolite concentrations, as well as how these factors can be integrated, i.e. the rate expressions, which is referred to as reaction kinetics. This encourages to integrate reaction kinetics with stoichiometric models, resulting in kinetic models of metabolism. Kinetic models are more explicit than stoichiometric models, and therefore suitable for some particular applications and analyses, especially metabolic control analysis (MCA), which will be detailed in Chapter 6. Here the kinetic models of metabolism are introduced followed by examples on the construction of kinetic models as well as applications. Note that most of the principles related with metabolism are also relevant for other dynamic processes in biology that could be described using kinetic models, including signaling pathways, pharmacokinetics, circadian rhythms, the cell-cycle, population dynamics, and so on.

5.2 Definition of Enzyme Kinetics

5.2.1 Michaelis–Menten Formula

The Michaelis–Menten formula is one of the earliest and best known mechanistic models to describe enzyme kinetics [1], which we will regard here as the basis for

Metabolic Engineering: Concepts and Applications, First Edition.
Edited by Sang Yup Lee, Jens Nielsen, and Gregory Stephanopoulos.
© 2021 WILEY-VCH GmbH. Published 2021 by WILEY-VCH GmbH.

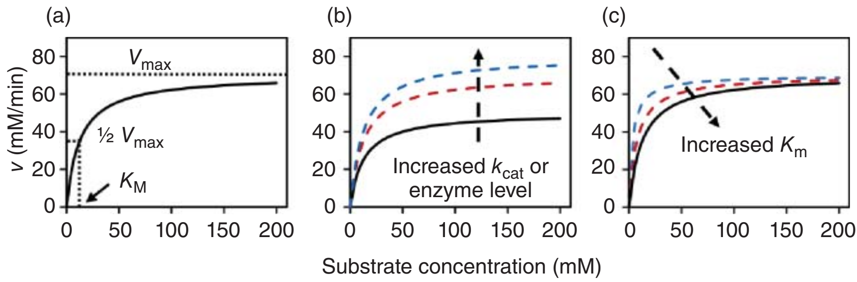


Figure 5.1 Relationship between parameters and reaction rates in Michaelis–Menten formulated enzyme kinetics. (a) K_M is equal to the substrate concentration when the reaction rate v reaches half of the maximum reaction rate (V_{max}). (b) The maximum reaction rate is elevated when increasing the k_{cat} or enzyme level. (c) The reaction rate is decreased when increasing the K_M.

the building of kinetic models that is introduced later in this chapter. Note that not all enzymes demonstrate Michaelis–Menten kinetics, but it is a common and versatile approximation for most reactions. At the first glance, the reaction rate dependency on substrate concentration can be accurately described using the Michaelis–Menten formula (Figure 5.1a). In general, this expression is suitable for one-substrate reactions without backward reactions and effectors:

$$E + c_S \underset{k_1}{\overset{k_{-1}}{\rightleftarrows}} Ec_S \rightarrow^{k_{cat}} E + c_P, \tag{5.1}$$

where k_1 and k_{-1} are the association and disassociation constants for the enzyme and the substrate, respectively; and k_{cat} is the catalytic rate constant, or the turnover number of the enzyme. c represents metabolite, where c_S is substrate and c_P is product, while E and Ec_S are the enzyme and enzyme–substrate complex, respectively.

By taking the assumption that the concentration of enzyme is much smaller than the concentration of substrate, the product formation rate can be calculated using the Michaelis–Menten formula:

$$v = \frac{dc_P}{dt} = \frac{V_{max} \cdot c_S}{c_S + K_M} \tag{5.2}$$

In the above expression,

$$K_M = \frac{k_{-1} + k_{cat}}{k_1} \tag{5.3}$$

$$V_{max} = k_{cat} E_{total} \tag{5.4}$$

$$\frac{dEc_S}{dt} + \frac{dE}{dt} = 0 \text{ or } E_{total} = E + Ec_S = \text{constant}, \tag{5.5}$$

where V_{max} is the maximum reaction rate and K_M is the Michaelis or affinity constant for the substrate, which are two important kinetic parameters of enzymes. E_{total} represents the total enzyme concentration, by combining the concentration.

5.3 Factors Affecting Intracellular Enzyme Kinetics

With the Michaelis–Menten formulation, the influence of enzyme properties (K_M, k_{cat}), enzyme abundance (E_{total}), and metabolite concentration (c) on the dynamic behavior of a reaction can be explained mechanistically. Firstly, the k_{cat} reflects the maximum catalytic efficiency of an enzyme. As k_{cat} increases, the maximum reaction rate is enhanced accordingly (Figure 5.1b). The k_{cat} of a specific enzyme by itself can be influenced through inhibition, activation, synergistic effects, and allosteric regulation. For example, so-called noncompetitive inhibition could slow down the enzyme catalytic efficiency, as binding of the inhibitor prevents the enzyme from catalyzing the reaction, thereby reducing the amount of effective enzyme and reducing the apparent k_{cat}.

As another key enzyme parameter, K_M characterizes the affinity between the enzyme and its reactant c. Its value is not constant across reactions and depends on enzyme structure, substrate structure, as well as the environmental pH, temperature, and ionic strength. Once the external conditions are fixed, K_M is equal to the substrate concentration at which the reaction rate reaches half of V_{max} (Figure 5.1a). If $c_S \gg K_M$, the reaction rate approaches V_{max} while if $c_S \ll K_M$, the reaction rate is then simplified as:

$$v = \frac{V_{max} \cdot c_S}{K_M} \tag{5.6}$$

In this case of low c_S concentration, there exists an approximately linear correlation between the reaction rate and substrate concentration (Figure 5.1a). Meanwhile, if the substrate concentration is fixed, an enzyme with a higher K_M value results in a lower reaction rate (Figure 5.1c). Generally, the ratio of k_{cat}/K_M can be used to evaluate the effectiveness of enzymes from different sources that all catalyze the same reaction: a higher k_{cat}/K_M indicates a higher catalytic efficiency of an enzyme, as a higher reaction rate can be reached with less substrate present.

