Chapter 2

Biochemical Reaction Networks

This chapter is a basic introduction to chemical reactions and chemical species. Different ways of quantifying the abundance of molecules lead to the notions of concentration. Similarly, a deterministic quantification of how fast a reaction proceeds in time leads to notions such as reaction rate and rate constant. Representation of biochemical reaction schemes is reviewed. Deterministic description of reaction networks in terms of reaction rates is described.

2.1 The Notion of a Chemical Reaction

Molecules inside the cell undergo various transformations. For example, a molecule can transform from one kind to another, two molecules of the same or different kinds can combine to form another molecule of a third kind, and so on. At the basic level these transformations are known as chemical reactions. In the context of chemical reactions, a molecule is said to be (an instance) of a certain species. Similarly, a chemical reaction is said to be (an instance) of a certain channel. The chemical species are denoted by roman uppercase letters. A single molecule of a species A is referred to as an A-molecule. The chemical reaction is written schematically as an arrow with reactants on the left and products on the right. Thus an A-molecule could transform to a B-molecule:

$$\mathrm{A} \rightarrow \mathrm{B},$$

a conversion or modification or isomerization. An A-molecule could associate with a B-molecule to form a non-covalently-bound complex:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C},$$

an *association* or synthesis. The complex C-molecule could dissociate into an A- and a B-molecule:

$$\mathrm{C} \rightarrow \mathrm{A} + \mathrm{B},$$

a dissociation or decomposition. A species that is not of interest to us (e.g., because its abundance does not change over time) is represented by the symbol

M. Ullah and O. Wolkenhauer, *Stochastic Approaches for Systems Biology*,
DOI 10.1007/978-1-4614-0478-1_2, © Springer Science+Business Media, LLC 2011

$\varnothing$ and referred to as the "null species." So the reaction

$$\mathrm{A} \to \varnothing$$

represents the degradation of an A-molecule to a form not of interest to us. Similarly, the production of a B-molecule is written as

$$\varnothing \to \mathrm{B}$$

when the reactants are disregarded. These reactions are said to be *elementary* and irreversible; elementary in the sense that each one takes one basic step (association, dissociation, conversion) to complete and irreversible because the change is only in one direction. They never exist in isolation, but always in combination with each other. So, what we usually describe as a chemical reaction can always be broken down into a mechanism that consists of combinations of these three elementary processes. For example, the probable mechanism of the chemical reaction

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C}$$

would be

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB} \to \mathrm{C},$$

where C is a covalent modification of AB. Each half ($\rightharpoonup$ or $\leftharpoondown$) of the double arrow ($\rightleftharpoons$) denotes one of the elementary reactions. Thermodynamically, all chemical reactions are reversible and consist of a forward reaction and a reverse reaction. Thus when we write an irreversible reaction, it will either represent the forward or backward step of a reversible reaction, or a simplification (i.e., approximation) of a reversible reaction by an irreversible one.

2.2 Networks of Reactions

Imagine molecules of s chemical species homogeneously distributed in a compartment of constant volume V at thermal equilibrium and interacting through r irreversible reaction channels. A reaction channel either is elementary or may represent a simplification of multiple elementary steps into a single step. Any reversible (bidirectional) reaction can be listed as two irreversible reactions. We symbolize the ith species with X_i and the jth reaction channel with R_j. The abundance of X_i present in the system at time t can be described by the copy number $N_i(t)$. The total copy number n^{tot} of all species indicates how large the system is. Since a large/small value of n^{tot} usually implies a large/small volume, the volume V can also indicate the size of the system. Any such parameter can be used as the *system size* and is usually denoted by

Ω. The copy number is usually divided by the system size, and the quantity thus obtained,

$$X_i(t) = \frac{N_i(t)}{\Omega},$$

is referred to as the concentration. The choice of the system size Ω depends on the kind of concentration one would like to define.

Molar Concentrations: For molar concentrations, in units $\mathrm{M} \equiv {}^{\mathrm{mol}}/_{\mathrm{L}}$, the system size is chosen as $\Omega = N_{\mathrm{A}} V$, where Avogadro's constant

$$N_{\mathrm{A}} = 6.022 \times 10^{23}\,\mathrm{mol}^{-1}$$

(correct to four significant digits) is the number of molecules (or any elementary entities) in one mole. If the volume is given in liters (L) and concentration in molar (M), then the unit of system size Ω is $\mathrm{mol}^{-1} \times \mathrm{L} = \mathrm{M}^{-1}$. The molar unit (M) is too large for very small concentrations, which are better specified in smaller units including nanomolar (nM), micromolar (μM), and millimolar (mM). Suppose the proteins in a cell of volume $V = 30\,\mathrm{fL}$ are measured in nanomolar $(\mathrm{nM})^{-1}$; then the computation of the system size proceeds like this:

$$\Omega = N_{\mathrm{A}} V = \left(6.022 \times 10^{14}\,(\mathrm{n\,mol})^{-1}\right) \times \left(3 \times 10^{-14}\,\mathrm{L}\right) \approx 18\,(\mathrm{nM})^{-1}\,.$$

Sometimes, the volume is chosen so that $\Omega = 1\,(\mathrm{nM})^{-1}$ for the resulting convenience that each nanomolar concentration is numerically equal to the corresponding copy number.

Relative concentrations: For relative concentrations, the system size is chosen to give dimensionless concentrations. One simpler way to obtain relative concentrations is by choosing $\Omega = n^{\mathrm{tot}}$, so that each concentration is just a fraction of two copy numbers. Take the isomerization reaction as an example whereby proteins are converted back and forth between the unmodified form U and the modified form W such that the total number n^{tot} of protein molecules remains constant. The relative concentrations in this example are the fractions

$$X_{\mathrm{U}}(t) = \frac{N_{\mathrm{U}}(t)}{n^{\mathrm{tot}}} \quad \text{and} \quad X_{\mathrm{W}}(t) = \frac{n^{\mathrm{tot}} - N_{\mathrm{U}}(t)}{n^{\mathrm{tot}}}$$

of proteins in the inactive and active form, respectively. For some systems it is more appropriate to introduce a different scaling parameter Ω_i for each component i if the copy numbers N_i differ in magnitude to keep X_i of the

same order $\mathcal{O}(1)$. That can be obtained by defining relative concentration as

$$X_i = \frac{N_i}{C_i \Omega},$$

that is, the concentration N_i/Ω divided by a characteristic concentration C_i. In that case, each scaling parameter can be expressed as $\Omega_i = C_i \Omega$. This will be of concern to us in the following chapter. In this chapter, we stick to the simpler case.

The reaction channel R_j will be represented by the general scheme

$$\underline{S}_{1j}\mathrm{X}_1 + \cdots + \underline{S}_{sj}\mathrm{X}_s \xrightarrow{k_j} \bar{S}_{1j}\mathrm{X}_1 + \cdots + \bar{S}_{sj}\mathrm{X}_s\,. \tag{2.1}$$

The participation of individual species in the reaction is indicated by *stoichiometries*, or *stoichiometric coefficients*, written beside them. Thus, the coefficient $\underline{S}_{ij}$ (on the left) represents the participation of X_i as a reactant and $\bar{S}_{ij}$ (on the right) is the corresponding participation as a product. The rate constant, or coefficient, k_j, written over the reaction arrow informs us about the assumed reaction kinetics, and will be explained later. The coefficient will be omitted when we do not want to attach any assumed reaction kinetics to the above reaction scheme. The progress of channel R_j is quantified in this text by the *reaction count* $Z_j(t)$, defined as the number of occurrences of R_j during the time interval $[0, t]$. One occurrence of R_j changes the copy number of X_i by $S_{ij} = \bar{S}_{ij} - \underline{S}_{ij}$, the (i, j)th element of the *stoichiometry matrix* S. During the time interval $[0, t]$, the change in the copy number of X_i contributed by R_j is thus $S_{ij} Z_j(t)$. The total change in the copy number is the sum of contributions from all reactions:

$$N_i(t) = N_i(0) + \sum_{j=1}^{r} S_{ij} Z_j(t)\,. \tag{2.2}$$

Thus changes in copy numbers are determined by stoichiometries and reaction counts. We need to caution the reader against a potential confusion between the term *reaction count* and a similar term *reaction extent*. Since the copy numbers appearing in the above equation are in units of molecules, we can also interpret the reaction count $Z_j(t)$ as number of molecules of a hypothetical substance in terms of which the other copy numbers are expressed. Dividing the above equation by N_A will change the measurements from molecules to moles, and the reaction count is replaced by the reaction extent $Z_j(t)/N_\mathrm{A}$. Following the usual vector notation, we write $N(t)$ for the s-vector of copy numbers, $X(t)$ for the s-vector of concentrations, and $Z(t)$ for the r-vector of reaction counts. The above conservation relation can be written in vector

notation:

$$N(t) = N(0) + S\,Z(t)\,. \tag{2.3}$$

Dividing by Ω gives the corresponding relation in concentrations:

$$X(t) = X(0) + \frac{S\,Z(t)}{\Omega}\,. \tag{2.4}$$

The quantity $Z_j(t)/\Omega$ is referred to as the *degree of advancement* of the reaction channel and replaces the role of reaction count in converting the progress of reaction to species concentration.

