2
Modeling of Biochemical Systems

2.1
Kinetic Modeling of Enzymatic Reactions

Summary

The rate of an enzymatic reaction, i.e., the velocity by which the execution of the reaction changes the concentrations of its substrates, is determined by concentrations of its substrates, concentration of the catalyzing enzyme, concentrations of possible modifiers, and by certain parameters. We introduce different kinetic laws for reversible and irreversible reactions, for reactions with varying numbers of substrates, and for reactions that are subject to inhibition or activation. The derivations of the rate laws are shown and the resulting restrictions for their validity and applicability. Saturation and sigmoidal kinetics are explained. The connection to thermodynamics is shown.

Deterministic kinetic modeling of individual biochemical reactions has a long history. The Michaelis–Menten model for the rate of an irreversible one-substrate reaction is an integral part of biochemistry, and the K_m value is a major characteristic of the interaction between enzyme and substrate. Biochemical reactions are catalyzed by enzymes, i.e., specific proteins which often function in complex with cofactors. They have a catalytic center, are usually highly specific, and remain unchanged by the reaction. One enzyme molecule can catalyze thousands of reactions per second (this so-called turnover number ranges from $10^2 \, \mathrm{s}^{-1}$ to $10^7 \, \mathrm{s}^{-1}$). Enzyme catalysis leads to a rate acceleration of about 10^6- up to 10^{12}-fold compared to the noncatalyzed, spontaneous reaction.

In this chapter, we make you familiar with the basic concepts of the mass action rate law. We will show how you can derive and apply more advanced kinetic expressions. The effect of enzyme inhibitors and activators will be discussed. The thermodynamic foundations and constraints are introduced.

The basic quantities are the concentration S of a substance S, i.e., the number n of molecules (or, alternatively, moles) of this substance per volume V, and the rate v of a reaction, i.e., the change of concentration S per time t. This type of modeling is

macroscopic and phenomenological, compared to the microscopic approach, where single molecules and their interactions are considered. Chemical and biochemical kinetics rely on the assumption that the reaction rate v at a certain point in time and space can be expressed as a unique function of the concentrations of all substances at this point in time and space. Classical enzyme kinetics assumes for sake of simplicity a spatial homogeneity (the "well-stirred" test tube) and no direct dependency of the rate on time

$$v(t) = v(S(t)). \tag{2.1}$$

In more advanced modeling approaches, longing toward whole-cell modeling, spatial inhomogeneities are taken into account, paying tribute to the fact that many components are membrane-bound or that cellular structures hinder the free movement of molecules. But, in the most cases, one can assume that diffusion is rapid enough to allow for an even distribution of all substances in space.

2.1.1
The Law of Mass Action

Biochemical kinetics is based on the mass action law, introduced by Guldberg and Waage in the nineteenth century [1–3]. It states that the reaction rate is proportional to the probability of a collision of the reactants. This probability is in turn proportional to the concentration of reactants to the power of the molecularity, that is the number in which they enter the specific reaction. For a simple reaction such as

$$S_1 + S_2 \rightleftharpoons 2P, \tag{2.2}$$

the reaction rate reads

$$v = v_+ - v_- = k_+ S_1 \cdot S_2 - k_- P^2. \tag{2.3}$$

where v is the net rate; v_+ and v_- are the rates of the forward and backward reactions; and k_+ and k_- are the *kinetic* or *rate constants*, i.e., the respective proportionality factors.

The molecularity is 1 for S_1 and for S_2 and 2 for P, respectively. If we measure the concentration in $mol\,l^{-1}$ (or M) and the time in seconds (s), then the rate has the unit $M\,s^{-1}$. Accordingly, the rate constants for bimolecular reactions have the unit $M^{-1}\,s^{-1}$. Rate constants of monomolecular reactions have the dimension s^{-1}. The general mass action rate law for a reaction transforming m_i substrates with concentrations S_i into m_j products with concentrations P_j reads

$$v = v_+ - v_- = k_+ \prod_{i=1}^{m_i} S_i^{n_i} - k_- \prod_{j=1}^{m_j} P_j^{n_j}, \tag{2.4}$$

where n_i and n_j denote the respective molecularities of S_i and P_j in this reaction.

The equilibrium constant K_{eq} (we will also use the simpler symbol q) characterizes the ratio of substrate and product concentrations in equilibrium (S_{eq} and P_{eq}), i.e., the state with equal forward and backward rate. The rate constants are related to K_{eq} in the

following way:

$$K_{eq} = \frac{k_+}{k_-} = \frac{\prod_{j=1}^{m_j} P_{j,eq}^{n_j}}{\prod_{i=1}^{m_i} S_{i,eq}^{n_i}} \tag{2.5}$$

The relation between the thermodynamic and the kinetic description of biochemical reactions will be outlined in Section 2.1.2.

The equilibrium constant for the reaction given in Eq. (2.2) is $K_{eq} = P_{eq}^2/(S_{1,eq} \cdot S_{2,eq})$. The dynamics of the concentrations away from equilibrium is described by the ODEs.

$$\frac{d}{dt} S_1 = \frac{d}{dt} S_2 = -v \quad \text{and} \quad \frac{d}{dt} P = 2v. \tag{2.6}$$

The time course of S_1, S_2, and P is obtained by integration of these ODEs (see Section 2.3).

Example 2.1

The kinetics of a simple decay like

$$S \rightarrow \tag{2.7}$$

is described by $v = kS$ and $dS/dt = -kS$. Integration of this ODE from time $t = 0$ with the initial concentration S_0 to an arbitrary time t with concentration $S(t)$, $\int_{S_0}^{S} dS/S = -\int_{t=0}^{t} k\, dt$, yields the temporal expression $S(t) = S_0 e^{-kt}$.

2.1.2
Reaction Kinetics and Thermodynamics

An important purpose of metabolism is to extract energy from nutrients, which is necessary for the synthesis of molecules, growth, and proliferation. We distinguish between energy-supplying reactions, energy-demanding reactions, and energetically neutral reactions. The principles of reversible thermodynamics and their application to chemical reactions allow understanding of energy circulation in the cell.

A biochemical process is characterized by the direction of the reaction, by whether it occurs spontaneously or not, and by the position of the equilibrium. The first law of thermodynamics, i.e., the law of energy conservation, tells us that the total energy of a closed system remains constant during any process. The second law of thermodynamics states that a process occurs spontaneous only if it increases the total entropy of the system. Unfortunately, entropy is usually not directly measurable. A more suitable measure is the Gibbs free energy G, which is the energy capable of carrying out work under isotherm–isobar conditions, i.e., at constant temperature and constant pressure. The change of the free energy is given as

$$\Delta G = \Delta H - T \Delta S, \tag{2.8}$$

where ΔH is the change in enthalpy, ΔS the change in entropy, and T the absolute temperature in Kelvin. ΔG is a measure for the driving force, the spontaneity of a chemical reaction. The reaction proceeds spontaneous under release of energy, if $\Delta G < 0$ (exergonic process). If $\Delta G > 0$, then the reaction is energetically not favorable and will not occur spontaneously (endergonic process). $\Delta G = 0$ means that the system has reached its equilibrium. Endergonic reactions may proceed if they obtain energy from a strictly exergonic reaction by energetic coupling. In tables, free energy is usually given for standard conditions (ΔG°), i.e., for a concentration of the reaction partners of 1 M, a temperature of $T = 298$ K, and, for gaseous reactions, a pressure of $p = 98$, 1 kPa $= 1$ atm. The unit is kJ mol^{-1}. Free energy differences satisfy a set of relations as follows. The free energy difference for a reaction can be calculated from the balance of free energies of formation of its products and substrates:

$$\Delta G = \sum G_P - \sum G_S. \tag{2.9}$$

The enzyme cannot change the free energies of the substrates and products of a reaction, neither their difference, but it changes the way the reaction proceeds microscopically, the so-called reaction path, thereby lowering the activation energy for the reaction. The *Transition State Theory* explains this as follows. During the course of a reaction, the metabolites must pass one or more transition states of maximal free energy, in which bonds are solved or newly formed. The transition state is unstable; the respective molecule configuration is called an activated complex. It has a lifetime of around one molecule vibration, 10^{-14}–10^{-13} s, and it can hardly be experimentally verified. The difference $\Delta G^{\#}$ of free energy between the reactants and the activated complex determines the dynamics of a reaction: the higher this difference, the lower the probability that the molecules may pass this barrier and the lower the rate of the reaction. The value of $\Delta G^{\#}$ depends on the type of altered bonds, on steric, electronic, or hydrophobic demands, and on temperature.

Figure 2.1 presents a simplified view of the reaction course. The substrate and the product are situated in local minima of the free energy; the active complex is assigned to the local maximum. The free energy difference ΔG is proportional to the logarithm of the equilibrium constant K_{eq} of the respective reaction:

$$\Delta G = -RT \ln K_{eq}, \tag{2.10}$$

where R is the gas constant, 8.314 J mol^{-1} K^{-1}. The value of $\Delta G^{\#}$ corresponds to the kinetic constant k_+ of the forward reaction (Eqs. (2.3)–(2.5)) by $\Delta G^{\#} = -RT \ln k_+$, while $\Delta G^{\#} + \Delta G$ is related to the rate constant k_- of the backward reaction.