In addition to the enzyme characteristic parameters, also the enzyme levels constrain the rate through each reaction. Increases in the level of enzyme translate to increases in reaction rates (Figure 5.1b), especially if the substrate concentration is high ($c_S > K_M$). *In vivo*, the intracellular enzyme levels are a tradeoff between the protein synthesis and protein degradation, which by themselves are affected by processes such as product inhibition, transcriptional, and feedback regulation. In metabolic engineering, enzyme levels can therefore be augmented through multiple strategies, i.e. increasing the gene copies, improving the enzyme stability, and relieve feedback inhibition of end products.

Although the enzyme kinetics are useful to describe enzyme dynamics in isolation, they are not sufficient to describe the dynamic changes in metabolite concentrations and reaction rates *in vivo*. In reality, enzymes function in the context of a metabolic network consisting of many different reactions and enzymes, and dynamics of enzymes will mutually influence each other. This gives rise to emerging properties of the metabolic network, which is a function of the enzyme kinetics but can only be observed when studying the whole system and not only its

constituent parts in isolation. Therefore, it is required to take the enzyme kinetics and combine these in kinetic models.

5.4 Kinetic Model: Definition and Scope

5.4.1 What Is a Kinetic Model?

Kinetic models mechanistically represent the processes that take place within a cell, and these models are made up of a series of ordinary differential equations (ODEs). The ODEs encompass the detailed rate expressions and kinetic parameters that describe the dynamic behavior of individual reactions within the model, as described in the previous section. The mathematical formalisms of kinetic models, whether they follow Michaelis–Menten or other kinetics, can be summarized using Eq. (5.7). In the kinetic model, the enzyme kinetics and levels are the parameters of the model while metabolite concentrations are variables, all of which are absent in purely stoichiometric models. Thus, compared with stoichiometric models, a kinetic model can predict changes and dynamics of reaction rates (fluxes) and metabolites concentration over time [2].

$$\frac{dc_i}{dt} = \boldsymbol{S} \cdot \boldsymbol{v}(E; c_i; k), i = 1, 2, 3, \ldots, n, c_i(0) = c_{i,0} \tag{5.7}$$

In Eq. (5.7), $\boldsymbol{S}$ represents the stoichiometric matrix, $\boldsymbol{v}$ represents the vector of metabolic reactions or fluxes, and $c_{\text{i},0}$ represents the initial metabolite concentration in the system. Each reaction rate (v_j) is determined by the enzyme abundance (E_j), metabolite concentration (c_i), and the corresponding kinetic parameters (k).

5.4.2 Scope of Kinetic Models

Typically, kinetic models consist of tens to hundreds of metabolic reactions with their detailed kinetic information from one or several sub-pathways. The scope and size of a kinetic model depends on the computational resources and the scientific questions to be answered. For long, kinetic models have been used to describe the dynamics of sub-pathways consisting of up to 10–20 reactions, such as glycolysis in beef heart supernatant by Garfinkel et al. in 1968 [3]. Since then, larger kinetic models have frequently been reconstructed while centering on multiple core metabolic pathways including glycolysis (EMP pathway), the pentose phosphate (PP) pathway, and tricarboxylic acid (TCA) cycle, which generally contain 50–100 reactions with detailed ODEs. To further increase the coverage of cellular metabolism, near genome-scale kinetic models with over 200 reactions have been developed for several intensively studied model organisms, i.e. *E. coli* [4] and *S. cerevisiae* [5]. However, to date no full genome-scale models (GEMs) with detailed enzyme kinetics have been built for *E. coli* and *S. cerevisiae*, let alone lesser studied organisms. A kinetic model requires the definition of rate equations and their respective parameters for each of the reactions, which are currently unknown for many of the reactions contained in GEMs. As an example, while there are about 700 Enzyme

Commission (EC) numbers associated to the *S. cerevisiae* GEM (Yeast8.3), only about 42 of these have K_M values recorded in the relevant databases. An alternative approach to overcome this lack of kinetic parameters for large size models has been to use approximative rate equations, which will be discussed below.

5.4.3 How to Build a Functional Kinetic Model?

The general procedures to establish a functional kinetic model are summarized in Figure 5.2. After describing the metabolic network for which a model will be constructed, the kinetic rate expressions for individual reactions are gathered and subsequently combined with their respective parameter values to build a complete model. In detail, the procedures to establish a kinetic model are divided into the following five steps:

Step 1. Define the metabolic network. In this step, one needs to decide on the scope of the metabolic network to be modeled, i.e. which sub-pathways to include. Based on this a detailed metabolic network structure is built, encompassing the stoichiometry of metabolites, reactions, and their respective enzymes. Ideally, also information regarding regulation and interaction among the components of the network should be gathered. The metabolic network is generated based on the genome annotation of the organism of interest and by consulting previous studies (Figure 5.2).

Step 2. Define the kinetic rate expressions. Each reaction in the metabolic model will be assigned a rate expression, of which the Michaelis–Menten equation is one example. To infer kinetic rate expressions, the biochemical and mechanistic information should be gathered from biological databases and literature.