The copy number $N(t)$, the concentration $X(t)$, and the reaction count $Z(t)$ are alternative ways to describe our system. Description in terms of these *macroscopic variables* is done in the hope that they approximately satisfy an autonomous set of deterministic (differential or difference) equations. Because of the ease of analysis, differential equations are always preferred over the difference equation. However, the reactions are discrete events in time, which means that the copy numbers do not vary continuously with time. That would require the adoption of difference equations. The situation is made even more complicated by two problems. Firstly, the occurrence time of a reaction is a random quantity because it is determined by a large number of microscopic factors (e.g., positions and momenta of the molecules involved). The second problem arises when more than one type of reaction can occur. The type of reaction to occur is also a random quantity for the same reasons mentioned above. Therefore, the deterministic description needs a few simplifying assumptions. Alternatively, the macroscopic variables are formulated as stochastic processes. Such a stochastic description in terms of macroscopic variables is *mesoscopic*.

Throughout this text, we will use a couple of academic examples. They are chosen to demonstrate different ideas and methods in the discussion. For further examples of simple biochemical networks and a discussion of their relevance to molecular and cell biology, the reader is referred to [157].

Example 2.1 (Standard modification) Consider a protein that can exist in two different conformations, or forms, an *unmodified* form U and a *modified* form W. The protein changes between the two forms by the reversible isomerization reaction

$$\mathrm{U} \underset{k_u}{\overset{k_w}{\rightleftharpoons}} \mathrm{W} \tag{2.5}$$

composed of a modification (forward) channel with rate constant k_u and a demodification (reverse) channel with rate constant k_w. The reaction scheme (2.5) also represents the opening and closing of an ion channel and similar systems with two-state conformational change. Since the two reactions are

not influenced by any external catalyst (e.g., an enzyme), the scheme (2.5) will be referred to as the *standard modification*. This example was used in the introductory chapter to illustrate ideas of identifiability and species extinction or depletion.

Example 2.2 (Heterodimerization) Consider the reversible heterodimerization

$$\mathrm{X}_1 + \mathrm{X}_2 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{X}_3\,. \tag{2.6}$$

Here the forward reaction is the association of a receptor X_1 and a ligand X_2 to form a heterodimer (complex) X_3. The backward reaction is the dissociation of the heterodimer back into the two monomers. The parameters k_1 and k_2 are the respective association and dissociation rate constants. This example is the simplest one with a bimolecular reaction.

Example 2.3 (Lotka–Volterra model) Consider the process whereby a reactant A, replenished at a constant rate, is converted into a product B that is removed at a constant rate. The reaction will reach a steady state but cannot reach a chemical equilibrium. Suppose the process can be decomposed into three elementary steps:

$$\begin{aligned} \mathrm{X}_1 + \mathrm{A} &\longrightarrow 2\mathrm{X}_1, \\ \mathrm{X}_1 + \mathrm{X}_2 &\longrightarrow 2\mathrm{X}_2, \\ \mathrm{X}_2 &\longrightarrow \mathrm{B}\,. \end{aligned}$$

The first two reactions are examples of an autocatalytic reaction: the first one is catalyzed by the reactant X_1 and the second by the reactant X_2. This simple reaction scheme was proposed as a simple mechanism of oscillating reactions [67, 94]. Although the scheme illustrates how oscillation may occur, known oscillating chemical reactions have mechanisms different from the above. For an in-depth treatment of biochemical oscillations, the reader is referred to [42, 59, 106, Chapter 9]. This type of process is found in fields other than chemistry; they were investigated in the context of population biology by Lotka [94] and Volterra [164]. Due to the frequent appearance of this latter context in the literature, we rewrite the above reaction scheme as

$$\left.\begin{aligned} \mathrm{X}_1 + \mathrm{A} &\xrightarrow{k_1} 2\mathrm{X}_1, \\ \mathrm{X}_1 + \mathrm{X}_2 &\xrightarrow{k_2} 2\mathrm{X}_2, \\ \mathrm{X}_2 &\xrightarrow{k_3} \varnothing, \end{aligned}\right\} \tag{2.7}$$

as a system of two interacting species: X_1 (the prey) and X_2 (the predator). The food (substrate) A is available for X_1, which reproduces, with rate

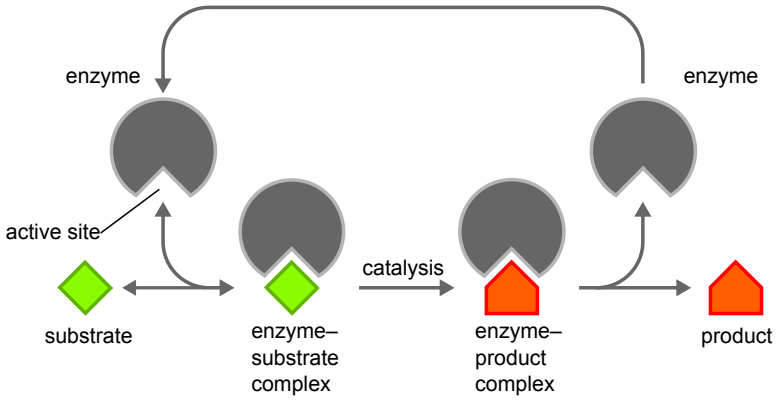

Figure 2.1 Enzyme-catalyzed conversion of a substrate to a product. The enzyme binds to the substrate to make its conversion to product energetically favorable. Figure based on an illustration in Alberts et al. [4].

coefficient k_1, after consuming one unit of A. An encounter between the two species, with rate coefficient k_2, results in the disappearance of X_1 and the replication of X_2. This is the only way X_1 dies (degrades), whereas X_2 has a natural death (degradation) with rate coefficient k_3. The food A is assumed to be constantly replenished so that the copy number n_A remains constant. This example serves the purpose of a simple system containing a bimolecular reaction and the resulting influence of (co)variance on the mean (Chapter 6).

Example 2.4 (Enzyme kinetic reaction) In biological systems, the conversion of a substrate to a product may not be a thermodynamically feasible reaction. However, specialized proteins called enzymes ease the job by binding to the substrate and lowering the activation energy required for conversion to the product, as depicted in Figure 2.1. Represented in reaction notation,

$$\mathrm{E} + \mathrm{S} \longrightarrow \mathrm{E} + \mathrm{P},$$

the enzymatic reaction is thought to be accomplished in three elementary steps:

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_3} \mathrm{E} + \mathrm{P}\,. \tag{2.8}$$

Here the enzyme E catalyzes a substrate S into a product P that involves an intermediary complex ES. Note that we have not placed any rate constant over the arrow in the original reaction because we do not specify any assumed kinetics in that notation. Later we will learn that it is possible to approximate

the three elementary reactions by a single reaction,

$$\mathrm{S} \xrightarrow{k_{\mathrm{eff}}} \mathrm{P},$$

with an effective rate coefficient k_{eff} that represents the assumed approximate kinetics. Intuitively, k_{eff} will be a function of the enzyme abundance. We include this example because this type of reaction appears frequently in the literature. It also serves the purpose of a simple system containing a bimolecular reaction and allows demonstration of how mass conservation leads to a simplified model.

Example 2.5 (Schlögl model) The Schlögl model is an autocatalytic, trimolecular reaction scheme, first proposed by Schlögl [137]:

$$\mathrm{A} + 2\mathrm{X} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3\mathrm{X}, \quad \mathrm{B} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{X}\,. \tag{2.9}$$

Here the concentrations of A and B are kept constant (buffered). This example, mentioned in the introduction, serves to illustrate the need for a stochastic approach to model systems with bistability and the associated behavior known as "stochastic switching."

Example 2.6 (Stochastic focusing) This example was first described in [121] to demonstrate a behavior phenomenon known as "stochastic focusing." The branched reaction network comprised the following reaction channels:

$$\varnothing \underset{k_d}{\overset{k_s}{\rightleftharpoons}} \mathrm{S}, \quad \varnothing \underset{k_a X_{\mathrm{S}}}{\overset{k_i}{\rightleftharpoons}} \mathrm{I} \xrightarrow{k_p} \mathrm{P} \xrightarrow{1} \varnothing\,. \tag{2.10}$$

Here the product P results from the irreversible isomerization of its precursor I, an intermediary chemical species. This isomerization is inhibited by a signaling chemical species S that is synthesized and degraded by independent mechanisms.