The interaction of the reactants with an enzyme may alter the reaction path and, thereby, lead to lower values of $\Delta G^{\#}$ as well as higher values of the kinetic constants. Furthermore, the free energy may assume more local minima and maxima along the path of reaction. They are related to unstable intermediary complexes. Values for the difference of free energy for some biologically important reactions are given in Table 2.1.
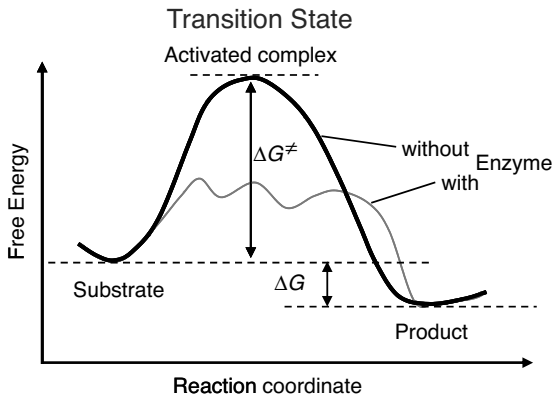
Figure 2.1 Change of free energy along the course of a reaction. The substrate and the product are situated in local minima of the free energy; the active complex is assigned to the local maximum. The enzyme may change the reaction path and thereby lower the barrier of free energy.

A biochemical reaction is reversible if it may proceed in both directions, leading to a positive or negative sign of the rate v. The actual direction depends on the current reactant concentrations. In theory, every reaction should be reversible. In practice, we can consider many reactions as irreversible since (i) reactants in cellular environment cannot assume any concentration, (ii) coupling of a chemical conversion to ATP consumption leads to a severe drop in free energy and therefore makes a reaction reversal energetically unfavorable, and (iii) for compound destruction, such as protein degradation, reversal by chance is extremely unlikely.

The detailed consideration of enzyme mechanisms by applying the mass action law for the single events has led to a number of standard kinetic descriptions, which will be explained in the following.

Table 2.1 Values of $\Delta G^{0'}$ and K_{eq} for some important reactions[a].

Reactions	$\Delta G^{0'}$ (kJ mol^{-1})
$2H_2 + O_2 \rightarrow 2H_2O$	-474
$2H_2O_2 \rightarrow 2H_2O + O_2$	-99
$PP_i + H_2O \rightarrow 2P_i$	-33.49
$ATP + H_2O \rightarrow ADP + P_i$	-30.56
Glucose-6-phosphate $+ H_2O \rightarrow$ Glucose $+ P_i$	-13.82
Glucose $+ P_i \rightarrow$ Glucose-6-phosphate $+ H_2O$	$+13.82$
Glucose-1-phosphate $\rightarrow$ Glucose-6-phosphate	-7.12
Glucose-6-phosphate $\rightarrow$ Fructose-6-phosphate	$+1.67$
Glucose $+ 6O_2 \rightarrow 6CO_2 + 6H_2O$	-2890

[a]*Source:* ZITAT: Lehninger, A.L. Biochemistry, 2nd edition, New York, Worth, 1975, p. 397.

2.1.3
Michaelis–Menten Kinetics

Brown [4] proposed an enzymatic mechanism for invertase, catalyzing the cleavage of saccharose to glucose and fructose. This mechanism holds in general for all one-substrate reactions without backward reaction and effectors, such as

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\longrightarrow} E + P. \tag{2.11}$$

It comprises a reversible formation of an enzyme–substrate complex ES from the free enzyme E and the substrate S and an irreversible release of the product P. The ODE system for the dynamics of this reaction reads

$$\frac{dS}{dt} = -k_1 E \cdot S + k_{-1} ES, \tag{2.12}$$

$$\frac{dES}{dt} = k_1 E \cdot S - (k_{-1} + k_2) ES, \tag{2.13}$$

$$\frac{dE}{dt} = -k_1 E \cdot S + (k_{-1} + k_2) ES, \tag{2.14}$$

$$\frac{dP}{dt} = k_2 ES. \tag{2.15}$$

The reaction rate is equal to the negative decay rate of the substrate as well as to the rate of product formation:

$$v = -\frac{dS}{dt} = \frac{dP}{dt}. \tag{2.16}$$

This ODE system (Eqs. (2.12)–(2.16)) cannot be solved analytically. Different assumptions have been used to simplify this system in a satisfactory way. Michaelis and Menten [5] considered a *quasi-equilibrium* between the free enzyme and the enzyme–substrate complex, meaning that the reversible conversion of E and S to ES is much faster than the decomposition of ES into E and P, or in terms of the kinetic constants,

$$k_1, k_{-1} \gg k_2. \tag{2.17}$$

Briggs and Haldane [6] assumed that during the course of reaction a state is reached where the concentration of the ES complex remains constant, the so-called quasi-steady state. This assumption is justified only if the initial substrate concentration is much larger than the enzyme concentration, $S(t = 0) \gg E$, otherwise such a state will never be reached. In mathematical terms, we obtain

$$\frac{dES}{dt} = 0. \tag{2.18}$$

In the following, we derive an expression for the reaction rate from the ODE system (2.12)–(2.15) and the quasi-steady-state assumption for ES. First, adding Eqs. (2.13) and (2.14) results in

$$\frac{\mathrm{d}ES}{\mathrm{d}t} + \frac{\mathrm{d}E}{\mathrm{d}t} = 0 \quad \text{or} \quad E_{\text{total}} = E + ES = \text{constant}. \tag{2.19}$$

This expression shows that enzyme is neither produced nor consumed in this reaction; it may be free or part of the complex, but its total concentration remains constant. Introducing (2.19) into (2.13) under the steady-state assumption (2.18) yields

$$ES = \frac{k_1 E_{\text{total}} S}{k_1 S + k_{-1} + k_2} = \frac{E_{\text{total}} S}{S + (k_{-1} + k_2)/k_1}. \tag{2.20}$$

For the reaction rate, this gives

$$v = \frac{k_2 E_{\text{total}} S}{S + ((k_{-1} + k_2)/k_1)}. \tag{2.21}$$

In enzyme kinetics, it is convention to present Eq. (2.21) in a simpler form, which is important in theory and practice

$$v = \frac{V_{\text{max}} S}{S + K_{\text{m}}}. \tag{2.22}$$

Equation (2.22) is the expression for Michaelis–Menten kinetics. The parameters have the following meaning: the *maximal velocity,*

$$V_{\text{max}} = k_2 E_{\text{total}}, \tag{2.23}$$

is the maximal rate that can be attained, when the enzyme is completely saturated with substrate. The *Michaelis constant,*

$$K_{\text{m}} = \frac{k_{-1} + k_2}{k_1}, \tag{2.24}$$

is equal to the substrate concentration that yields the half-maximal reaction rate. For the quasi-equilibrium assumption (Eq. (2.17)), it holds that $K_{\text{m}} \cong k_{-1}/k_1$. The maximum velocity divided by the enzyme concentration (here $k_2 = v_{\text{max}}/E_{\text{total}}$) is often called the turnover number, k_{cat}. The meaning of the parameters is illustrated in the plot of rate versus substrate concentration (Figure 2.2).

2.1.3.1 How to Derive a Rate Equation

Below, we will present some enzyme kinetic standard examples to derive a rate equation. Individual mechanisms for your specific enzyme of interest may be more complicated or merely differ from these standards. Therefore, we summarize here the general way of deriving a rate equation.

1. Draw a wiring diagram of all steps to consider (e.g., Eq. (2.11)). It contains all substrates and products (S and P) and n free or bound enzyme species (E and ES).
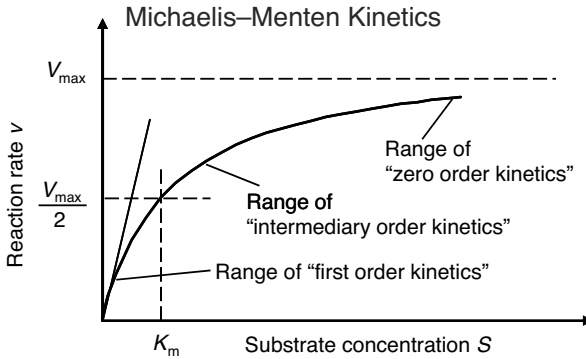
Figure 2.2 Dependence of reaction rate ν on substrate concentration S in Michaelis–Menten kinetics. V_{max} denotes the maximal reaction rate that can be reached for large substrate concentration. K_m is the substrate concentration that leads to half-maximal reaction rate. For low substrate concentration, ν increases almost linearly with S, while for high substrate concentrations ν is almost independent of S.