Step 3. Assign parameter values. The rate expressions from Step 2 require parameterization, and these enzyme-specific parameters are either measured from experiments or queried from literature and/or databases. For unknown parameters, their value should be obtained by, e.g. taking reported values from the same reaction but a different organism, or through inference by simulating the

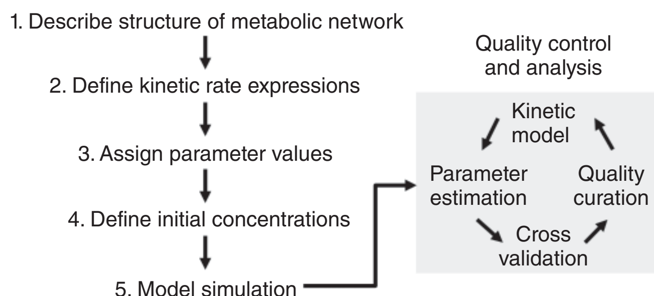


Figure 5.2 Framework to build a functional kinetic model.

model with arbitrary parameter values and compare the simulation results with measured data.

Step 4. Define initial concentrations of metabolites and enzyme levels, based on measured values or reported data.

Step 5. Conduct simulation with the complete kinetic model. With the information from Steps 1 to 4, a kinetic model containing ODEs are defined (Eq. (5.7)). During simulations with this model, measured physiological data, metabolite concentrations, enzyme levels, and ^{13}C labeled fluxes can all be used to evaluate the predictive performance of the kinetic model. Once a kinetic model of high quality is obtained, it will be further used for practical applications in metabolic engineering and biological discoveries, as discussed later in this chapter.

As the reaction rate expressions and their related kinetic parameters, Steps 2 and 3, are arguably the most important components of a kinetic model, they will be discussed in more detail below.

5.5 Main Mathematical Expressions in Description of Reaction Rates

Reaction kinetics can be described with mathematical expressions where the reaction rates are functions of kinetic parameters and the concentration of metabolites. These rate expressions are of varying complexity, referring to the catalytic mechanism they describe, potential regulatory properties, and the number of required parameters, and include mechanistic, approximate, and stochastic formulas. Of these, the stochastic formulas are typically the most complex and computationally intensive, while their use should be deliberated if stochasticity is expected to play an important role in the simulated system, such as noise in signal transduction or gene expression. However, such processes can often be disregarded when studying a large population of cells, such as a bioreactor cultivation of micro-organisms, as the stochastic behavior is normalized over the sheer number of cells. Instead, mechanistic (Box 5.1) and approximate (Box 5.2) rate expression are most frequently used for kinetic models of metabolism.

5.5.1 Mechanistic Rate Expressions

Underlying mechanistic rate expressions is the law of mass action, which assumes that a reaction rate is proportional to the concentration of its reactants. While mass action expressions are suitable for enzymatic or transporter reactions, the equations become very large if every step of the enzyme catalytic process is to be described and this results in many parameters that are hard to measure (see Eq. (5.1), there would also be association and dissociation constants for the product and a reverse catalytic constant). Instead, mass action kinetics is typically reduced to the aforementioned Michaelis–Menten equation with its apprehensible parameters, while the Hill equation is another example of a mechanistic rate expression (Box 5.1). These mechanistic expressions distinctly

clarify the roles of various factors on the reaction rates, as shown in the above Michaelis–Menten equations. More importantly, such mechanistic expressions can easily be extended based on new experimental evidences, to cover reactions with more reactants and products, as well as with complex regulatory mechanisms, i.e. activation and inhibition in enzyme activity by metabolites (Box 5.1).

Box 5.1 Mechanistic rate expressions: two typical examples.

Michaelis–Menten. As introduced earlier in this chapter, the Michaelis–Menten formulation is a mechanistic rate expression that can be used when the reaction kinetics follow the distinct hyperbolic saturation curve (Figure 5.1). Based on the simplified case shown in Eq. (5.2), the Michaelis–Menten formulation can be extended to describe more complex cases, such as competitive substrate inhibition, where a nonreactant metabolite competes with the substrate to bind the same part of the enzyme.

$$v = \frac{dc}{dt} = \frac{V_{\max} \cdot c}{K_M + c(1 + (c/K_i))}, \quad K_i \text{ is the inhibition constant} \tag{5.8}$$

Hill equation. Enzymes that have sigmoidal saturation curves can have their kinetics described using a Hill equation. A typical case of such an enzyme is a homomultimer, where the affinity of the subunits for the substrate increases if one or more other subunits are already bound to the substrate. This cooperativity is represented in the unitless Hill coefficient n which represents positive cooperativity with a value >1:

$$v = \frac{dc}{dt} = \frac{V_{\max} \cdot c^n}{{K_{0.5}}^n + c^n}, \tag{5.9}$$

where $K_{0.5}$ is the half-maximal concentration constant. If $n = 1$, $K_{0.5}$ is equal to the Michaelis constant (K_M).

5.6 Approximative Rate Expressions

While mechanistic rate expressions are simplifications of mass action kinetics, they can still be very complex for reactions with multiple reactants, products, and/or intertwined metabolic regulation, such as the influences of pH, temperature, and cofactor concentrations. Moreover, detailed mechanistic knowledge and measured parameter values are typically inadequate for enzymes that are not catalyzing reactions in central carbon metabolism. For nonmodel organisms, this situation is further exacerbated by the sheer lack of measured kinetic parameters. To circumvent these issues, approximative rate expressions is also adopted in the kinetic model reconstruction. There are various approximative rate expressions, including generalized mass action, log-lin, and lin-log (Box 5.2). Compared with the mechanistic Michaelis–Menten formulation, the approximative rate expressions are of lower complexity, thereby enhancing the computation efficiency for large-scale kinetic models. Simulations of central carbon metabolism in *E. coli*

using a lin-log model are consistent with a mechanistic model [6], even when simple structures and few parameters were used in the lin-log model. Due to the above advantages, lin-log expressions have been used to reconstruct a kinetic model of yeast metabolism with 240 reactions [5].

Box 5.2 Approximate expressions: two typical examples.