Example 2.7 (Gene regulation) As pointed out earlier, oscillating chemical reactions have mechanisms different from the simple and intuitive Lotka–Volterra scheme. Those familiar with dynamical systems theory will recall that such a kinetic system can oscillate only if both activation and inhibition are present in the form of a feedback loop. Such feedback loops exist in gene expression, where the protein product serves as a transcription factor and represses transcription. A simplified regulatory mechanism is illustrated in Figure 2.2. The protein product from gene expression binds to a regulatory region on the DNA and represses transcription. The regulatory mechanism

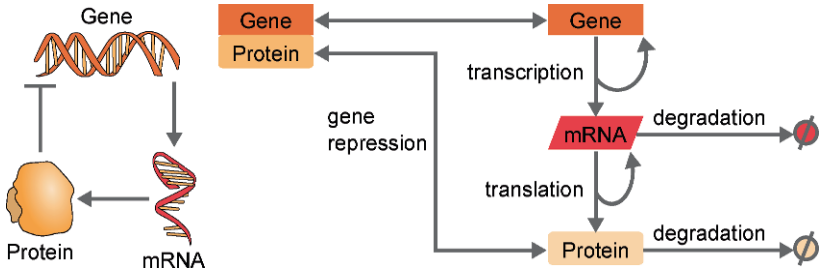

Figure 2.2 Gene regulation: a simplified model. *Left*: cartoon representation. *Right*: reaction pathways.

is simplified by not showing the contributions of RNA polymerase and any cofactors. The reaction scheme for the system is

$$\left.\begin{aligned} \mathrm{G} &\xrightarrow{k_m} \mathrm{G}+\mathrm{M} && \text{(transcription)},\\ \mathrm{M} &\xrightarrow{k_p} \mathrm{M}+\mathrm{P} && \text{(translation)},\\ \mathrm{G}+\mathrm{P} &\underset{k_u}{\overset{k_b}{\rightleftharpoons}} \mathrm{GP} && \text{(binding/unbinding)},\\ \mathrm{M} \xrightarrow{k_m^-} \varnothing,&\quad \mathrm{P} \xrightarrow{k_p^-} \varnothing && \text{(degradation)}, \end{aligned}\right\} \tag{2.11}$$

where the gene G is transcribed to the mRNA M with rate constant k_m, the mRNA is translated to the protein P with rate constant k_p, and the protein binds to (and represses) the gene with rate constant k_b and unbinds back with rate constant k_u. The mRNA and protein are degraded with respective rate constants k_m^- and k_p^-.

Synthetic gene regulation: The above idea of a simple feedback loop has motivated several researchers to construct such feedback transcriptional regulatory networks in living cells [11]. These investigators found that an increased delay in the feedback loop increases the dynamic complexity of the synthetic transcription system. A feedback loop with one repressor protein constructed by Becskei and Serrano [12] exhibited on and off transitions. Another loop with two repressor proteins, constructed by Gardner, Cantor, and Collins [48], manifested bistability in the on and off states. Yet another loop with three repressor proteins, constructed and termed "repressilator" by Elowitz and Leibler [40], exhibited oscillations.

Repressilator: The repressilator is a milestone of synthetic biology because its shows that gene regulatory networks can be designed and implemented to

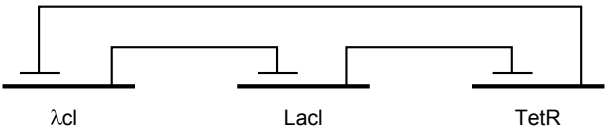

Figure 2.3 The repressilator gene regulatory network.

perform a novel desired function [40]. The repressilator consists of three genes, λcl, LacI, and TelR, connected in a feedback loop. As depicted in Figure 2.3, each gene product represses the next gene in the loop, and is repressed by the previous gene. In addition, not shown in the figure, green fluorescent protein is used as a reporter so that the behavior of the network can be observed using fluorescence microscopy.

Cell cycle: The cycle through which cells grow and duplicate their DNA before eventually dividing into two daughter cells is of central importance to the realization of higher levels of biological organization. Underlying the cell cycle and its regulation are complex mechanisms, realized through large reaction networks. Due to its complexity, the cell cycle is investigated as a case study in Chapter 7.

2.3 Deterministic Description

Suppose that reactions occur so frequently that the reaction count $Z(t)$ can be approximated by a continuous quantity $z(t)$. This assumption requires that a large number of reactant molecules be freely available (no crowding) in a large volume so that they can react easily. It also requires that the energy and orientation of reactant molecules favor the reaction, a fact summarized in a rate constant. Large numbers of molecules also mean that a change resulting from a single occurrence of a reaction is relatively small. This means that the copy number $N(t)$ can be approximated by a continuous quantity $n(t)$. The concentration $X(t)$ is similarly approximated by a continuous quantity $x(t)$. In a deterministic description, equations (2.3) and (2.4) respectively translate to

$$n(t) = n(0) + S\, z(t) \tag{2.12}$$

and

$$x(t) = x(0) + \frac{S\, z(t)}{\Omega}\,. \tag{2.13}$$

Taking the time derivatives gives us the net chemical fluxes:

$$\dot{n}(t) = S\, \dot{z}(t), \quad \dot{x}(t) = S\, \frac{\dot{z}(t)}{\Omega}\,. \tag{2.14}$$

Here the time derivative $\dot{x}$ is the net concentration flux and $\dot{n}$ is the net copy-number flux. Note that our usage of the term "chemical flux" differs from IUPAC[1], which defines it in terms of moles. The above equations are useful only if a relationship between the time derivative on the right and the abundance variable ($n(t)$ or $x(t)$) is established. Suppose a relation can be mathematically represented as

$$\dot{z} = \hat{v}(n) = \Omega v(x), \tag{2.15}$$

where the vectors $\hat{v}(x)$ and $v(x)$ are referred to here as the *conversion rate* and the *reaction rate*, respectively. The conversion rate is here defined as reaction count per unit time, a slight difference with the standard definition in [1] as the time derivative $\dot{z}/N_{\mathrm{A}}$ of the extent of reaction. The reaction rate is defined as reaction count per unit time divided by the system size. The notation $v\left(x(t)\right)$ is based on the assumption that the reaction rate depends only on the concentrations of the reactants. This is a realistic assumption in many reactions at constant temperature. In general, the reaction rate can depend on temperature, pressure, and the concentrations or partial pressures of the substances in the system.

The functional form $v_j(\cdot)$ of the rate of R_j is called the *rate law* (or kinetic law), which is a result of the modeling assumptions about the particular reaction channels. It is only after specifying a rate law that the above ODEs can characterize a particular biochemical reaction network. Without that specification, the above ODEs only represent a consistency condition imposed by mass (or substance) conservation of reactants and products. Incorporating the rate law specification (2.15) into the ODEs (2.14) leads to the deterministic *chemical kinetic equations*

$$\dot{n}(t) = S\,\hat{v}\left(n(t)\right), \quad \dot{x}(t) = S\,v\left(x(t)\right)\,. \tag{2.16}$$

There is a large class of chemical reactions in which the reaction rate is proportional to the concentration of each reactant raised to some power:[1]

$$v_j(x) = k_j \prod_{i=1}^{s} x_i^{g_{ij}}, \quad \hat{v}_j(n) = \hat{k}_j \prod_{i=1}^{s} n_i^{g_{ij}}, \tag{2.17}$$

which is called *a rate law with definite orders* [102]. The rate constant k_j summarizes factors such as activation energy and proper orientation of the reactant molecules for an encounter leading to the reaction. The rate constant k_j can be interpreted as the factor of the reaction rate that does not

[1] Since 0^0 is undefined, the product $\prod_{i=1}^{s}$ must exclude i for which both x_i and g_{ij} are zero.

depend on reactant concentrations. The conversion rate constant $\hat{k}_j$ has a similar interpretation as the factor of the extensive reaction rate that does not depend on the reactant copy numbers. Recall that while the units of $\hat{k}_j$ are always $\sec^{-1}$, the units of k_j additionally depend on the units used for the concentration x. The exponent g_{ij} is the order with respect to the species X_i. The sum of orders for a particular reaction channel is the overall order. For elementary reactions, the orders g_{ij} are the same as the reactant stoichiometries $\underline{S}_{ij}$:

$$v_j(x) = k_j \prod_{i=1}^{s} x_i^{S_{ij}}, \quad \hat{v}_j(n) = \hat{k}_j \prod_{i=1}^{s} n_i^{S_{ij}} . \tag{2.18}$$