2. The right sites of the ODEs for the concentrations changes sum up the rates of all steps leading to or away from a certain substance (e.g., Eqs. (2.12)–(2.15)). The rates follow mass action kinetics (Eq. (2.3)).
3. The sum of all enzyme-containing species is equal to the total enzyme concentration E_{total} (the right site of all differential equations for enzyme species sums up to zero). This constitutes one equation.
4. The assumption of quasi-steady state for $n-1$ enzyme species (i.e., setting the right sites of the respective ODEs equal to zero) together with (3.) result in n algebraic equations for the concentrations of the n enzyme species.
5. The reaction rate is equal to the rate of product formation (e.g., Eq. (2.16)). Insert the respective concentrations of enzyme species resulting from (4.).

2.1.3.2 Parameter Estimation and Linearization of the Michaelis–Menten Equation

To assess the values of the parameters V_{max} and K_m for an isolated enzyme, one measures the initial rate for different initial concentrations of the substrate. Since the rate is a nonlinear function of the substrate concentration, one has to determine the parameters by nonlinear regression. Another way is to transform Eq. (2.22) to a linear relation between variables and then apply linear regression.

The advantage of the transformed equations is that one may read the parameter value more or less directly from the graph obtained by linear regression of the measurement data. In the plot by Lineweaver and Burk [7] (Table 2.2), the values for V_{max} and K_m can be obtained from the intersections of the graph with the ordinate and the abscissa, respectively. The Lineweaver–Burk plot is also helpful to easily discriminate different types of inhibition (see below). The drawback of the transformed equations is that they may be sensitive to errors for small or high substrate

Table 2.2 Different approaches for the linearization of Michaelis–Menten enzyme kinetics.

	Lineweaver–Burk	Eadie–Hofstee	Hanes–Woolf
Transformed equation	$\dfrac{1}{v} = \dfrac{K_m}{V_{max}}\dfrac{1}{S} + \dfrac{1}{V_{max}}$	$v = V_{max} - K_m\dfrac{v}{S}$	$\dfrac{S}{v} = \dfrac{S}{V_{max}} + \dfrac{K_m}{V_{max}}$
New variables	$\dfrac{1}{v}, \dfrac{1}{S}$	$v, \dfrac{v}{S}$	$\dfrac{S}{v}, S$
Graphical representation			

concentrations or rates. Eadie and Hofstee [8] and Hanes and Woolf [9] have introduced other types of linearization to overcome this limitation.

2.1.3.3 The Michaelis–Menten Equation for Reversible Reactions

In practice, many reactions are reversible. The enzyme may catalyze the reaction in both directions. Consider the following mechanism:

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E} + \mathrm{P} \tag{2.25}$$

The product formation is given by

$$\frac{\mathrm{d}P}{\mathrm{d}t} = k_2 ES - k_{-2} E \cdot P = v. \tag{2.26}$$

The respective rate equation reads

$$
\begin{aligned}
v &= E_{\text{total}} \frac{Sq - P}{Sk_1/(k_{-1}k_{-2}) + 1/k_{-2} + k_2/(k_{-1}k_{-2}) + P/k_{-1}} \\
&= \frac{(V_{\max}^{\text{for}}/K_{\mathrm{m}S})S - (V_{\max}^{\text{back}}/K_{\mathrm{m}P})P}{1 + S/K_{\mathrm{m}S} + P/K_{\mathrm{m}P}}.
\end{aligned}
\tag{2.27}
$$

While the parameters $k_{\pm 1}$ and $k_{\pm 2}$ are the kinetic constants of the individual reaction steps, the phenomenological parameters $V_{\max}^{\text{for}}$ and $V_{\max}^{\text{back}}$ denote the maximal velocity in forward or backward direction, respectively, under zero product or substrate concentration, and the phenomenological parameters $K_{\mathrm{m}S}$ and $K_{\mathrm{m}P}$ denote the substrate or product concentration causing half maximal forward or backward rate. They are related in the following way [10]:

$$K_{\text{eq}} = \frac{V_{\max}^{\text{for}} K_{\mathrm{m}P}}{V_{\max}^{\text{back}} K_{\mathrm{m}S}}. \tag{2.28}$$

2.1.4
Regulation of Enzyme Activity by Effectors

Enzymes may immensely increase the rate of a reaction, but this is not their only function. Enzymes are involved in metabolic regulation in various ways. Their production and degradation is often adapted to the current requirements of the cell. Furthermore, they may be targets of effectors, both inhibitors and activators.

The effectors are small molecules, or proteins, or other compounds that influence the performance of the enzymatic reaction. The interaction of effector and enzyme changes the reaction rate. Such regulatory interactions that are crucial for the fine-tuning of metabolism will be considered here [11].

Basic types of inhibition are distinguished by the state, in which the enzyme may bind the effector (i.e., the free enzyme E, the enzyme–substrate complex ES, or both), and by the ability of different complexes to release the product. The general pattern of inhibition is schematically represented in Figure 2.3. The different types result, if some of the interactions may not occur.
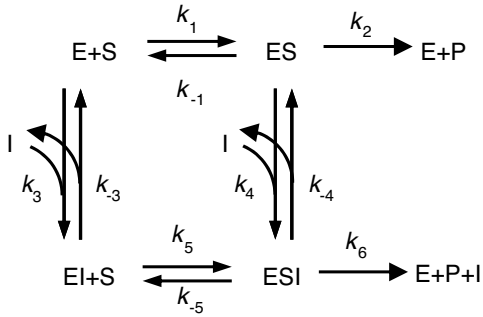
Figure 2.3 General scheme of inhibition in Michaelis–Menten kinetics. Reactions 1 and 2 belong to the standard scheme of Michaelis–Menten kinetics. Competitive inhibition is given, if in addition reaction 3 (and not reactions 4, 5, or 6) occurs. Uncompetitive inhibition involves reactions 1, 2, and 4, and noncompetitive inhibition comprises reactions 1, 2, 3, 4, and 5. Occurrence of reaction 6 indicates partial inhibition.

The rate equations are derived according to the following scheme:

1. Consider binding equilibriums between compounds and their complexes:

$$K_{\mathrm{m}} \cong \frac{k_{-1}}{k_1} = \frac{E \cdot S}{ES}, \ K_{\mathrm{I},3} = \frac{k_{-3}}{k_3} = \frac{E \cdot I}{EI}, \ K_{\mathrm{I},4} = \frac{k_{-4}}{k_4} = \frac{ES \cdot I}{ESI}, \ K_{\mathrm{I},5} = \frac{k_{-5}}{k_5} = \frac{EI \cdot S}{ESI}. \tag{2.29}$$

Note that, if all reactions may occur, the Wegscheider condition [12] holds in the form

$$\frac{k_1 k_4}{k_{-1} k_{-4}} = \frac{k_3 k_5}{k_{-3} k_{-5}}, \tag{2.30}$$

which means that the difference in the free energies between two compounds (e.g., E and ESI) is independent of the choice of the reaction path (here via ES or via EI).

2. Take into account the moiety conservation for the total enzyme (include only those complexes, which occur in the course of reaction):

$$E_{\mathrm{total}} = E + ES + EI + ESI. \tag{2.31}$$

3. The reaction rate is equal to the rate of product formation

$$v = \frac{\mathrm{d}P}{\mathrm{d}t} = k_2 ES + k_6 ESI. \tag{2.32}$$

Equations (2.29)–(2.31) constitute four independent equations for the four unknown concentrations of E, ES, EI, and ESI. Their solution can be inserted into Eq. (2.32). The effect of the inhibitor depends on the concentrations of substrate and inhibitor and on the relative affinities to the enzyme. Table 2.3 lists the different types of inhibition for irreversible and reversible Michaelis–Menten kinetics together with the respective rate equations.

Table 2.3 Types of inhibition for irreversible and reversible Michaelis–Menten kinetics[a].

Name	Implementation	Equation – irreversible case	Equation – reversible case	Characteristics
Competitive inhibition	I binds only to free E; P-release only from ES complex $k_{\pm4} = k_{\pm5} = k_6 = 0$	$v = \dfrac{V_{max} S}{K_m \cdot i_3 + S}$	$v = \dfrac{V_{max}^f (S/K_{mS}) - V_{max}^r (P/K_{mP})}{(S/K_{mS}) + (P/K_{mP}) + i_3}$	K_m changes, V_{max} remains same. S and I compete for the binding place; high S may out compete I.
Uncompetitive inhibition	I binds only to the ES complex; P-release only from ES complex $k_{\pm3} = k_{\pm5} = k_6 = 0$	$v = \dfrac{V_{max} S}{K_m + S \cdot i_4}$	$v = \dfrac{V_{max}^f (S/K_{mS}) - V_{max}^r (P/K_{mP})}{1 + ((S/K_{mS}) + (P/K_{mP})) i_4}$	K_m and V_{max} change, but their ratio remains same. S may not out compete I
Noncompetitive inhibition	I binds to E and ES; P-release only from ES $K_{I,3} = K_{I,4},\ k_6 = 0$	$v = \dfrac{V_{max} S}{(K_m + S) i_3}$	$v = \dfrac{V_{max}^f (S/K_{mS}) - V_{max}^r (P/K_{mP})}{(1 + (S/K_{mS}) + (P/K_{mP})) i_4}$	K_m remains, V_{max} changes. S may not out compete I
Mixed inhibition	I binds to E and ES; P-release only from ES $K_{I,3} \neq K_{I,4},\ k_6 = 0$	$v = \dfrac{V_{max} S}{K_m \cdot i_4 + S \cdot i_3}$		K_m and V_{max} change. $K_{I,3} > K_{I,4}$: competitive–noncompetitive inhibition $K_{I,3} < K_{I,4}$: noncompetitive–uncompetitive inhibition
Partial Inhibition	I may bind to E and ES; P-release from ES and ESI $K_{I,3} \neq K_{I,4},\ K_6 \neq 0$	$v = \dfrac{V_{max} S[1 + \{(k_6 I)/(k_2 K_{I,3})\}]}{K_m i_4 + S i_3}$		K_m and V_{max} change. if $k_6 > k_2$: activation instead of inhibition.