Power-laws. Generalized mass action (GMA) is a so-called power law formalism with noninteger exponents (Eq. (5.10)) [7, 8]. In the GMA expressions, the reaction rate is proportional to the enzyme activity, as well as to the power law involving dependent and independent concentrations of metabolites. Compared with mechanistic rate expressions, the GMA expressions reduce the parameters in the formula.

$$v_i = k_i \cdot E_i \cdot \prod_{j=1}^{m} c_{\text{in},j}^{a_{ij}} \cdot \prod_{k=1}^{n} c_{\text{ex},k}^{b_{ik}}, \tag{5.10}$$

where E_i represents the enzyme level; c_{in} and c_{ex} are dependent (intracellular) and independent (extracellular, or constant) metabolite concentrations, respectively; k_i, a_{ij}, and b_{ik} denote kinetic coefficients that can be obtained from fitting the equation to observed enzyme dynamics.

Lin-log. The lin-log modeling approach is basically the same as the power law formulation, but through transformation to the logarithmic domain a set of linear equations are obtained, which significantly simplify parameter estimations. Thus, even though the relation between the reaction rate and enzyme level, metabolites concentration and kinetic parameters is highly nonlinear, the linear logarithmic (lin-log) approximation handles reaction rates in a simple and linear manner to obtain an analytic solution. In the lin-log expressions, the rate is proportional to enzyme level, as well as to a linear sum of logarithms of metabolites concentration [9] (the so-called lin-log), which is thus beneficial to reduce the kinetic parameters used in the rate expressions.

$$\begin{aligned} v_i = E_i(a_i &+ p_{i.1} \ln c_{\text{in},1} + p_{i.2} \ln c_{\text{in},2} + \cdots + p_{i.m} \ln c_{\text{in},m} + q_{i.1} \ln c_{\text{ex},1} \cdot \\ &\cdot + q_{i.2} \ln c_{\text{ex},2} + \cdots + q_{i,r} \ln c_{\text{ex},r}), \end{aligned} \tag{5.11}$$

where E_i represents the enzyme level; c_{in} and c_{ex} are dependent (intracellular) and independent (extracellular, or constant) metabolite concentrations, respectively; p and q represent the independent kinetic coefficients that are obtained by fitting the rate expression to observed enzyme dynamics.

5.7 Approaches to Assign Parameters in the Rate Expressions

Estimation of parameters in rate expressions is essential for having good predictive performance of a kinetic model. However, the determination of parameter through experimentation, estimation, or fitting is one of most challenging aspects of kinetic model reconstruction. There are three main procedures to

obtain kinetic parameters that are included in a rate expression, namely direct experimental measurement through enzyme assays; querying databases and literature of previously reported enzyme parameters; and parameter inference using statistical analysis.

5.7.1 Direct Measurements of Kinetic Parameters in Enzyme Assays

Enzyme parameters can be measured in *in vitro* enzyme assays, using either purified enzymes or whole-cell extracts. The reaction rate is typically measured colorimetrically, e.g. by coupling the product formation with a chemical reaction that causes a color change. Sequentially changing the substrate concentration in the enzyme assays can then yield the K_M and V_{max} parameters that are part of Michaelis–Menten expressions. Here it should be noted that the *in vitro* measured parameters can be different from those observed *in vivo*, even while a positive correlation between the *in vitro* and *in vivo* kinetic parameters exists [10]. Such inconsistencies are likely due to the fact that *in vitro* assay conditions are distinct from the *in vivo* intracellular microenvironment, e.g. metabolite concentrations, temperature, pH, osmotic strength, and presence of potential inhibitors. Nonetheless, it is not uncommon that V_{max} values are reported at assay conditions that give the highest activity but that are far from *in vivo*-like conditions, and Van Eunen et al. showed that the use of such unrealistic parameters resulted in unrealistic metabolite concentrations in a model of yeast glycolysis [11]. It is therefore important to perform enzyme assays in conditions that are close to *in vivo* conditions.

5.7.2 Querying Databases

Instead of measuring all kinetic parameters that are required for a kinetic model, one can also leverage previous work and download kinetic data of specific enzymes from the BRENDA (https://www.brenda-enzymes.org) and SABIO-RK (http://sabio.villa-bosch.de/SABIORK) databases (Figure 5.3 as an example), which now contain 68 963 and 648 732 enzyme entries, respectively, with detailed annotation information [12, 13]. When querying and using this public data in kinetic models, one implicitly assumes that the measured kinetic parameters are conserved or similar for strains of different genetic background, as the databases collate data from many different studies and not measurements from one single strain isolate. Moreover, the parameters compiled in these databases are measured under many different experimental conditions, of which many are not representing relevant *in vivo*-like conditions. This does not render the information from these databases unusable, but it does indicate that one should be critical about the provided parameter values. In addition, the available parameters are not evenly distributed across all enzyme catalyzed reactions: many enzymes have been subjected to few studies, while for some enzymes many experimentally measured parameter values are provided. To fully address all these issues, it becomes near indispensable to perform parameter estimation based on large scale of physiological data, as will be discussed later in this chapter.