This rate law is called *mass-action kinetics* [66] and is justified by collision theory and transition state theory [71, 102, 171]. The mass-action kinetics should not be confused with the closely related law of mass action, which is obtained by equating the forward and backward reaction rates (according to the above rate law) of a reversible reaction. Reactions that cannot be described by rate laws like (2.17) are said to have *no definite order*. For such a reaction, the rate law depends on the assumptions involved in the approximation of the constituent reaction channels. Examples of such rate laws include Michaelis–Menten kinetics, Hill kinetics, and competitive inhibition [25, 43, 66]. A family tree of deterministic ODE models is sketched in Figure 2.4. The ODEs in their most general form are rarely used in systems biology. Equation (2.16) is the most common representation to describe the continuous changes in concentration $x(t)$ in terms of the network structure, encoded by the stoichiometry matrix S, and the network kinetics, encoded by the rate law $v(\cdot)$. Note that the kinetic parameters such as the rate constant k and the kinetic order g are incorporated in the rate law. Further variations emerge through an implicit assumption about the underlying biophysical environment in which reactions take place. Assuming basic mass-action-type kinetics, the kinetic order g_{ij} of the rate law will typically take the value 1 or 2 (dimerization). Further quasi-steady-state assumptions for intermediate complexes can simplify into Michaelis–Menten type kinetic models. The left branch allows for noninteger kinetic orders and takes two routes that depend on the semantics [161]. Simplified power-law models (e.g., S-Systems [163]) assume very little knowledge about the biophysical structure of the environment in which reactions take place. These models distinguish between positive and negative contributions (pos/neg kinetic orders) and different strengths of activation/inhibition. On the other hand, criticizing the assumption of a homogeneous and well mixed environment (underlying the right branch) leads to noninteger (but positive) kinetic orders. A detailed kinetic power-law model would thus arguably represent the biophysical environment more accurately

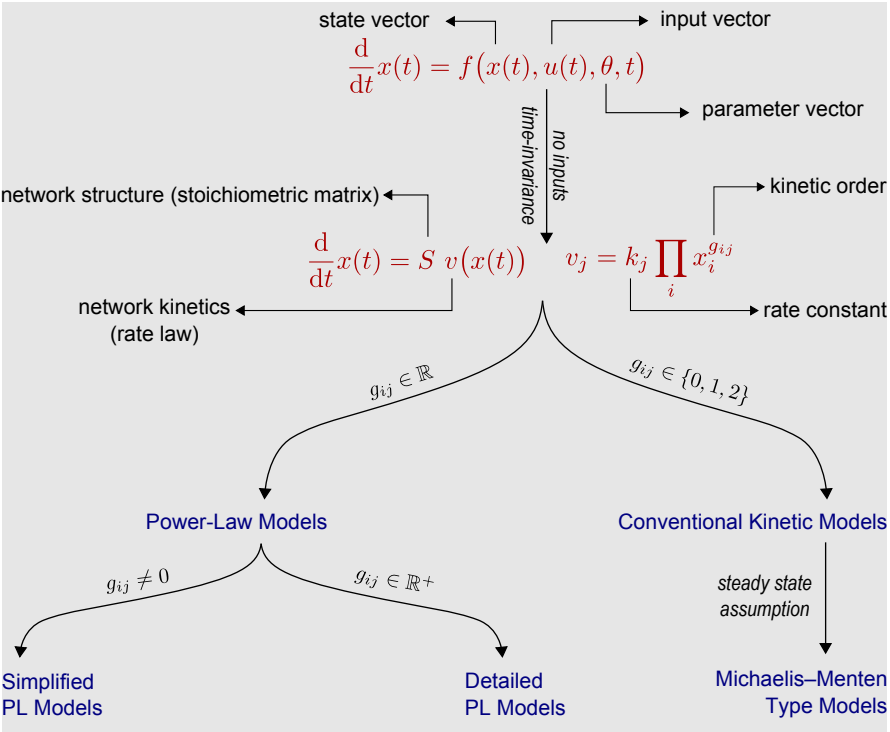

Figure 2.4 Family tree of deterministic ODE models. For chemical reaction networks, the general ODE formulation simplifies to a decomposition into the stoichiometry matrix (encoding the network structure) and the rate law (encoding the network kinetics). A large class of chemical reactions have a rate law with definite (kinetic) orders, of the form (2.17). Restricting and broadening the range of values of the kinetic order g_{ij} allows further classification.

than the conventional mass-action model. On the other hand the simplified power-law model admits a more phenomenological interpretation. A drawback of the power-law models is that of additional parameters, the kinetic orders, they introduce. The more parameters a model has, the more difficult it is to identify a unique set of parameter values from experimental time-course data.

Relationship between k and $\hat{k}$: We can combine the defining relationship (2.15) with the rate law (2.17) to get a relationship between the rate constant

Table 2.1 Relationship between the rate constant and the conversion rate constant for example reactions.

R_j	Relation
$\varnothing \xrightarrow{k_j} X$	$\hat{k}_j = \Omega k_j$
$X \xrightarrow{k_j} ?$	$\hat{k}_j = k_j$
$X_1 + X_2 \xrightarrow{k_j} ?$	$v = \frac{k_j}{\Omega}$
$2X \xrightarrow{k_j} ?$	$\hat{k}_j = \frac{k_j}{\Omega}$
$X_1 + X_2 + X_3 \xrightarrow{k_j} ?$	$\hat{k}_j = \frac{k_j}{\Omega^2}$
$X_1 + 2X_2 \xrightarrow{k_j} ?$	$\hat{k}_j = \frac{k_j}{\Omega^2}$

k and the conversion rate constant $\hat{k}$:

$$\hat{k}_j \prod_{i=1}^{s} n_i^{g_{ij}} = \hat{v}(n) = \Omega v(x) = \Omega k_j \prod_{i=1}^{s} x_i^{g_{ij}} .$$

Now invoke the defining relationship $n = \Omega x$ to obtain

$$\hat{k}_j = \frac{k_j}{\Omega^{K_j - 1}}, \tag{2.19}$$

where $K_j = \sum_{i=1}^{s} g_{ij}$, which, for elementary reactions, is simply $K_j = \sum_{i=1}^{s} S_{ij}$. The relationship for sample elementary reactions is illustrated in Table 2.1. The table suggests that the two types of rate constants are equal for monomolecular reactions.

Matlab implementation: To implement rate laws of the form (2.17) in Matlab [96], the standard Matlab data type *function handle* can be employed. We will need Matlab representations of our mathematical quantities. Let us collect the species concentrations x_i (at a certain time) in an $s \times 1$ column vector `x`, the reaction rate constants k_j in an $r \times 1$ column vector `k`, and the exponents g_{ij} (which equal S_{ij} for mass-action kinetics) of the rate law (2.17) in an $s \times r$ matrix `G`. Then the Matlab representation `v` of the rate law $v(\cdot)$ defined elementwise in (2.17) takes the following form:

M-code 2.1 `makeRateLaw`: implements rate law with definite orders (2.17).

```
function v = makeRateLaw(k,G)
r = size(G,2);
i0 = (G==0);
i = ~i0 & (G~=1);
v = @RateLaw;
    function vofx = RateLaw(x)
        X = repmat(x,1,r);
        X(i0) = 1;
        X(i) = X(i2.^G(i);
        vofx = k.*prod(X)';
    end
end
```

```
v = @(x) k.*prod(repmat(x,1,r).^G)';
```

where `r` is the Matlab representation of the number r of reaction channels. Here the function handle `v` stores the mathematical expression following `@(x)`. The standard Matlab notations `.*` and `.^` represent the elementwise operations multiplication and exponentiation. The compact code above may not be efficient in dealing with a large network of many species and reactions. Specifically, the exponentiation and multiplication are computationally demanding. To avoid these unnecessary computations, the code is replaced by Matlab function `makeRateLaw` in M-code 2.1. Here the output `v` returned by the main function `makeRateLaw` is a function handle to the *nested function* `RateLaw`. Note how exponentiation is avoided for the obvious cases $g_{ij} = 0$ and $g_{ij} = 1$. In general, a rate law may not be expressible in the form (2.17) and has to be written on a case-by-case basis. Once such function (or handle) has been written for the rate law, a Matlab representation of the chemical kinetic equations (2.16) can be written and numerically solved with the following piece of Matlab code:

```
dxdt = @(t,x) S*v(x); % concentration ODE
[tout,xout] = ode15s(dxdt, [0 tf], x0); % solution
```

Here `x0` is a column vector of initial concentrations and `tf` is the final (stop) time of the simulation. The solver `ode15s` returns the column vector `tout` of time points and the solution array `xout` with a row of concentrations for each time point.