[a]These abbreviations are used: $K_{I,3} = \frac{k_{-3}}{k_3}$, $K_{I,4} = \frac{k_{-4}}{k_4}$, $i_3 = 1 + \frac{I}{K_{I,3}}$, $i_4 = 1 + \frac{I}{K_{I,4}}$.

In the case of *competitive* inhibition, the inhibitor competes with the substrate for the binding site (or inhibits substrate binding by binding elsewhere to the enzyme) without being transformed itself. An example for this type is the inhibition of succinate dehydrogenase by malonate. The enzyme converts succinate to fumarate forming a double bond. Malonate has two carboxyl groups, like the proper substrates, and may bind to the enzyme, but the formation of a double bond cannot take place. Since substrates and inhibitor compete for the binding sites, a high concentration of one of them may displace the other one. For very high substrate concentrations, the same maximal velocity as without inhibitor is reached, but the effective K_m value is increased.

In the case of *uncompetitive* inhibition, the inhibitor binds only to the ES complex. The reason may be that the substrate binding caused a conformational change, which opened a new binding site. Since S and I do not compete for binding sites, an increase in the concentration of S cannot displace the inhibitor. In the presence of inhibitor, the original maximal rate cannot be reached (lower V_{max}). For example, an inhibitor concentration of $I = K_{I,4}$ halves the K_m-value as well as V_{max}. Uncompetitive inhibition occurs rarely for one-substrate reactions, but more frequently in the case of two substrates. One example is inhibition of arylsulphatase by hydracine.

Noncompetitive inhibition is present, if substrate binding to the enzyme does not alter the binding of the inhibitor. There must be different binding sites for substrate and inhibitor. In the classical case, the inhibitor has the same affinity to the enzyme with or without bound substrate. If the affinity changes, this is called mixed inhibition. A standard example is inhibition of chymotrypsion by H^+-ions.

If the product may also be formed from the enzyme–substrate–inhibitor complex, the inhibition is only partial. For high rates of product release (high values of k_6), this can even result in an activating instead of an inhibiting effect.

The general types of inhibition, competitive, uncompetitive, and noncompetitive inhibition also apply for the reversible Michaelis–Menten mechanism. The respective rate equations are also listed in Table 2.3.

2.1.4.1 Substrate Inhibition

A common characteristic of enzymatic reaction is the increase of the reaction rate with increasing substrate concentration S up to the maximal velocity V_{max}. But in some cases, a decrease of the rate above a certain value of S is recorded. A possible reason is the binding of a further substrate molecule to the enzyme–substrate complex yielding the complex ESS that cannot form a product. This kind of inhibition is reversible if the second substrate can be released. The rate equation can be derived using the scheme of uncompetitive inhibition by replacing the inhibitor by another substrate. It reads

$$v = k_2 ES = \frac{V_{max}S}{K_m + S(1 + (S/K_I))}. \tag{2.33}$$

This expression has an optimum, i.e., a maximal value of v, at

$$S_{opt} = \sqrt{K_m K_I} \quad \text{with} \quad v_{opt} = \frac{V_{max}}{1 + 2\sqrt{K_m/K_I}}. \tag{2.34}$$
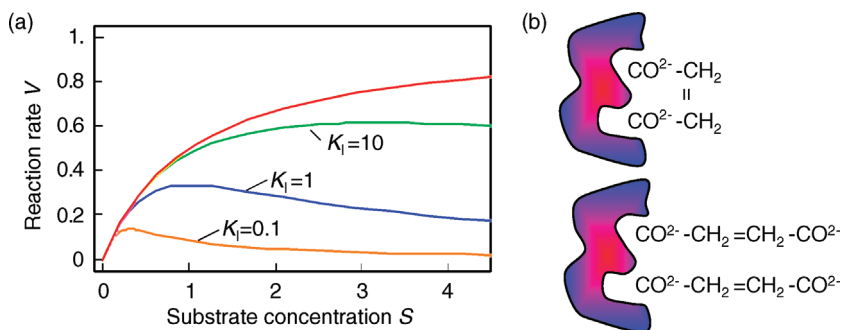
Figure 2.4 Plot of reaction rate v against substrate concentration S for an enzyme with substrate inhibition. The upper curve shows Michaelis–Menten kinetics without inhibition, the lower curves show kinetics for the indicated values of binding constant K_I. Parameter values: $V_{max} = 1$, $K_m = 1$. The left part visualizes a possible mechanism for substrate inhibition: The enzyme (gray item) has two binding pockets to bind different parts of a substrate molecule (upper scheme). In case of high substrate concentration, two different molecules may enter the binding pockets, thereby preventing the specific reaction (lower scheme).

The dependence of v on S is shown in Figure 2.4. A typical example for substrate inhibition is the binding of two succinate molecules to malonate dehydrogenase, which possesses two binding pockets for the carboxyl group. This is schematically represented in Figure 2.4.

2.1.4.2 Binding of Ligands to Proteins

Every molecule that binds to a protein is a ligand, irrespective of whether it is subject of a reaction or not. Below we consider binding to monomer and oligomer proteins. In oligomers, there may be interactions between the binding sites on the subunits.

Consider binding of one ligand (S) to a protein (E) with only one binding site:

$$E + S \rightleftharpoons ES \tag{2.35}$$

The binding constant K_B is given by

$$K_B = \left(\frac{ES}{E \cdot S} \right)_{eq}. \tag{2.36}$$

The reciprocal of K_B is the dissociation constant K_D. The fractional saturation Y of the protein is determined by the number of subunits that have bound ligands, divided by the total number of subunits. The fractional saturation for one subunit is

$$Y = \frac{ES}{E_{total}} = \frac{ES}{ES + E} = \frac{K_B \cdot S}{K_B \cdot S + 1}. \tag{2.37}$$

The plot of Y versus S at constant total enzyme concentration is a hyperbola, like the plot of v versus S in the Michaelis–Menten kinetics (Eq. (2.22)). At a process where the binding of S to E is the first step followed by product release and where the initial concentration of S is much higher than the initial concentration of E, the rate is proportional to the concentration of ES and it holds

$$\frac{v}{V_{\text{max}}} = \frac{ES}{E_{\text{total}}} = Y. \tag{2.38}$$

If the protein has several binding sites, then interactions may occur between these sites, i.e., the affinity to further ligands may change after binding of one or more ligands. This phenomenon is called *cooperativity*. Positive or negative cooperativity denote increase or decrease in the affinity of the protein to a further ligand, respectively. Homotropic or heterotropic cooperativity denotes that the binding to a certain ligand influences the affinity of the protein to a further ligand of the same or another type, respectively.

2.1.4.3 Positive Homotropic Cooperativity and the Hill Equation

Consider a dimeric protein with two identical binding sites. The binding to the first ligand facilitates the binding to the second ligand.

$$\begin{aligned} E_2 + S &\xrightarrow{\text{slow}} E_2 S \\ E_2 S + S &\xrightarrow{\text{fast}} E_2 S_2 \end{aligned} \tag{2.39}$$

where E is the monomer and E_2 is the dimer. The fractional saturation is given by

$$Y = \frac{E_2 S + 2 \cdot E_2 S_2}{2 \cdot E_{2,\text{total}}} = \frac{E_2 S + 2 \cdot E_2 S_2}{2 \cdot E_2 + 2 \cdot E_2 S + 2 \cdot E_2 S_2}. \tag{2.40}$$

If the affinity to the second ligand is strongly increase by binding to the first ligand, then $E_2 S$ will react with S as soon as it is formed and the concentration of $E_2 S$ can be neglected. In the case of complete *cooperativity*, i.e., every protein is either empty or fully bound, Eq. (2.39) reduces to

$$E_2 + 2S \rightarrow E_2 S_2 \tag{2.41}$$

The binding constant reads

$$K_B = \frac{E_2 S_2}{E_2 \cdot S^2}, \tag{2.42}$$

and the fractional saturation is

$$Y = \frac{2 \cdot E_2 S_2}{2 \cdot E_{2,\text{total}}} = \frac{E_2 S_2}{E_2 + E_2 S_2} = \frac{K_B \cdot S^2}{1 + K_B \cdot S^2}. \tag{2.43}$$

Generally, for a protein with n subunits, it holds:

$$v = V_{\text{max}} \cdot Y = \frac{V_{\text{max}} \cdot K_B \cdot S^n}{1 + K_B \cdot S^n}. \tag{2.44}$$

This is the general form of the *Hill equation*. To derive it, we assumed complete homotropic cooperativity. The plot of the fractional saturation Y versus substrate concentration S is a sigmoid curve with the inflection point at $1/K_B$. The quantity n (often "h" is used instead) is termed the *Hill coefficient*.