General information	
Organism	Saccharomyces cerevisiae
Strain	BY4700 transformed in Y258
Tissue	-
EC Class	5.3.1.9
SABIO reaction id	1123
Variant	wildtype
Recombinant	expressed in Saccharomyces cerevisiae
Experiment Type	in vitro
Pathways	Glycolysis classical Glycolysis/Gluconeogenesis
Event Description	-

Substrates		
name	location	comment
D-Fructose 6-phosphate	-	-

Products		
name	location	comment
D-Glucose 6-phosphate	-	-

Modifiers				
name	location	effect	comment	protein complex
glucose-6-phosphate isomerase(Enzyme)	-	Modifier-Catalyst	-	(P12709)*2;

Enzyme (protein data)				
	UniProtKB_AC	name	mol. weight (kDa)	deviation (kDa)
subunit	P12709	-	-	-
complex	-	-	-	-

Kinetic Law		
type	formula	annotation
Michaelis-Menten	kcat * Et * S / (Ks + S)	SBO:0000029

Parameter							
name	type	species	start val.	end val.	deviat.	unit	comment
S	concentration	D-Fructose 6-phosphate	0.0	2.0	-	mM	-
Et	concentration	Enzyme	6.67E-7	-	-	mM	-
kcat	kcat	-	247.2	-	5.1	s^(-1)	-
Ks	Km	D-Fructose 6-phosphate	0.307	-	0.021	mM	-

Experimental conditions			
	start value	end value	unit
pH	6.5	-	-
temperature	30.0	-	°C
buffer	0.0 mM Glucose-6-phosphate 1-dehydrogenase, 100.0 mM 2-(N-morpholino)ethanesulfonic acid, 5.0 mM magnesium dichloride, 100.0 mM potassium chloride, 0.4 mM NADP		

Figure 5.3 An entry example from SABIO-RK[13] (http://sabio.villa-bosch.de/SABIORK) with detailed reaction and kinetic information for the enzyme glucose-6-phosphate isomerase. Source: HITS, gGmbH.

5.7.3 Inferring from Measured Fluxes

Besides directly measuring kinetic parameters, *in vivo* k_{cat} values can be estimated from accurate measurements of fluxes and protein abundances [10] (Figure 5.4). Fluxes through the individual reactions can be determined based on ^{13}C metabolic flux analysis, or alternatively from flux balance analysis (FBA). Enzyme levels can subsequently be determined by proteomics, where all protein levels are quantitatively measured by mass spectrometry. From this, the apparent k_{cat} is calculated as the ratio between the fluxes and the corresponding enzyme levels. By applying this approach across a series of different cultivation conditions, across maximum apparent k_{cat} values are regarded as the maximum

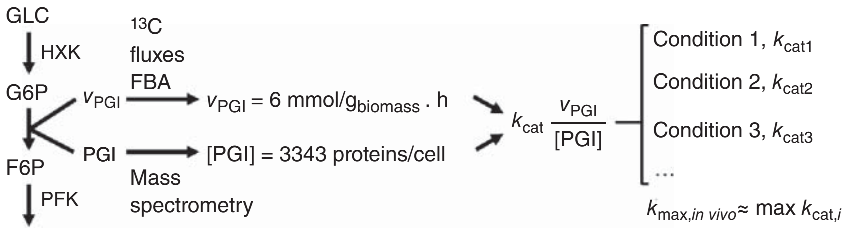


Figure 5.4 Procedure to calculate *in vivo* k_{cat} values. For each reaction, the k_{cat} is calculated based on fluxes and enzyme abundances. By performing this procedure under a series of cultivation conditions, reaction specific maximum k_{cat} values are assigned.

in vivo k_{cat} (Figure 5.4). Through this procedure, relatively reliable k_{cat} values can be obtained from their relevant *in vivo* conditions without laborious enzyme assays. A drawback of this procedure is that it requires accurate simultaneous determination of fluxes and proteomics data.

5.7.4 Parameters Inference Using the Statistical Analysis

While kinetic parameters can be obtained from databases or measured *in vivo* or *in vitro*, it is not unlikely that a large number of parameters are still unknown when constructing a kinetic model, especially for a model of large size. Consequently, parameter inference through data fitting is increasingly used to get a fully functional kinetic model. In general, parameter estimation aims to sample reasonable parameter sets that minimize the distance between model predictions and experimental data. Parameter estimation can be a complex process for large-scale kinetic models, as the number of rate expressions with unknown parameter values increases drastically. Thus, effective parameter fitting methods become important to balance computation cost and accuracy in the parameter inference. Various algorithms have been conceived to conduct the parameter estimation, of which Maximum Likelihood Estimator (MLE)- and Monte Carlo-based approaches are widely used.

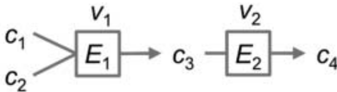
MLE-based approaches try to find the best estimate for each parameter and quantify the uncertainties in these estimates, while Monte Carlo-based approaches sample from probability distributions for each parameter, to extract values that result in a reasonable output based on the optimized objective function. Because of the nonlinear structures of kinetic models and the typical large number of unknown kinetic parameters, MLE-based approaches are not suitable for reasonable parameter fitting of complex and large kinetic models. In contrast, Monte Carlo-based approaches estimate parameter distributions instead of singular parameter values; this partially overcomes the drawbacks of MLE, and is therefore commonly used for parameter estimation of large kinetic models. Regardless which algorithm is used for parameter inference, the estimated parameters should be comprehensively evaluated. For this, simulation results that are obtained from the constructed kinetic models are to be compared with relevant experimental measurements, until the model is able to yield acceptable results.

Example 5.1 ***A toy kinetic model.***

A toy model with two reactions, two enzymes, and four metabolites, is used here to illustrate how to generate a simple kinetic model for specific sub-pathways, and how to use this kinetic model to predict the dynamic behaviors of the system (Figure 5.5). The five steps of kinetic model reconstruction are:

1) Describe the structure of the toy network, made up of two reactions that are catalyzed by enzymes E_1 and E_2, with their corresponding reaction rates v_1 and v_2.
2) Define the kinetic rate expressions. Michaelis–Menten kinetics are defined for both reactions, where the first reaction has two substrates and therefore a more expanded rate expression. Based on the network structure and rate expressions, mass balance equations are defined for the four metabolites (c_1, c_2, c_3, c_4), which describe how their respective concentrations are affected by the reaction rates.
3) Assign parameters values for the two rate expressions from measurements or databases.
4) Define initial concentrations of metabolites and enzymes, according to the experimental condition or measurement.
5) Model simulation. The toy kinetic model can now be used to predict the evolution of fluxes and metabolites concentration data over time. The output of such a time-course simulation is shown in Figure 5.5, where the concentration of the first set of metabolites (c_1, c_2) decrease gradually, while the concentration of