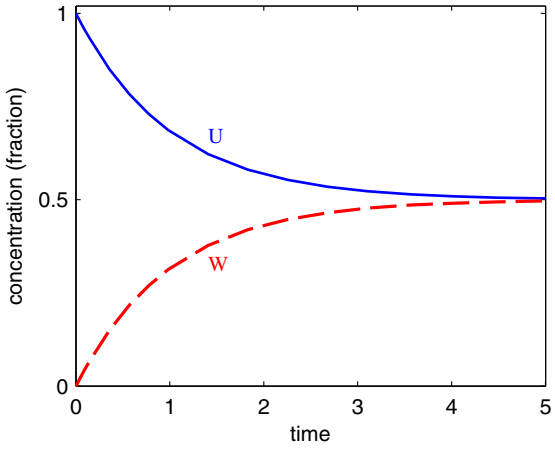

Figure 2.5 Time course of concentrations in the standard modification (2.20). Initially all molecules are assumed to be unmodified (U). The ordinate is the fraction of molecules in (un)modified form. Equilibrium is reached when the two fractions are equal. Both the rate constants were taken as $2\,\mathrm{sec}^{-1}$.

Example 2.8 (Standard modification) Consider the (de)modification of a protein between two forms by the reaction scheme (2.5). Suppose there are n^{tot} copies of this protein in a container, $n(t)$ of them being unmodified (in form U) at time t. The two reaction channels progress at the following conversion rates (listed on the right)

$$\begin{array}{l|l} \mathrm{U} \xrightarrow{k_w} \mathrm{W} & \hat{v}_w = k_w n \\ \mathrm{W} \xrightarrow{k_u} \mathrm{U} & \hat{v}_u = \left(n^{\mathrm{tot}} - n\right) k_u \end{array} \tag{2.20}$$

and their difference gives the rate equation

$$\dot{n} = -\hat{v}_w + \hat{v}_u = k_u n^{\mathrm{tot}} - (k_w + k_u) n\,.$$

The rate equation for the unmodified fraction $x = {}^{n}/_{n^{\mathrm{tot}}}$ of all proteins is then

$$\dot{x} = k_u - (k_w + k_u) x\,. \tag{2.21}$$

The Matlab implementation of this differential equation and its numerical solution will look like the following piece of code:

```
k = [2;2]; % rate constants
dxdt = @(t,n) k(2)-(k(1)+k(2))*x; % ODE
x0 = 1; % initial condition
```

```
[tout,xout] = ode15s(dxdt, [0 tf], x0); % solution
```

with the understanding that the Matlab workspace has values of variables `k`, `tf`, and `x0`, which correspond respectively to the rate constant $k = [k_w, k_u]$, the simulation stop time, and the initial fraction x^{init}. A typical time course is plotted in (2.21) wherein the fractions of molecules in the two forms are plotted against time. The above Matlab code can be rewritten in a way that lends itself to automatic code-writing. Toward that end, we write down the stoichiometry matrix S and the reaction rate vector v for this example:

$$S = \begin{bmatrix} -1 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} v_w \\ \\ v_u \end{bmatrix} = \begin{bmatrix} k_w x \\ \\ (1-x)k_u \end{bmatrix}.$$

With these two quantities available, the above Matlab code can be replaced by

```
S = [-1 1]; % stoichiometry matrix
k = [2;2]; % rate constants
v = @(x) [k(2)*x; (1-x)*k(1)]; % reaction rate
dxdt = @(t,x) S*v(x)'; % rate equation
x0 = 1; % initial condition
[tout,xout] = ode15s(dxdt, [0 tf], x0); % solution
```

Here the first line assigns values to (the array) `S`, which corresponds to the stoichiometry matrix S. The second line assigns an expression to the function handle `v`, which corresponds to the rate law $v(\cdot)$. The next line defines the function handle `dndt` to represent the system of ODEs in question. The last line calls an ODE solver to solve the problem and returns the output arrays `tout` of time points and `xout` of concentration values. It can be seen from the above Matlab code that all we need is a representation `S` (a Matlab matrix) of the stoichiometry matrix S and a representation `v` (a Matlab function handle) of the reaction rate law $v(\cdot)$.

For the remainder of the text, we will mostly specify such quantities with an understanding that the reader can translate that information into the corresponding Matlab code.

Chemical equilibrium: When the modification rate v_w (in the last example) is balanced by the demodification rate v_u, chemical equilibrium is said to have occurred. In other words, the reversible reaction equilibrates or reaches the steady state. The steady-state fraction x^{ss} is the value of x that makes the

time derivative in (2.21) zero, that is,

$$x^{\mathrm{ss}} = \frac{k_u}{k_w + k_u}\,.$$

Thus, in the steady state, a fraction $P_{\mathrm{U}} = {k_u}/{(k_u+k_w)}$ of proteins are in the unmodified form and a fraction $P_{\mathrm{W}} = {k_w}/{(k_u+k_w)}$ of them in the modified form. We can also say that a protein spends, on average, a fraction P_{W} of time in the modified form and a fraction P_{U} of time in the unmodified form. This interpretation proves very useful in reducing complicated reactions to single steps. Suppose the W form participates in another reaction $\mathrm{W} \xrightarrow{k_b} \mathrm{B}$ that occurs on a much slower time scale than two-state conformational changes between U and W. The overall complicated reaction

$$\mathrm{U} \underset{k_u}{\overset{k_w}{\rightleftharpoons}} \mathrm{W} \xrightarrow{k_b} \mathrm{B}$$

can be reduced to a single step $\varnothing \xrightarrow{k_b P_{\mathrm{W}}} \mathrm{B}$ under the fast equilibration assumption for the reversible reaction.

Example 2.9 (Heterodimerization) Recall the reversible heterodimerization depicted in the reaction scheme (2.6). Let $x_1(t)$, $x_2(t)$, and $x_3(t)$ denote the respective time-dependent molar concentrations of receptor X_1, ligand X_2, and heterodimer X_3. The reaction network has to satisfy two conservation relations:

$$x_1 + x_3 = q_1, \quad x_2 + x_3 = q_2, \tag{2.22}$$

where q_1 and q_2 are constants determined by the initial conditions. Using these to express x_1 and x_2 in terms of x_3, the system state can be represented by tracking only species X_3. The reaction rates according to the mass-action kinetics follow from (2.18) to be (each listed to the right of the corresponding reaction channel)

$$\begin{aligned} \mathrm{X}_1 + \mathrm{X}_2 &\xrightarrow{k_1} \mathrm{X}_3, \\ \mathrm{X}_3 &\xrightarrow{k_2} \mathrm{X}_1 + \mathrm{X}_2, \end{aligned} \qquad \left| \quad \begin{aligned} v_1 &= k_1\,(q_1 - x_3)\,(q_2 - x_3)\,, \\ v_2 &= k_2 x_3\,. \end{aligned} \right.$$

As far as X_3 is concerned, the stoichiometry matrix S and the reaction rate v

can be written as[2]

$$S = \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} k_1 (q_1 - x_3) (q_2 - x_3) \\ k_2 x_3 \end{bmatrix}.$$

The concentration $x_3(t)$ of the complex thus evolves according to

$$\frac{\mathrm{d}x_3}{\mathrm{d}t} = Sv = k_1 (q_1 - x_3) (q_2 - x_3) - k_2 x_3 .$$

Example 2.10 (Lotka–Volterra model) Revisit the mutual interactions (2.7) between the prey X_1 and the predator X_2. Let $n_1(t)$ and $n_2(t)$ denote the copy numbers of X_1 and X_2, respectively. The number n_A of the food items A is assumed to be unchanged by consumption during the time scale of interest. The reaction rates according to the mass-action kinetics follow from (2.18) to be (listed to the right)

$$\begin{aligned} \mathrm{X}_1 + \mathrm{A} &\xrightarrow{\hat{k}_1} 2\mathrm{X}_1, \\ \mathrm{X}_1 + \mathrm{X}_2 &\xrightarrow{\hat{k}_2} 2\mathrm{X}_2, \\ \mathrm{X}_2 &\xrightarrow{\hat{k}_3} \varnothing, \end{aligned} \quad \left| \quad \begin{aligned} \hat{v}_1 &= \hat{k}_1 n_\mathrm{A} n_1, \\ \hat{v}_2 &= \hat{k}_2 n_1 n_2, \\ \hat{v}_3 &= \hat{k}_3 n_2 . \end{aligned} \right.$$

As far as X_1 and X_2 are concerned, the stoichiometry matrix S and the reaction rate v can be written as

$$S = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \quad \hat{v} = \begin{bmatrix} \hat{k}_1 n_\mathrm{A} n_1 \\ \hat{k}_2 n_1 n_2 \\ \hat{k}_3 n_2 \end{bmatrix}.$$

The ODEs governing the time courses of $n_1(t)$ and $n_2(t)$ can be constructed from the vector Sv as

$$\left. \begin{aligned} \frac{\mathrm{d}n_1}{\mathrm{d}t} &= \left(\hat{k}_1 n_\mathrm{A} - \hat{k}_2 n_2 \right) n_1, \\ \frac{\mathrm{d}n_2}{\mathrm{d}t} &= \left(\hat{k}_2 n_1 - \hat{k}_3 \right) n_2 . \end{aligned} \right\} \tag{2.23}$$

[2]The full stoichiometry matrix for the 3-species 2-reaction scheme has three rows and two columns.