The derivation of this expression was based on experimental findings concerning the binding of oxygen to hemoglobin (Hb) [13, 14]. In 1904, Bohr *et al.* found that the

plot of the fractional saturation of Hb with oxygen against the oxygen partial pressure had a sigmoid shape. Hill (1913) explained this with interactions between the binding sites located at the Hb subunits [14]. At this time, it was already known that every subunit Hb binds one molecule of oxygen. Hill assumed complete cooperativity and predicted an experimental Hill coefficient of 2.8. Today it is known that Hb has four binding sites, but that the cooperativity is not complete. The sigmoid binding characteristic has the advantage that Hb binds strongly to oxygen in the lung with a high oxygen partial pressure while it can release O_2 easily in the body with low oxygen partial pressure.

2.1.4.4 The Monod–Wyman–Changeux Model for Sigmoid Kinetics

The Monod model [15] explains sigmoid enzyme kinetics by taking into account the interaction of subunits of an enzyme. We will show here the main characteristics and assumptions of this kinetics. The full derivation is given in the web material. It uses the following assumptions: (i) the enzyme consists of n identical subunits, (ii) each subunit can assume an active (R) or an inactive (T) conformation, (iii) all subunits change their conformations at the same time (concerted change), and (iv) the equilibrium between the R and the T conformation is given by an allosteric constant

$$L = \frac{T_0}{R_0}. \tag{2.45}$$

The binding constants for the active and inactive conformations are given by K_R and K_T, respectively. If substrate molecules can only bind to the active form, i.e., if $K_T = 0$, the rate can be expressed as

$$V = \frac{V_{max}K_R S}{(1 + K_R S)} \frac{1}{[1 + \{L/((1 + K_R S)^n)\}]}, \tag{2.46}$$

where the first factor $(V_{max}K_R S)/(1 + K_R S)$ corresponds to the Michaelis–Menten rate expression, while the second factor $[1 + (L/(1 + K_R S)^n)]^{-1}$ is a regulatory factor (Figure 2.5).

For $L = 0$, the plot v versus S is hyperbola as in Michaelis–Menten kinetics. For $L > 0$, we obtain a sigmoid curve shifted to the right. A typical value for the allosteric constant is $L \cong 10^4$.

Up to now we considered in the model of Monod, Wyman, and Changeux only homotropic and positive effects. But this model is also well suited to explain the dependence of the reaction rate on activators and inhibitors. Activators A bind only to the active conformation and inhibitors I bind only to the inactive conformation. This shifts the equilibrium to the respective conformation. Effectively, the binding to effectors changes L:

$$L' = L\frac{(1 + K_I I)^n}{(1 + K_A A)^n}, \tag{2.47}$$

where K_I and K_A denote binding constants. The interaction with effectors is a heterotropic effect. An activator weakens the sigmoidity, while an inhibitor strengthens it.
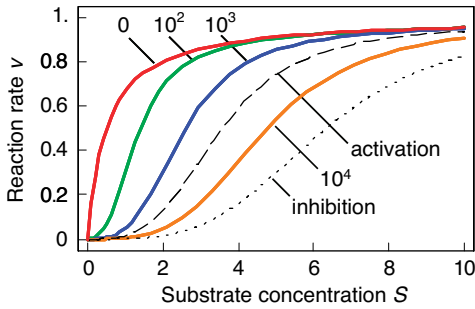
Figure 2.5 Model of Monod, Wyman, and Changeux: Dependence of the reaction rate on substrate concentration for different values of the allosteric constant L, according to equation. The binding constants for the active and inactive conformations are given by K_R and K_T, respectively. If substrate molecules can only bind to the active form, i.e., if $K_T = 0$, the rate can be expressed as

$$V = \frac{V_{max} K_R S}{(1 + K_R S)} \frac{1}{[1 + \{L/((1 + K_R S)^n)\}]}, \quad (2.46).$$

Parameters: $V_{max} = 1$, $n = 4$, $K_R = 2$, $K_T = 0$. The value of L is indicated at the curves. Obviously, increasing value of L causes stronger sigmoidity. The influence of activators or inhibitors (compare Eq. (2.47)) is illustrated with the dotted line for $K_I I = 2$ and with the dashed line for $K_A A = 2$ ($L = 10^4$ in both cases).

A typical example for an enzyme with sigmoid kinetics that can be described with the Monod model is the enzyme phosphofructokinase, which catalyzes the transformation of fructose-6-phosphate and ATP to fructose-1,6-bisphosphate. AMP, NH_4, and K^+ are activators, ATP is an inhibitor.

2.1.5
Generalized Mass Action Kinetics

Mass action kinetics (see Section 2.1.1) has experienced refinements in different ways. The fact that experimental results frequently do not show the linear dependence of rate on concentrations as assumed in mass action laws is acknowledged in power law kinetics used in the S-systems approach [16]. Here, the rate reads

$$\frac{v_j}{v_j^0} = k_j \prod_{i=1}^{n} \left(\frac{S_i}{S_i^0} \right)^{g_{j,i}}, \quad (2.48)$$

where the concentrations S_i and rates v_j are normalized to some standard value denoted by superscript 0, and $g_{i,j}$ is a real number instead of an integer as in Eq. (2.4). The normalization yields dimensionless quantities. The power law kinetics can be considered as a generalization of the mass action rate law. The exponent $g_{i,j}$ is equal to the concentration elasticities, i.e., the scaled derivatives of rates with respect to substrate concentrations (see Section 2.3, Eq. (2.107)). Substrates and effectors (their concentrations both denoted by S_i) enter expression (2.48) in the same formal way, but the respective exponents $g_{i,j}$ will be different. The exponents $g_{i,j}$ will be positive for substrates and activators, but should assume a negative value for inhibitors.

2.1.6
Approximate Kinetic Formats

In metabolic modeling studies, approximate kinetic formats are used (for a recent review, see [17]). They preassume that each reaction rate v_j is proportional to the enzyme concentration E_j. The rates, enzyme concentrations, and substrate concentrations are normalized with respect to a reference state, which is usually a steady state. This leads to the general expression

$$\frac{v_j}{v_j^0} = \frac{E_j}{E_j^0} \cdot f\left(\frac{S}{S^0}, \varepsilon_c^0\right), \tag{2.49}$$

where ε_c is the matrix of concentration elasticities as explained in Section 2.3. One example is the so-called lin-log kinetics

$$\frac{v}{v^0} = \frac{E}{E^0}\left(I + \varepsilon_c^0 \ln\frac{S}{S^0}\right), \tag{2.50}$$

where I is the $r \times r$ identity matrix. Another example is an approximation of the power-law kinetics

$$\ln\frac{v}{v^0} = \ln\frac{E}{E^0} + \varepsilon_c^0 \ln\frac{S}{S^0}. \tag{2.51}$$

Approximative kinetics simplify the determination of model parameters and, especially, of concentration elasticities, since Eq. (2.51) is a set of linear equations in the elasticity coefficients.

2.1.7
Convenience Kinetics

The convenience kinetics [18] has been introduced to ease parameter estimation and to have a kinetic mechanism, where all parameters are independent of each other and not related via the Haldane relation (Eq. (2.28)). It is a generalized form of Michaelis–Menten kinetics that covers all possible stoichiometries, and describes enzyme regulation by activators and inhibitors. For a reaction with stoichiometry

$$n_{-1}S_1 + n_{-2}S_2 + \cdots \leftrightarrow n_{+1}P_1 + n_{+2}P_2 + \cdots, \tag{2.52}$$
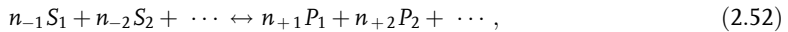
it reads

$$v = E_{\text{total}} \cdot f_{\text{reg}}$$

$$\cdot \frac{k_{\text{cat}}^{\text{for}}\prod_i(S_i/K_{\text{m},S_i})^{n_{-i}} - k_{\text{cat}}^{\text{back}}\prod_j(P_j/K_{\text{m},P_j})^{n_{+j}}}{\prod_i(1+(S_i/K_{\text{m},S_i})+\cdots+(S_i/K_{\text{m},S_i})^{n_{-i}})+\prod_j(1+(P_j/K_{\text{m},P_j})+\cdots+(P_j/K_{\text{m},P_j})^{n_{+j}})-1}, \tag{2.53}$$

with enzyme concentration E_{total} and turnover rates $k_{\text{cat}}^{\text{for}}$ and $k_{\text{cat}}^{\text{back}}$. The regulatory prefactor f_{reg} is either 1 (in case of no regulation) or a product of terms $M/(K_A + M)$ or $1 + M/K_A$ for activators and $K_I/(K_I + M)$ for inhibitors. Activation constants K_A and

inhibition constants K_I are measured in concentration units. M is the concentration of the modifier.