1. A toy model



2.1 Rate expressions

$$v_1 = \frac{k_{cat1} \cdot E_1 \cdot c_1 \cdot c_2}{(K_{M1} + c_1)(K_{M2} + c_2)}$$

$$v_2 = \frac{k_{cat2} \cdot E_2 \cdot c_3}{K_{M3} + c_3}$$

2.2 Mass balance

$$\frac{dc_1}{dt} = -v_1 \qquad \frac{dc_2}{dt} = -v_1$$

$$\frac{dc_3}{dt} = v_1 - v_2 \qquad \frac{dc_4}{dt} = v_2$$

3. Assign parameters

For enzyme 1 (E_1)
$k_{cat1} = 0.18$
$K_{M1} = 6$
$K_{M2} = 3$

For enzyme 2 (E_2)
$k_{cat2} = 0.04$
$K_{M3} = 1.5$

4. Initial values

$c_{1,t=0} = 6$ $\quad c_{2,t=0} = 4$ $\quad c_{3,t=0} = 0$
$c_{4,t=0} = 0$ $\quad E_{1,t=0} = 6$ $\quad E_{2,t=0} = 8$

5. Simulation

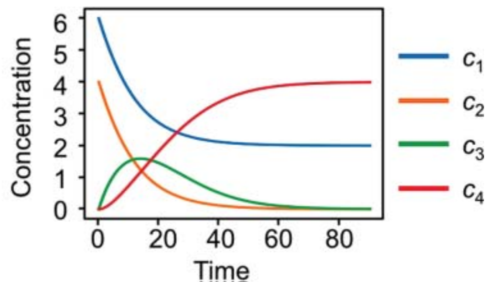


Figure 5.5 A toy example to show how to build kinetic model.

the intermediate metabolite (c_3) first rises and then falls after 20 minutes, and the concentration of end metabolite (c_4) is gradually increasing.

Example 5.2 ***A functional kinetic model for core metabolic pathway of yeast.***

To display kinetic model reconstruction of a real sub-pathway, a kinetic model covering glycolysis in yeast is taken as an example (Figure 5.6). To obtain a functional kinetic model predicting the dynamic behavior of this pathway, the following steps are essential:

1) Define the metabolic network. From literature reports, we can define the metabolic pathway that we want to model, including detailed annotation on reactions, metabolites, and enzymes. It is not uncommon to lump very long pathways into simpler reactions, to reduce the complexity of the subsequent kinetic model and decrease the number of unknown parameters. Here the reaction network of EMP pathway is listed as a simple example (Figure 5.6) of the real metabolic models, which consists of metabolites and enzymes for each reaction from EMP pathway.
2) Define the kinetic rate expressions. Shown are the Michaelis–Menten rate expressions for phosphoglucose isomerase (PGI) and phosphoglycerate mutase (PGM), who among others affect the concentrations of glucose 6-phosphate and 3-phosphoglycerate, as detailed in their mass balances.
3) Assign parameters values, where enzymes that have had their kinetics characterized can have their kinetic parameters collected from literature or databases, shown here for PGI. Unknown parameters are estimated by parameter inference while the known parameters are set.

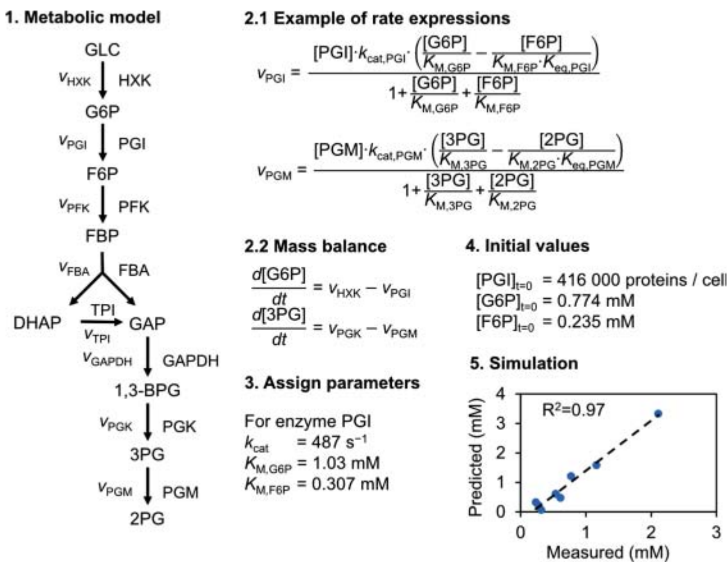


Figure 5.6 Detailed steps to build a functional kinetic model which could predict the cellular physiology. Source: Adapted from Smallbone et al. [14].

4) Define initial concentrations based on the measured metabolite concentrations (or assumption thereof) and measured enzyme levels.
5) Model simulation. Predicted fluxes, enzyme levels, and metabolite concentrations should be compared with experimental values to validate the model. Typically, multiple rounds of curation are required to result in a model that is highly consistent with experimental measurements. Once the quality of the kinetic model has been deemed sufficient, it will be used in the further applications, such as metabolic control analysis (Chapter 6) or strain design in metabolic engineering.

5.8 Applications

Kinetic models can be used for various applications, which are largely divided into three groups, i.e. (i) metabolic control analysis (MCA)-based methods; (ii) time-course simulations of dynamic processes; and (iii) integrative analysis of omics data.