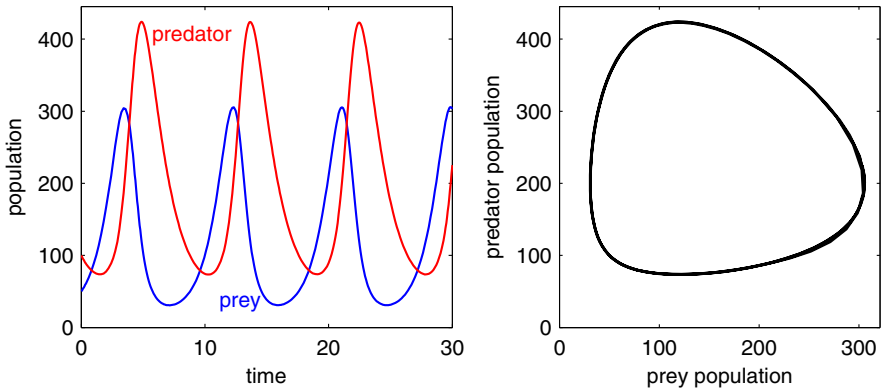

Figure 2.6 Deterministic simulation of the Lotka–Volterra model. *Left*: time course, *Right*: phase plot. Parameters (in $\sec^{-1}$): $\hat{k}_1 = 1$, $\hat{k}_2 = 0.005$, $\hat{k}_3 = 0.6$. Initial population is taken as 50 individuals of prey for 100 individuals of predator.

A numerical solution of the ODEs above is the time plot shown in Figure 2.6 side by side with the associated phase plot.

Example 2.11 (Enzyme kinetic reaction) For the enzyme kinetic reaction (2.8), we write $x_{\mathrm{E}}(t)$, $x_{\mathrm{S}}(t)$, $x_{\mathrm{ES}}(t)$, and $x_{\mathrm{P}}(t)$ for the respective time-dependent molar concentrations of E, S, ES, and P. The solution is usually assumed to respect two conservation laws:

$$x_{\mathrm{E}}(t) + x_{\mathrm{ES}}(t) = x_{\mathrm{E}}^{\mathrm{tot}} \quad \text{and} \quad x_{\mathrm{S}}(t) + x_{\mathrm{ES}}(t) + x_{\mathrm{P}}(t) = x_{\mathrm{S}}^{\mathrm{tot}}, \tag{2.24}$$

where $x_{\mathrm{E}}^{\mathrm{tot}}$ and $x_{\mathrm{S}}^{\mathrm{tot}}$ are, respectively, the total concentrations of the enzyme and substrate determined by the initial conditions. We can choose $x = (x_{\mathrm{S}}, x_{\mathrm{ES}})^T$ as the state vector sufficient to describe the system because the remaining two variables can be determined from the conservation relations above. The channelwise mass-action kinetic laws for the reaction scheme (2.8) are (list on the right):

$$\begin{aligned} \mathrm{E} + \mathrm{S} &\xrightarrow{k_1} \mathrm{ES}, \\ \mathrm{ES} &\xrightarrow{k_2} \mathrm{E} + \mathrm{S}, \\ \mathrm{ES} &\xrightarrow{k_3} \mathrm{E} + \mathrm{P}, \end{aligned} \qquad \left| \qquad \begin{aligned} v_1 &= \left(x_{\mathrm{E}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}}, \\ v_2 &= k_2 x_{\mathrm{ES}}, \\ v_3 &= k_3 x_{\mathrm{ES}}\,. \end{aligned} \right.$$

As far as S and ES are concerned, the stoichiometry matrix S and the reaction

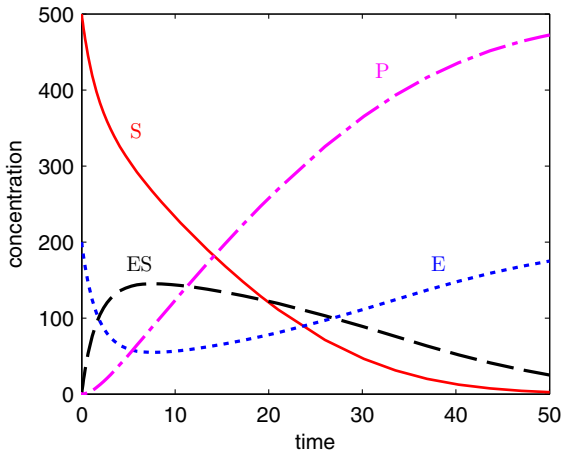

Figure 2.7 Deterministic time course of the enzyme kinetic reaction. Parameters: $k_1 = 10^{-3}\,(\mathrm{nM\,sec})^{-1}$, $k_2 = 10^{-4}\,\mathrm{sec}^{-1}$, $k_3 = 0.1\,\mathrm{sec}^{-1}$. Initial concentrations: $x_\mathrm{S} = 500\,\mathrm{nM}$, $x_\mathrm{E} = 200\,\mathrm{nM}$, $x_\mathrm{ES} = x_\mathrm{P} = 0\,\mathrm{nM}$.

rate v can be written as

$$S = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & -1 \end{bmatrix}, \quad v = \begin{bmatrix} (x_\mathrm{E}^\mathrm{tot} - x_\mathrm{ES})\, k_1 x_\mathrm{S} \\ k_2 x_\mathrm{ES} \\ k_3 x_\mathrm{ES} \end{bmatrix}.$$

The concentrations evolve according to the following set of nonlinear coupled ODEs (constructed from the vector Sv)

$$\begin{aligned} \frac{\mathrm{d}x_\mathrm{S}}{\mathrm{d}t} &= k_2 x_\mathrm{ES} - \left(x_\mathrm{E}^\mathrm{tot} - x_\mathrm{ES}\right) k_1 x_\mathrm{S}, \\ \frac{\mathrm{d}x_\mathrm{ES}}{\mathrm{d}t} &= \left(x_\mathrm{E}^\mathrm{tot} - x_\mathrm{ES}\right) k_1 x_\mathrm{S} - \left(k_2 + k_3\right) x_\mathrm{ES}\,. \end{aligned} \tag{2.25}$$

A numerical solution of the ODEs above is the time plot shown in Figure 2.7.

Michaelis–Menten kinetics: Following Michaelis and Menten [99] and Briggs and Haldane [19], in addition to the assumption of a constant total enzyme concentration $x_\mathrm{E}^\mathrm{tot}$, we make an additional assumption that the concentration x_ES of the substrate-bound enzyme changes little over time,

assuming a quasi steady state, that is,

$$\frac{\mathrm{d}x_{\mathrm{ES}}}{\mathrm{d}t} = \left(x_{\mathrm{E}}^{\mathrm{tot}} - x_{\mathrm{ES}}\right) k_1 x_{\mathrm{S}} - \left(k_2 + k_3\right) x_{\mathrm{ES}} \approx 0,$$

which is reasonable if the concentration x_{ES} of the substrate-bound enzyme changes much more slowly than those of the product and substrate. The above steady-state assumption can rearranged to form an algebraic expression for the steady-state concentration of the complex:

$$x_{\mathrm{ES}} = \frac{x_{\mathrm{E}}^{\mathrm{tot}} x_{\mathrm{S}}}{\left(\frac{k_2+k_3}{k_1}\right) + x_{\mathrm{S}}} = \frac{x_{\mathrm{E}}^{\mathrm{tot}} x_{\mathrm{S}}}{K_{\mathrm{M}} + x_{\mathrm{S}}},$$

where $K_{\mathrm{M}} = (k_2 + k_3)/k_1$ is known as the Michaelis–Menten constant. This can be combined with the fact that the product concentration x_{P} changes at the rate

$$\frac{\mathrm{d}x_{\mathrm{P}}}{\mathrm{d}t} = v_3 = k_3 x_{\mathrm{ES}} = \frac{k_3 x_{\mathrm{E}}^{\mathrm{tot}} x_{\mathrm{S}}}{K_{\mathrm{M}} + x_{\mathrm{S}}}\,.$$

Thus the 3-reaction enzymatic network has been reduced to a single reaction channel $\mathrm{S} \to \mathrm{P}$ with reaction rate

$$\frac{\mathrm{d}x_{\mathrm{P}}}{\mathrm{d}t} = -\frac{\mathrm{d}x_{\mathrm{S}}}{\mathrm{d}t} = v\left(x_{\mathrm{S}}\right) = \frac{v_{\max} x_{\mathrm{S}}}{K_{\mathrm{M}} + x_{\mathrm{S}}},$$

where $v_{\max} = k_3 x_{\mathrm{E}}^{\mathrm{tot}}$ is the initial (maximum) reaction rate.