In analogy to Michaelis–Menten kinetics, K_m values denote substrate concentrations, at which the reaction rate is half-maximal if the reaction products are absent; K_I and K_A values denote concentrations, at which the inhibitor or activator has its half-maximal effect. In this respect, many parameters in convenience kinetics are comparable to the kinetic constants measured in enzyme assays. This is important for parameter estimation (see Section 4.2).

To facilitate thermodynamic independence of the parameters, we introduce new system parameters that can be varied independently, without violating any thermodynamic constraints (see Section 2.1.1). For each reaction, we define the velocity constant $K_V = (k_{cat}^{for} \cdot k_{cat}^{back})^{1/2}$ (geometric mean of the turnover rates in both directions). Given the equilibrium and velocity constants, the turnover rates can be written as $k_{cat}^{for} = K_V(K_{eq})^{-1/2}$, $k_{cat}^{back} = K_V(K_{eq})^{1/2}$. The equilibrium constants K_{eq} can be expressed by independent parameters such as the Gibbs free energies of formation: for each substance i, we define the dimensionless energy constant $K_i^G = \exp(G_i(0)/(RT))$ with Boltzmann's gas constant $R = 8.314\,\text{J}\,(\text{mol}^{-1}\,\text{K}^{-1})$ and absolute temperature T. The equilibrium constants then satisfy $\ln K_{eq} = -N^T \ln K^G$.

2.2
Structural Analysis of Biochemical Systems

Summary

We discuss basic structural and dynamic properties of biochemical reaction networks. We introduce a stoichiometric description of networks and learn how moieties and fluxes are balanced within networks.

The *basic elements* of a metabolic or regulatory network model are

1. the compounds with their concentrations or activities and
2. the reactions or transport processes changing the concentrations or activities of the compounds.

In biological environments, reactions are usually catalyzed by enzymes, and transport steps are carried out by transport proteins or pores, thus they can be assigned to identifiable biochemical compounds. In the following, we will mainly refer to metabolic networks. However, the analysis can also be applied to regulatory networks, if different activity states or complexes of regulatory molecules are considered as individual compounds that are converted into each other by modifying reactions.

2.2.1
System Equations

Stoichiometric coefficients denote the proportion of substrate and product molecules involved in a reaction. For example, for the reaction

$$S_1 + S_2 \rightleftharpoons 2P, \tag{2.54}$$

the stoichiometric coefficients of S_1, S_2, and P are -1, -1, and 2, respectively. The assignment of stoichiometric coefficients is not unique. We could also argue that for the production of one mole P, half a mole of each S_1 and S_2 have to be used and, therefore, choose $-1/2$, $-1/2$, and 1. Or, if we change the direction of the reaction, then we may choose 1, 1, and -2.

The change of concentrations in time can be described using ODEs. For the reaction depicted in Eq. (2.54) and the first choice of stoichiometric coefficients, we obtain

$$\frac{dS_1}{dt} = -v, \qquad \frac{dS_2}{dt} = -v, \quad \text{and} \quad \frac{dP}{dt} = 2v. \tag{2.55}$$

This means that the degradation of S_1 with rate v is accompanied by the degradation of S_2 with the same rate and by the production of P with the double rate.

For a metabolic network consisting of m substances and r reactions, the system dynamics is described by the *system equations* (or *balance equations*, since the balance of substrate production and degradation is considered) [19, 20]:

$$\frac{dS_i}{dt} = \sum_{j=1}^{r} n_{ij} v_j \quad \text{for} \quad i = 1, \dots, m. \tag{2.56}$$

The quantities n_{ij} are the stoichiometric coefficients of the ith metabolite in the jth reaction. Here, we assume that the reactions are the only reason for concentration changes and that no mass flow occurs due to convection or to diffusion. The balance equations (2.56) can also be applied, if the system consists of several compartments. In this case, every compound in different compartments has to be considered as an individual compound and transport steps are formally considered as reactions transferring the compound belonging to one compartment into the same compound belonging to the other compartment. In case, volume differences must be considered (see Section 3.4).
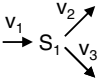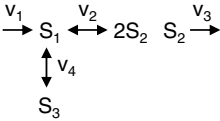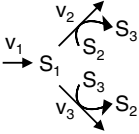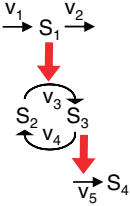
The stoichiometric coefficients n_{ij} assigned to the compounds S_i and the reactions v_j can be comprehended into the *stoichiometric matrix*

$$N = \{n_{ij}\} \quad \text{for} \quad i = 1, \dots, m \quad \text{and} \quad j = 1, \dots, r, \tag{2.57}$$

where each column belongs to a reaction and each row to a compound. Table 2.4 shows some examples for reaction networks and their respective stoichiometric matrices.

Note that all reactions may be reversible. In order to determine the signs is N, the direction of the arrows is artificially assigned as positive "from left to right" and "from top down." If the net flow of a reaction proceeds in the opposite direction as the arrow indicates, the value of rate v is negative.

Table 2.4 Different reaction networks and their stoichiometric matrices[a].

	Network	Stoichiometric matrix
N1	$S_1 + S_2 + S_3 \overset{v_1}{\longleftrightarrow} S_4 + 2S_5$	$N = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 2 \end{pmatrix}$
N2	$\overset{v_1}{\longrightarrow} S_1 \overset{v_2}{\longrightarrow} S_2 \overset{v_3}{\longrightarrow} S_3 \overset{v_4}{\longrightarrow} S_4 \overset{v_5}{\longrightarrow}$	$N = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$
N3	$\overset{v_1}{\longrightarrow} S_1 \overset{v_2}{\nearrow} \underset{v_3}{\searrow}$	$N = \begin{pmatrix} 1 & -1 & -1 \end{pmatrix}$
N4	$\overset{v_1}{\longrightarrow} S_1 \overset{v_2}{\longleftrightarrow} 2S_2 \quad S_2 \overset{v_3}{\longrightarrow}$ $\updownarrow v_4$ S_3	$N = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
N5	$\overset{v_1}{\longrightarrow} S_1 \overset{v_2}{\nearrow} S_2 \nearrow S_3$ $\underset{v_3}{\searrow} S_2$	$N = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$
N6	$\overset{v_1}{\longrightarrow} S_1 \overset{v_2}{\longrightarrow}$ $S_2 \overset{v_3}{\underset{v_4}{\rightleftarrows}} S_3$ $\overset{v_5}{\longrightarrow} S_4$	$N = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

[a]Note that external metabolites are neither drawn in the network nor included in the stoichiometric matrix. Thin arrows denote reactions, bold arrows denote activation.

Altogether, the mathematical description of the metabolic system consists of a vector $S = (S_1, S_2, S_n)^T$ of concentrations values, a vector $v = (v_1, v_2, ..., v_r)^T$ of reaction rates, a parameter vector $p = (p_1, p_2, ..., p_m)^T$, and the stoichiometric matrix N. If the system is in steady state, we can also consider the vector $J = (J_1, J_2, ..., J_r)^T$ containing the steady-state fluxes. With these notions, the balance equation reads

$$\frac{dS}{dt} = Nv, \tag{2.58}$$

a compact form that is suited for various types of analysis.

2.2.2
Information Encoded in the Stoichiometric Matrix *N*

The stoichiometric matrix contains important information about the structure of the metabolic network. Using the stoichiometric matrix, we may calculate which combinations of individual fluxes are possible in steady state (i.e., calculate the admissible steady-state flux space). We may easily find out dead ends and unbranched reaction pathways. In addition, we may find out the conservation relations for the included reactants.

In steady state, it holds that

$$\frac{\mathrm{d}S}{\mathrm{d}t} = Nv = 0. \tag{2.59}$$

The right equality sign denotes a linear equation system for determination of the rates *v*. From linear algebra, it is known that this equation has nontrivial solutions only for Rank $N < r$. A kernel matrix *K* fulfilling

$$NK = 0 \tag{2.60}$$

shows the respective linear dependencies [21]. The choice of the kernel is not unique. It can be determined using the Gauss Algorithm (see mathematical textbooks). It contains as columns *r*–Rank *N* basis vectors. Every possible set *J* of steady-state fluxes can be expressed as linear combination of the columns k_i of *K*

$$J = \sum_{i=1}^{r - \text{Rank } N} \alpha_i \cdot k_i. \tag{2.61}$$

The coefficients must have units corresponding to the units of reaction rates ($\mathrm{M\,s^{-1}}$ or $\mathrm{mol\,l^{-1}\,s^{-1}}$).