Kinetic models enable the calculation of control of each enzyme on the flux through a pathway through MCA-based approaches (Figure 5.7a). Key enzymes that exert high control on a particular pathway indicate promising targets in the fields of biotechnology and systems medicine. Indeed, kinetic models have succeeded in predicting metabolic engineering targets that improve productivity in microbial cell factories [15–17]. For example, a kinetic model was used to identify limonene synthase as a key metabolic flux-controlling enzyme for limonene biosynthesis in the cyanobacterium *Synechococcus elongatus*, resulting in improved limonene production by increasing the enzyme level through genetic engineering [17]. In systems medicine, kinetic models and MCA approaches have been used to identify putative drug targets in biochemical networks [18, 19].

Kinetic models are also uniquely suitable for simulating time-dependent behavior that cannot be captured by steady state models, thereby possibly providing optimization strategies for industrial bioprocesses (Figure 5.7b). Ideally, also the extracellular conditions in which the cells are cultured should be modeled. An example of this is the use of a kinetic model of Chinese hamster ovary (CHO) cells to simulate a fed-batch cultivation, which was able to capture time-dependent extracellular metabolite concentrations and the effect of various process variables on antibody production. By simulating over 9000 combinations of process variables, e.g. cell density at inoculation; day at which the culture was shifted to a lower temperature (a strategy to helps to balance cell growth and protein productivity); how many days after inoculation the temperature shift took place; and knockdowns of metabolic enzymes, the researchers we able to optimize antibody production by modifying some of the process parameters [20].

With the increasing ease of high-throughput data generation, kinetic models can provide a framework for the analysis of omics data, which are large-scale measurements of cellular components, e.g. protein, mRNA, and metabolite concentrations. Notably, some of the omics data have direct connections to

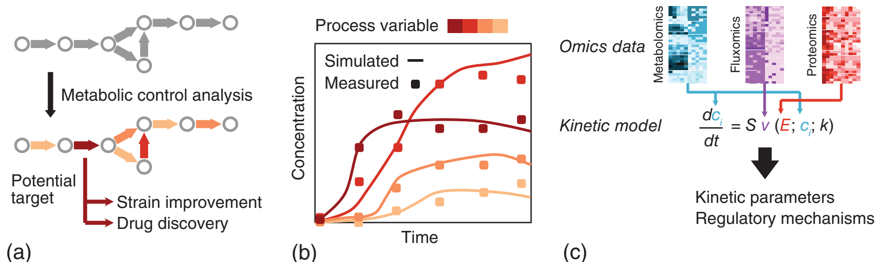


Figure 5.7 Applications of kinetic models of metabolism. (a) In metabolic control analysis, kinetic models are able to calculate the control of each enzyme on the flux through a network. Enzymes with high control are potential targets for strain improvement and drug discovery. (b) Kinetic models are able to simulate dynamic processes. The example illustrates measured and simulated time-course concentrations of a metabolite in response to different levels of a process variable. (c) Kinetic models are suitable for analyzing omics data as rate expressions contain cellular components, i.e. metabolite (c_i); enzyme (E); as well as metabolic rate ($\boldsymbol{v}$), which correspond to metabolomics, proteomics, and fluxomics data, respectively. Integrated analysis of multiomics data can estimate enzyme kinetics and regulatory mechanisms. The equation corresponds to Eq. (5.7).

the parameters and variables of kinetic models (Figure 5.7c). When analyzing single-type omics data, kinetic models can give deeper understanding that cannot be obtained from the data alone. For example, researchers found that estimating kinetic parameters in personalized kinetic models of erythrocyte metabolism are better representations of the individual's genotype than based on metabolomics data [21]. Furthermore, kinetic models enable integration of multiple different types of omics data (multiomics), thereby bringing systematic insights on metabolism and regulation. For example, researchers used a kinetic modeling framework to perform integrated analysis of proteomics, metabolomics, and fluxomics data and identified that substrate concentrations are the strongest drivers of metabolic fluxes [22].

5.9 Perspectives

In essence, kinetic models can be regarded as stoichiometric models that are augmented with rate expressions that account for kinetic information, which can render this model approach advantageous over *classical* stoichiometric GEMs in particular aspects. Kinetic models can quantitatively simulate metabolite concentrations which cannot be achieved by stoichiometric GEMs, while kinetic models are also more suitable to simulate dynamic responses where GEMs rely on the steady state assumption. Meanwhile, the augmentation with kinetic information also has drawbacks. Simulations with kinetic models are relatively computationally expensive as nonlinear optimization problems need to be solved. In addition, the construction of kinetic models requires copious experimental knowledge on rate expressions and kinetic parameters, and while these can be assumed, estimated, or simplified, this introduces more uncertainties.

Considering both the pros and cons of kinetic models and stoichiometric GEMs, researchers have proposed hybrid modeling approaches that utilize simplified rate expressions where enzyme turnover rates are set as constraints on fluxes through stoichiometric networks [23]. Despite these advanced modeling approaches, challenges remain including unknown parameters, which for such models are purely turnover rates. This has raised calls for measurements of the kinetome [24], i.e. all enzyme turnover rates in a cell, information that is also required in so-called proteome-constrained models (Chapter 4).

With continuing progress in the generation and utilization of kinetic models, two directions can be envisioned for future advances. First, the quality of kinetic models can be greatly improved. While pathway and network stoichiometries can now largely be readily obtained from genome annotation and reaction databases, the quality of kinetic models instead depends greatly on the availability of knowledge on rate expressions, kinetic parameters, and concentrations of cellular components such as metabolites. Although the kinetics of many enzymes have been characterized, the obtained parameters are mostly *in vitro*, which are not necessarily representative of their *in vivo* behavior [25]. While inference from measured fluxes and omics data is promising, the amount of absolutely quantified omics data is still sparse, and it remains challenging to measure metabolites from different subcellular compartmentation in eukaryotes [26]. All of these points currently hamper improvement of kinetic models, but likewise indicate where significant breakthroughs can be made. In addition to the approaches of parameter estimation and inference as mentioned above, it is also anticipated that machine learning algorithms will be able to effectively determine parameter values in the near future, as such approaches have already shown applications in various biological studies [27].