Example 2.12 (Schlögl model) For the Schlögl reaction scheme (2.9), write x_{A} and x_{B} for the constant respective concentrations of chemicals A and B, and $x(t)$ for the time-dependent concentration of chemical X. The reaction rates according to the mass-action kinetics follow from (2.18) to be (listed on the right)

$$\begin{array}{rcl|l}
\mathrm{A} + 2\mathrm{X} & \xrightarrow{k_1} & 3\mathrm{X}, & v_1 = k_1 x_{\mathrm{A}} x^2, \\
3\mathrm{X} & \xrightarrow{k_2} & \mathrm{A} + 2\mathrm{X}, & v_2 = k_2 x^3, \\
\mathrm{B} & \xrightarrow{k_3} & \mathrm{X}, & v_3 = k_3 x_{\mathrm{B}}, \\
\mathrm{X} & \xrightarrow{k_4} & \mathrm{B}, & v_4 = k_4 x\,.
\end{array}$$

As far as X_3 is concerned, the stoichiometry matrix S and the reaction rate v

can be written as

$$S = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}, \quad v = \begin{bmatrix} k_1 x_{\mathrm{A}} x^2 \\ k_2 x^3 \\ k_3 x_{\mathrm{B}} \\ k_4 x \end{bmatrix}.$$

The deterministic ODE turns out to be

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Sv = k_1 x_{\mathrm{A}} x^2 - k_2 x^3 + k_3 x_{\mathrm{B}} - k_4 x\,. \tag{2.26}$$

Example 2.13 (Stochastic focusing) The branched reaction scheme (2.10):

$$\varnothing \underset{k_d}{\overset{k_s}{\rightleftharpoons}} \mathrm{S}, \quad \varnothing \underset{k_a x_{\mathrm{S}}}{\overset{k_i}{\rightleftharpoons}} \mathrm{I} \xrightarrow{k_p} \mathrm{P} \xrightarrow{1} \varnothing\,.$$

Write $x_{\mathrm{S}}(t)$, $x_{\mathrm{I}}(t)$, and $x_{\mathrm{P}}(t)$ for the respective time-dependent molar concentrations of the signal S, the intermediary precursor I, and product P. The reaction rates based on mass-action kinetics are k_s for synthesis of S and $k_d x_{\mathrm{S}}$ for its degradation, k_i for synthesis of I and $k_a x_{\mathrm{S}} x_{\mathrm{I}}$ for its degradation, $k_p x_{\mathrm{I}}$ for the $\mathrm{I} \to \mathrm{P}$ conversion and $-x_{\mathrm{P}}$ the product degradation. Ordering the species as $\{\mathrm{S}, \mathrm{I}, \mathrm{P}\}$, the stoichiometry matrix S and the reaction rate v take the forms

$$S = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad v = \begin{bmatrix} k_s \\ k_d x_{\mathrm{S}} \\ k_i \\ k_a x_{\mathrm{S}} x_{\mathrm{I}} \\ k_p x_{\mathrm{I}} \\ -x_{\mathrm{P}} \end{bmatrix}.$$

The deterministic system of ODEs for the system can now be read from the

vector Sv:

$$\left.\begin{aligned} \frac{\mathrm{d}x_\mathrm{S}}{\mathrm{d}t} &= k_s - k_d x_\mathrm{S}, \\ \frac{\mathrm{d}x_\mathrm{I}}{\mathrm{d}t} &= k_i - (k_p + k_a x_\mathrm{S})\, x_\mathrm{I}, \\ \frac{\mathrm{d}x_\mathrm{P}}{\mathrm{d}t} &= k_p x_\mathrm{I} - x_\mathrm{P}\,. \end{aligned}\right\} \tag{2.27}$$

Example 2.14 (Hyperbolic control) If the pool of I-molecules is insignificant, the two reactions involving their loss are fast enough, and X_S does not change significantly during the life span of an individual I-molecule, then we can assume the steady state of ending up in P or A to be reached immediately. The steady-state abundance of I-molecules, obtained by setting to zero the right side of the second equation in (2.27), is $x_I^\mathrm{ss} = {}^{k_i}\!/_{(k_p + k_a X_\mathrm{S})}$. That leads to the following simplification of (2.27):

$$\left.\begin{aligned} \frac{\mathrm{d}x_\mathrm{S}}{\mathrm{d}t} &= k_s - k_d x_\mathrm{S}, \\ \frac{\mathrm{d}x_\mathrm{P}}{\mathrm{d}t} &= \frac{k_p k_i}{k_p + k_a x_\mathrm{S}} - x_\mathrm{P}, \end{aligned}\right\} \tag{2.28}$$

and a corresponding reduction of the branched reaction scheme (2.10):

$$\varnothing \underset{k_d}{\overset{k_s}{\rightleftharpoons}} \mathrm{S}, \quad \varnothing \underset{1}{\overset{k_i/(1+x_\mathrm{S}/K)}{\rightleftharpoons}} \mathrm{P}, \tag{2.29}$$

where $K = {}^{k_p}\!/_{k_a}$ is the inhibition constant. The denominator $1 + {}^{x_\mathrm{S}}\!/_K$ in the expression for the new effective rate coefficient suggests the name "hyperbolic control" for the product molecule by the signal molecule.

Example 2.15 (Gene regulation) For the gene regulation scheme (2.11):

$$\begin{aligned} \mathrm{G} &\xrightarrow{k_m} \mathrm{G} + \mathrm{M} && \text{(transcription)}, \\ \mathrm{M} &\xrightarrow{k_p} \mathrm{M} + \mathrm{P} && \text{(translation)}, \\ \mathrm{G} + \mathrm{P} &\underset{k_u}{\overset{k_b}{\rightleftharpoons}} \mathrm{GP} && \text{(binding/unbinding)}, \\ \mathrm{M} &\xrightarrow{k_m^-} \varnothing, \quad \mathrm{P} \xrightarrow{k_p^-} \varnothing && \text{(degradation)}, \end{aligned}$$

write $x_\mathrm{M}(t)$, $x_\mathrm{G}(t)$, and $x_\mathrm{P}(t)$ for the respective time-dependent molar concentrations of mRNA M, the unbound gene G, and protein P. The total gene concentration $x_\mathrm{G}^\mathrm{tot}$ is assumed to be constant, so that the bound (repressed) protein concentration is simply $x_\mathrm{G}^\mathrm{tot} - x_\mathrm{G}$. The reaction rates based on mass-

action kinetics are $k_m x_{\mathrm{G}}$ for transcription, $k_p x_{\mathrm{M}}$ for translation, $k_b x_{\mathrm{G}} x_{\mathrm{P}}$ for the gene–protein binding, $k_u \left(x_{\mathrm{G}}^{\mathrm{tot}} - x_{\mathrm{G}}\right)$ for the gene–protein unbinding, $k_m^- x_{\mathrm{M}}$ for mRNA degradation, and $k_p^- x_{\mathrm{P}}$ for protein degradation. Ordering the species as {M, G, P}, the stoichiometry matrix S and the reaction rate v take the forms

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \end{bmatrix}, \quad v = \begin{bmatrix} k_m x_{\mathrm{G}} \\ k_p x_{\mathrm{M}} \\ k_b x_{\mathrm{G}} x_{\mathrm{P}} \\ k_u \left(x_{\mathrm{G}}^{\mathrm{tot}} - x_{\mathrm{G}}\right) \\ k_m^- x_{\mathrm{M}} \\ k_p^- x_{\mathrm{P}} \end{bmatrix}.$$

The deterministic system of ODEs for the system can now be constructed from the vector Sv:

$$\left.\begin{aligned} \frac{\mathrm{d}x_{\mathrm{M}}}{\mathrm{d}t} &= k_m x_{\mathrm{G}} - k_m^- x_{\mathrm{M}}, \\ \frac{\mathrm{d}x_{\mathrm{G}}}{\mathrm{d}t} &= k_u \left(x_{\mathrm{G}}^{\mathrm{tot}} - x_{\mathrm{G}}\right) - k_b x_{\mathrm{G}} x_{\mathrm{P}}, \\ \frac{\mathrm{d}x_{\mathrm{P}}}{\mathrm{d}t} &= k_p x_{\mathrm{M}} + k_u \left(x_{\mathrm{G}}^{\mathrm{tot}} - x_{\mathrm{G}}\right) - \left(k_b x_{\mathrm{G}} + k_p^-\right) x_{\mathrm{P}}. \end{aligned}\right\} \tag{2.30}$$

2.4 The Art of Modeling

To do mathematical modeling at the life sciences interface is to engage in an act of discovery and conjecture. The art of modeling is not in the accuracy of a mathematical model but in the explanation, that is, in the argument that is developed in the process outlined in Figure 1.4. It is this argument and its context that give the model its validity. Mathematical modeling of cell-biological systems is an art—the art of asking suitable questions, choosing an appropriate conceptual framework to formulate and test hypotheses, and making appropriate assumptions and simplifications. Our goal is to improve the understanding of living systems, and we believe that there is nothing more practical in addressing the complexity of living systems than mathematical modeling.