Example 2.2

For the network N2 in Table 2.4, we have $r = 5$ reactions and Rank $N = 4$. The kernel matrix contains just $1 = 5 - 4$ basis vectors, which are multiples of $k = (1 \quad 1 \quad 1 \quad 1 \quad 1)^{\mathsf{T}}$. This means that in steady state, the flux through all reactions must be equal. Network N3 comprises $r = 3$ reactions and has Rank $N = 1$. Each representation of the kernel matrix contains $3 - 1 = 2$ basis vectors, e.g.,

$$K = (k_1 \quad k_2) \quad \text{with} \quad k_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad k_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tag{2.62}$$

and for the steady-state flux holds

$$J = \alpha_1 \cdot k_1 + \alpha_2 \cdot k_2. \tag{2.63}$$

Network N6 can present a small signaling cascade. It has five reactions and Rank $N = 3$. The resulting two basis vectors of the kernel are linear combinations of

$$k_1 = (1 \quad 1 \quad 0 \quad 0 \quad 0)^T, \quad k_2 = (0 \quad 0 \quad 1 \quad 1 \quad 0)^T. \tag{2.64}$$

If we calculate the possible steady-state fluxes according to Eq. (2.63), we can easily see that in every steady state, it holds that production and degradation of S_1 are balanced ($J_1 = J_2$) and that the fluxes through the cycle are equal ($J_3 = J_4$). In addition, J_5 must be equal to zero, otherwise S_4 would accumulate. One could prevent the last effect by also including the degradation of S_4 into the network.

If the entries in a certain row are zero in all basis vectors, we have found an equilibrium reaction. In any steady state, the net rate of this reaction must be zero. For the reaction system N4 in Table 2.4, it holds that $r = 4$ and Rank $N = 3$. Its kernel consists of only one column $K = (1 \quad 1 \quad 1 \quad 0)^T$. Hence, $v_4 = \sum_{i=1}^{1} \alpha \cdot 0 = 0$. In any steady state, the rates of production and degradation of S_3 must equal.

If all basis vectors contain the same entries for a set of rows, this indicates an unbranched reaction path. In each steady state, the net rate of all respective reactions is equal.

Example 2.3

Consider the reaction scheme

$$(2.65)$$
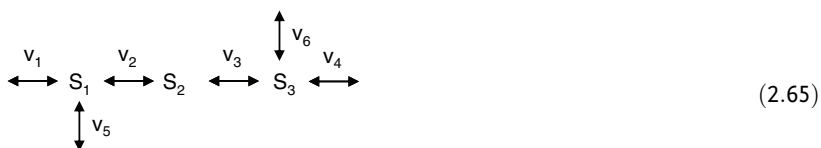
The system comprises $r = 6$ reactions. The stoichiometric matrix reads

$$N = \begin{pmatrix} 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{pmatrix}$$

with Rank $N = 3$. Thus, the kernel matrix is spanned by three basis vectors, for example, $k_1 = (1 \quad 1 \quad 1 \quad 0 \quad 0 \quad -1)^T$, $k_2 = (1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0)^T$, and $k_3 = (-1 \quad -1 \quad -1 \quad -1 \quad 0 \quad 0)^T$. The entries for the second and third reactions are always equal, thus in any steady state, the fluxes through reactions 2 and 3 must be equal.

Up to now, we have not been concerned about (ir)reversibility of reactions in the network. If a certain reaction is considered irreversible, this has no consequences for the stoichiometric matrix N but rather for the kernel K. The set of vectors belonging to K is restricted by the condition that some values may not become negative (or positive – depending on the definition of flux direction).

2.2.3
Elementary Flux Modes and Extreme Pathways

The definition of the term "pathway" in a metabolic network is not straightforward. A descriptive definition of a pathway is a set of subsequent reactions that are linked by common metabolites. Typical examples include glycolysis or different amino acid synthesis pathways. More detailed inspection of metabolic maps like the *Boehringer Chart* [22] shows that metabolism is highly interconnected. Pathways that are known for a long time from biochemical experience are already hard to recognize, and it is even harder to find out new pathways, for example in metabolic maps that have been reconstructed from sequenced genomes of bacteria.

This problem has been elaborated in the concept of *elementary flux modes* [21, 23–27]. Here, the stoichiometry of a metabolic network is investigated to find out which direct routes are possible that lead from one external metabolite to another external metabolite. The approach takes into account that some reactions are reversible, while others are irreversible.

A *flux mode* M is set of flux vectors that represent such direct routes through the metabolic networks. In mathematical terms, it is defined as the set

$$M = \{v \in R^r | v = \lambda v^*, \lambda > 0\}, \tag{2.66}$$

where v^* is an r-dimensional vector (unequal to the null vector) fulfilling two conditions: (i) steady state, i.e., Eq. (2.59), and (ii) sign restriction, i.e., the flux directions in v^* fulfill the prescribed irreversibility relations.

A flux mode M comprising v is called reversible if the set M' comprising $-v$ is also a flux mode. A flux mode is an elementary flux mode if it uses a minimal set of reactions and cannot be further decomposed, i.e., the vector v cannot be represented as nonnegative linear combination of two vectors that fulfill conditions (i) and (ii) but contain more zero entries than v. An elementary flux mode is a minimal set of enzymes that could operate at steady state, with all the irreversible reactions used in the appropriate direction. The number of elementary flux modes is at least as high as the number of basis vectors of the null space.

Example 2.4

The systems (A) and (B) differ by the (ir)reversibility of reaction 2.
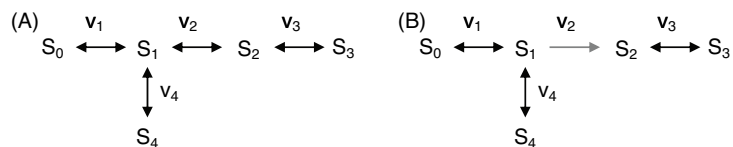
$$\tag{2.67}$$

The elementary flux modes connect the external metabolites S_0 and S_3, S_0 and S_4, or S_3 and S_4. The stoichiometric matrix and the flux modes read for case (A) and case (B)

$$N = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \quad \boldsymbol{v}^A = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix},$$

$$\text{and } \boldsymbol{v}^B = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

$$(2.68)$$

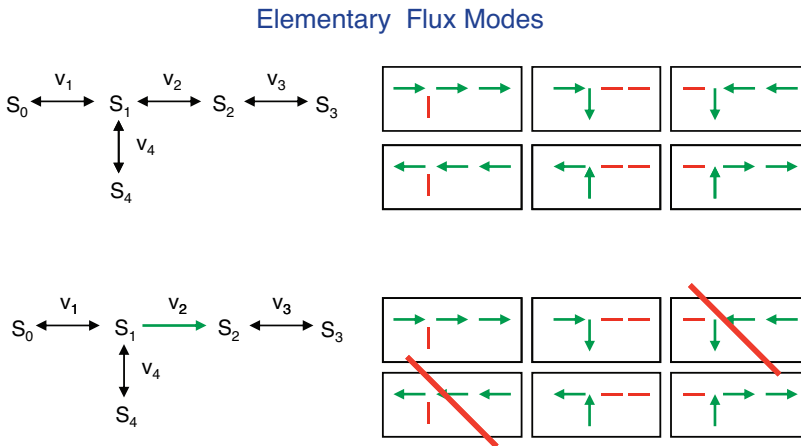The possible routes are illustrated in Figure 2.6.

Elementary Flux Modes

Figure 2.6 Schematic representation of elementary flux modes for the reaction network depicted in Eq. (2.67).