The second direction for further advancement in kinetic models would be the efficacious implementation of kinetic models on a genome scale. A number of recent efforts have been made to build genome-scale kinetic models, where the kinetic model *k-ecoli457* covers major parts of *E. coli* metabolism and has shown even better predictive power than constraint-based approaches in terms of predicting yields of many products [4]. There are, however, several obstacles on the way to genome-scale kinetic models. In addition to the high computational cost of model simulations, again the lack of large-scale data and missing rate expressions for individual reactions needs to be overcome for model construction and calibration. As several approaches are taken to address these obstacles [28–30], it is anticipated that genome-scale kinetic models have promise for wide application in the future.

References

1 Ainsworth, S. (1977). Michaelis–Menten kinetics. In: *Steady-State Enzyme Kinetics*, 43–73. London: Macmillan Education UK.

2 Stalidzans, E., Seiman, A., Peebo, K. et al. (2018). Model-based metabolism design: constraints for kinetic and stoichiometric models. *Biochem. Soc. Trans.* 46: 261–267.

3 Garfinkel, D., Frenkel, R.A., and Garfinkel, L. (1968). Simulation of the detailed regulation of glycolysis in a heart supernatant preparation. *Comput. Biomed. Res.* 2: 68–91.

4 Khodayari, A. and Maranas, C.D. (2016). A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat. Commun.* 7: 13806.

5 St. John, P.C., Strutz, J., Broadbelt, L.J. et al. (2019). Bayesian inference of metabolic kinetics from genome-scale multiomics data. *PLoS Comput. Biol.* e1007424: 15.

6 Visser, D., Schmid, J.W., Mauch, K. et al. (2004). Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab. Eng.* 6: 378–390.

7 Savageau, M.A. (1976). *Biochemical Systems Analysis*. Reading, MA: Addison-Wesley Pub. Co.

8 Tucker, W., Kutalik, Z., and Moulton, V. (2007). Estimating parameters for generalized mass action models using constraint propagation. *Math. Biosci.* 208: 607–620.

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

10 Davidi, D., Noor, E., Liebermeister, W. et al. (2016). Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro k(cat) measurements. *Proc. Natl. Acad. Sci. U S A* 113: 3401–3406.

11 van Eunen, K., Kiewiet, J.A.L., Westerhoff, H.V., and Bakker, B.M. (2012). Testing biochemistry revisited: how in vivo metabolism can be understood from in vitro enzyme kinetics. *PLoS Comput. Biol.* 8: e1002483.

12 Jeske, L., Placzek, S., Schomburg, I. et al. (2019). BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Res.* 47: 542–549.

13 Wittig, U., Rey, M., Weidemann, A. et al. (2017). SABIO-RK: an updated resource for manually curated biochemical reaction kinetics. *Nucleic Acids Res.* 46: D656–D660.

14 Smallbone, K., Messiha, H.L., Carroll, K.M. et al. (2013). A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. *FEBS Lett.* 587: 2832–2841.

15 Cintolesi, A., Clomburg, J.M., Rigou, V. et al. (2012). Quantitative analysis of the fermentative metabolism of glycerol in *Escherichia coli*. *Biotechnol. Bioeng.* 109: 187–198.

16 Andreozzi, S., Chakrabarti, A., Soh, K.C. et al. (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: 148–159.

17 Wang, X., Liu, W., Xin, C. et al. (2016). Enhanced limonene production in cyanobacteria reveals photosynthesis limitations. *Proc. Natl. Acad. Sci. U. S. A.* 113: 14225–14230.

18 Murabito, E., Smallbone, K., Swinton, J. et al. (2011). A probabilistic approach to identify putative drug targets in biochemical networks. *J. R. Soc. Interface* 8: 880–895.

19 Haanstra, J.R., Gerding, A., Dolga, A.M. et al. (2017). Targeting pathogen metabolism without collateral damage to the host. *Sci. Rep.* 7: 40406.

20 Nolan, R.P. and Lee, K. (2012). Dynamic model for CHO cell engineering. *J. Biotechnol.* 158: 24–33.

21 Bordbar, A., McCloskey, D., Zielinski, D.C. et al. (2015). Personalized whole-cell kinetic models of metabolism for discovery in genomics and pharmacodynamics. *Cell Syst.* 1: 283–292.

22 Hackett, S.R., Zanotelli, V.R.T., Xu, W. et al. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* 354: aaf2786. https://doi.org/10.1126/science.aaf2786.

23 Sanchez, B.J., Zhang, C., Nilsson, A. et al. (2017). Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* 13: 935.

24 Nilsson, A., Nielsen, J., and Palsson, B.O. (2017). Metabolic models of protein allocation call for the kinetome. *Cell Syst.* 5: 538–541.

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

26 Verhagen, K.J., van Gulik, W.M., and Wahl, S.A. (2020). Dynamics in redox metabolism, from stoichiometry towards kinetics. *Curr. Opin. Biotechnol.* 64: 116–123.

27 Kim, G.B., Kim, W.J., Kim, H.U., and Lee, S.Y. (2020). Machine learning applications in systems metabolic engineering. *Curr. Opin. Biotechnol.* 64: 1–9.

28 Tummler, K. and Klipp, E. (2018). 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.* 8: 1–6.

29 Saa, P.A. and Nielsen, L.K. (2017). Formulation, construction and analysis of kinetic models of metabolism: a review of modelling frameworks. *Biotechnol. Adv.* 35: 981–1003.

30 Srinivasan, S., Cluett, W.R., and Mahadevan, R. (2015). Constructing kinetic models of metabolism at genome-scales: a review. *Biotechnol. J.* 10: 1345–1359.