What we are seeking is an understanding of the functioning of cells, of

their behavior and the mechanisms underlying it. When we speak of mechanisms and principles as being the goal of our scientific quest, we really mean that we are interested in the system's organization [168]. In living systems there are two forms of interlinked organization: The *structural organization* of a cell refers to the arrangement and structural (material or biophysical) properties of its components— organelles and macromolecules. Inseparable from the cell's structural organization is its *functional organization*, describing the processes that determine the cell's behavior or '(mal)functioning'. Interacting with other cells and/or its environment, the cell realizes four key functions: growth, proliferation, apoptosis, and differentiation. The processes that realize these functions of a cell can be further organized into three process levels: gene regulation, signal transduction, and metabolism (Figure 1.3). The experimental study of any one of these cell functions and any one of these process levels is subject to high degrees of specialization. These specialized research fields are often separated by technology, methodology, and culture. This depth of specialization is a hurdle to a comprehensive understanding of how cells and cell populations (mal)function.
In summary, systems theory is the study of organization, using mathematical modeling. With respect to systems biology, the key challenges are:

- Depending on the data and question at hand, what approach to choose and why?
- How do I decompose a complex system intro tractable subsystems?
- Given an understanding of subsystems, how can one integrate these data and models into an understanding of the system as a whole?

Techniques for coupling/embedding models of components built on disparate time and length scales, and often with different modeling techniques, into larger models spanning much longer scales are in their infancy and require further investigation. We limit ourselves in this text to a small subset of these challenges and focus on one particular approach to studying small subsystems.

Problems

2.1. When the volume is not known or important, it is convenient to choose a value so that each nanomolar concentration is numerically equal to the corresponding copy number. Compute that value of the volume.

2.2. Suppose species concentration is measured in molecules per μm^3 (cubic micrometers) of volume. What can you say about the magnitude and unit of the system size?

2.3. Consider the irreversible bimolecular reaction

$$\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{X} + \mathrm{Y}\,.$$

Temporal changes in species concentration for this reaction are restricted by a conservation relation.

1. Write down the conservation relation for concentrations in terms of initial concentrations.
2. Express the reaction rate law in terms of time-dependent concentration of X.
3. Implement the rate law as a Matlab function handle. Assume that $k = 1\,\mathrm{sec}^{-1}$ and initial abundances are 2 M for A, 3 M for B, and 0.5 M for X.
4. Call the function handle in an ODE solver to compute and plot the time-course concentration of X for the first 5 seconds.

2.4. Consider the consecutive reaction

$$\mathrm{X}_1 \xrightarrow{k_1} \mathrm{X}_2 \xrightarrow{k_2} \mathrm{X}_3\,.$$

1. Write down the differential equation for the concentration X_2.
2. Assume zero initial concentrations except for the first reactant, which is 10 M, and take $1\,\mathrm{sec}^{-1}$ for both rate constants. Run the following script:

```
x0 = [10;0]; % initial concentrations
k1 = 1; k2 = k1; % rate constants
v = @(t,x) [-k1*x(1);k1*x(1)-k2*x(2)]; % rate law
[t,x] = ode45(v,[0 5],x0); % solver
plot(t,x(:,2)) % plot x2
```

   Repeat the simulation for $k_2 = 0.1k_1$ and $k_2 = 10k_1$. Relate the relative magnitudes of the rate constants to the relative reaction rates.

3. If one of the two reactions is much faster than the other, the overall reaction rate is determined by the slower reaction, which is then called the "rate-determining step." For each value of k_2, which reaction is rate-determining?

2.5. Recall the rate law

$$v(x) = k \prod_{i=1}^{s} x_i^{g_i}$$

with definitive orders for a chemical reaction. It can be implemented as a function handle:

```
k = 2; % rate constant
g = [0 1 1 0 0 1]'; % reaction stoichiometry
v = @(x) k*prod(x.^g); % rate law
```

for the specified values of k and g.

1. Evaluate the rate expression for

$$x = \begin{bmatrix} 2 & 0.5 & 0 & 1.5 & 0 & 3 \end{bmatrix}^T.$$

   What problem did you encounter? Can you figure out why?

2. Reimplement the rate law as a function that accounts for the pitfall you encountered.

2.6. Consider a simple network

$$2\mathrm{X}_1 \xrightarrow{k_1} \mathrm{X}_2, \quad \mathrm{X}_2 + \mathrm{X}_3 \xrightarrow{k_2} \mathrm{X}_4,$$

of metabolites. The metabolite concentrations are measured in molecules per $\mu\mathrm{m}^3$ (cubic micrometers).

1. Set up the stoichiometry matrix S.

2. Write down the expression, based on mass-action kinetics, for the two reaction rates v_1 and v_2 in terms of species concentrations.

3. How would you combine the two results to construct the ODEs that describe how species concentrations change with time.

4. Complete the following script based on the quantities in the above steps in order to compute and plot the species concentrations against time over 500 seconds:

```
% initial abundance (molecules per cubic micrometer)
x0 = [10;0;5;0];
% rate constants (per cubic micrometer per second)
k = [1e-3;3e-3];
```

```
% S = ?; % stoichiometry matrix
% v = @(x) ?; % rate law
% dxdt = ?; % ODEs
[t,x] = ode15s(v,[0 500],x0); % solver
plot(t,x) % plot x
```

5. Discover the conservation relations in the reaction scheme and utilize them to rewrite the rate equations so that they involve concentrations of X_2 and X_3 only.

6. Modify the code accordingly and check the result by plotting and comparing with the previous implementation.

2.7. The repressilator consists of three genes connected in a feedback loop such that each gene product represses the next gene in the loop and is repressed by the previous gene [40]. If we use subscripts $i = 1, 2, 3$ to denote the three genes; M_i represents mRNAs, and P_i the proteins. The gene network can be represented by the reaction scheme

$$M_i \underset{\alpha(M_{i-1})}{\overset{1}{\rightleftharpoons}} \varnothing, \quad P_i \xrightarrow{b} \varnothing, \quad M_i \xrightarrow{b} M_i + P_i$$

where i runs through 1, 2, 3 and $P_0 = P_3$. For simplicity, assume relative (nondimensional) concentrations. The mRNA transcription rate is

$$\alpha(x) = a_0 + \frac{a_1}{(1+x)^h},$$

where a_0 is the transcription rate in the presence of saturating repressor and $a_0 + a_1$ represents the maximal transcription rate in the absence of the repressor. The exponent h in the denominator is the Hill coefficient. The parameter b appears as the protein degradation rate constant and translation rate constant.

1. Set up the stoichiometry matrix S by adopting the ordering M_1, M_2, M_3, P_1, P_2, P_3 for species and the ordering $M_1 \to \varnothing$, $M_2 \to \varnothing$, $M_3 \to \varnothing$, $P_1 \to \varnothing$, $P_2 \to \varnothing$, $P_3 \to \varnothing$, $\varnothing \to M_1$, $\varnothing \to M_2$, $\varnothing \to M_3$, $M_1 \to M_1 + P_1$, $M_2 \to M_2 + P_2$, $M_3 \to M_3 + P_3$ for reactions.

2. Write down the expressions for channelwise reaction rates v_j in terms of species concentrations.

3. Combine the two results to construct the ODEs that describe how species concentrations change with time.

4. Complete the following script based on the quantities in the above steps in order to compute and plot the protein levels for 50 time units:

```
% parameters
a0 = 0.25; a1 = 250; b = 5; h = 2.1;
% S = ?; % stoichiometry matrix
% v = @(x) ?; % rate law
dxdt = @(t,x) S*v(x); % ODEs
tmax = 50; % time
x0 = [0 0 0 4 0 15]'; % initial concentration
[t,x] = ode45(dxdt,[0 tmax],x0); % solution
plot(t,x(:,4:6)) % plot protein levels
```

5. Do you see oscillations in the protein levels? Play with the parameter values and initial conditions to see whether you always get oscillations.

6. Looking at time plots for checking oscillations is one way to solve part 5 above. An alternative is to look at the phase plot. Extend the code to plot the phase plots for each mRNA–protein pair. What do these phase plots reflect?

2.8. The repressilator model in the last exercise is a nondimensional version of the original model available on the biomodel database http://biomodels.caltech.edu/BIOMD0000000012. Run the online simulation provided. Do you see oscillations in the protein levels? Play with the parameter values and initial conditions to see whether you always get oscillations.

http://www.springer.com/978-1-4614-0477-4

Stochastic Approaches for Systems Biology

Ullah, M.; Wolkenhauer, O.
2011, XXXII, 290 p., Hardcover
ISBN: 978-1-4614-0477-4