2.2.3.1 Flux Cone

The stoichiometric analysis of biochemical network analysis can be modified by considering only irreversible reactions (e.g., by splitting reversible reactions into two irreversible ones). Based on such a unidirectional representation, the basis vectors (Eq. (2.61)) form a convex cone in the flux space. This mapping relates stoichiometric analysis to the concepts of convex geometry as follows. The steady-state assumption requires that a flux vector is an element of the null space of the stoichiometry matrix N spanned by matrix K. A row of K can be interpreted as a hyperplane in flux space. The intersection of all these hyperplanes forms the null space. From thermodynamic
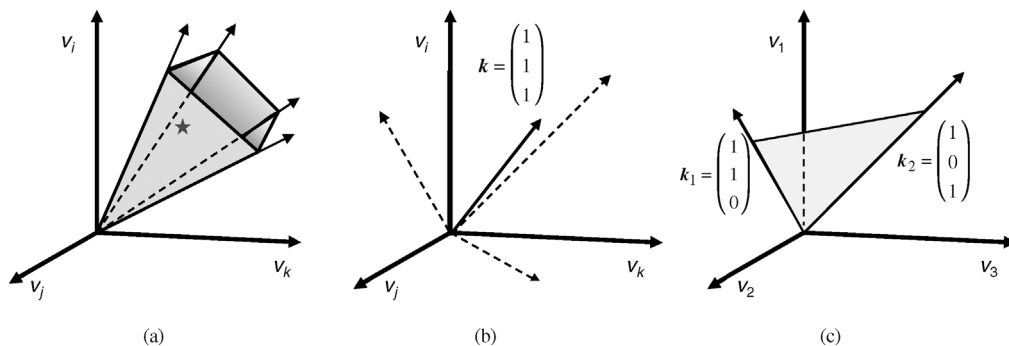
(a) (b) (c)

Figure 2.7 Flux cone: schematic representation of the subspace of feasible steady states within the space spanned by all positive-valued vectors for rates of irreversible reactions, v_i, $i = 1, \ldots, r$. Only three dimensions are shown. Feasible solutions are linear combinations of basis vectors of matrix K (see text). (a) Illustrative representation of the flux cone for a higher dimensional system (with r–Rank (N) $= 4$)). The basis vectors of K are rays starting at the origin. The line connecting the four rays indicates possible limits for real flux distributions set by constraints. The little star indicates one special feasible solution for the fluxes. (b) The flux cone for an unbranched reaction chain of arbitrary length, such as the network N2 in Table 2.4, is just a ray since K is represented by a single basis vector containing only 1s. (c) The flux cone for network N3 in Table 2.4 is the plane spanned by the basis vectors $k_1 = (1 \quad 1 \quad 0)^T$, $k_2 = (1 \quad 0 \quad 1)^T$.

considerations, some of the reactions can be assumed to proceed only in one direction so that the backward reaction can be neglected. Provided that all reactions are unidirectional or irreversible, the intersection of the null space with the semipositive orthant of the flux space forms a polyhedral cone, the flux cone. The intersection procedure results in a set of rays or edges starting at 0, which fully describe the cone. The edges are represented by vectors and any admissible steady state of the system is a positive combination of these vectors. An illustration is presented in Figure 2.7.

The set of elementary flux modes is uniquely defined. Pfeiffer *et al.* [23] developed a software ("Metatool") to calculate the elementary flux modes for metabolic networks. The concept of *extreme pathways* [28–30] is analogous to the concept of elementary flux modes, but here all reactions are constrained by flux directionality, while the concept of elementary flux modes allows for reversible reactions. To achieve this, reversible reactions are broken down into their forward and backward components. This way, the set of extreme pathways is a subset of the set of elementary flux modes and the extreme pathways are systemically independent.

Elementary flux modes and extreme pathways can be used to understand the range of metabolic pathways in a network, to test a set of enzymes for production of a desired product and detect nonredundant pathways, to reconstruct metabolism from annotated genome sequences and analyze the effect of enzyme deficiency, to reduce drug effects, and to identify drug targets. A specific application, the flux balance analysis, will be explained in Section 8.1.

2.2.4
Conservation Relations: Null Space of N^T

If a substance is neither added to nor removed from the reaction system (neither produced nor degraded), its total concentration remains constant. This also holds if the substance interacts with other compounds by forming complexes. We have seen already as an example the constancy of the total enzyme concentration (Eq. (2.19)) when deriving the Michaelis–Menten rate equation. This was based on the assumption that enzyme production and degradation takes place on a much faster timescale than the catalyzed reaction.

For the mathematical derivation of the conservation relations [21], we consider a matrix G fulfilling

$$GN = 0. \tag{2.69}$$

Due to Eq. (2.58), it follows

$$G\dot{S} = GNv = 0. \tag{2.70}$$

Integrating this equation leads directly to the conservation relations

$$GS = \text{constant}. \tag{2.71}$$

The number of independent rows of G is equal to n–Rank N, where n is the number of metabolites in the system. G^T is the kernel matrix of N^T, hence it has similar properties as K. Matrix G can also be found using the Gauss algorithm. It is not unique, but every linear combination of its rows is again a valid solution. There is a simplest representation $G = (\,G_0 \quad I_{n-\text{Rank }N}\,)$. Finding this representation may be helpful for a simple statement of conservation relations, but this may necessitate renumbering and reordering of metabolite concentrations (see below).

Example 2.5

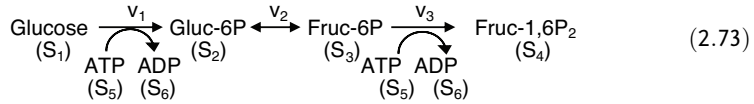Consider a set of two reactions comprising a kinase and a phosphatase reaction

$$\text{ATP} \quad \text{ADP} \tag{2.72}$$
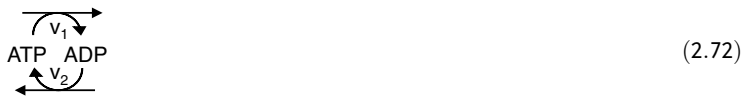
The metabolite concentration vector reads $S = (\,ATP \quad ADP\,)^T$, the stoichiometric matrix is $N = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ yielding $G = (\,1 \quad 1\,)$. From the condition $GS = \text{con-}$stant, it follows $ATP + ADP = \text{constant}$. Thus, we have a conservation of adenine

nucleotides in this system. The actual values of $ATP + ADP$ must be determined from the initial conditions.

Example 2.6

For the following model of the upper part of glycolysis

$$\text{Glucose} \xrightarrow{v_1} \text{Gluc-6P} \xleftrightarrow{v_2} \text{Fruc-6P} \xrightarrow{v_3} \text{Fruc-1,6P}_2 \qquad (2.73)$$
$$(S_1) \quad \text{ATP ADP} \quad (S_2) \quad (S_3) \quad \text{ATP ADP} \quad (S_4)$$
$$(S_5) \quad (S_6) \qquad\qquad (S_5) \quad (S_6)$$

the stoichiometric matrix N (note the transpose!) and a possible representation of the conservation matrix G are given by

$$N^{\mathrm{T}} = \begin{pmatrix} -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}.$$
$$(2.74)$$

The interpretation of the second and third row is straightforward, showing the conservation of adenine nucleotides (g_2, $ADP + ATP =$ constant) and the conservation of sugars (g_3), respectively. The interpretation of the first row is less intuitive. If we construct the linear combination $g_4 = -g_1 + 3 \cdot g_2 + 2 \cdot g_3 = (0 \ 1 \ 1 \ 2 \ 3 \ 2)$, we find the conservation of phosphate groups.

Importantly, conservation relations can be used to simplify the system of differential equations $\dot{S} = Nv$ describing the dynamics of our reaction system. The idea is to eliminate linear dependent differential equations and to replace them by appropriate algebraic equations. Below the procedure is explained systematically [20].

First we have to rearrange the rows in the stoichiometric matrix N as well as in the concentration vector S such that a set of independent rows is on top and the dependent rows are at the bottom. Then the matrix N is split into the independent part N_{R} and the dependent part N' and a *link matrix* L is introduced in the following way:

$$N = \begin{pmatrix} N_{\mathrm{R}} \\ N' \end{pmatrix} = LN_{\mathrm{R}} = \begin{pmatrix} I_{\mathrm{Rank}\,N} \\ L' \end{pmatrix} N_{\mathrm{R}}. \qquad (2.75)$$

$I_{\mathrm{Rank}\,N}$ is the identity matrix of size Rank N. The differential equation system may be rewritten accordingly

$$\dot{S} = \begin{pmatrix} \dot{S}_{\mathrm{indep}} \\ \dot{S}_{\mathrm{dep}} \end{pmatrix} = \begin{pmatrix} I_{\mathrm{Rank}\,N} \\ L' \end{pmatrix} N_{\mathrm{R}} v, \qquad (2.76)$$

and the dependent concentrations fulfil

$$\dot{S}_{\text{dep}} = L' \cdot \dot{S}_{\text{indep}}. \tag{2.77}$$

Integration leads to

$$S_{\text{dep}} = L' \cdot S_{\text{indep}} + \text{constant}. \tag{2.78}$$

This relation is fulfilled during the entire time course. Thus, we may replace the original system by a reduced differential equation system

$$\dot{S}_{\text{indep}} = N_R v, \tag{2.79}$$

supplemented with the set of algebraic equations (2.78).

Example 2.7

For the reaction system,

$$\tag{2.80}$$

the stoichiometric matrix, the reduced stoichiometric matrix, and the link matrix read

$$N = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \quad N_R = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix},$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \quad L' = \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}$$

The conservation relation $S_3 + S_4 =$ constant is expressed by $G = (0 \quad 0 \quad 1 \quad 1)$. The ODE system

$$\begin{aligned} \dot{S}_1 &= v_1 - v_2 \\ \dot{S}_2 &= v_2 - v_3 \\ \dot{S}_3 &= v_4 - v_2 \\ \dot{S}_4 &= v_2 - v_4 \end{aligned}$$

can be replaced by the differential-algebraic system

$$
\begin{aligned}
\dot{S}_1 &= v_1 - v_2 \\
\dot{S}_2 &= v_2 - v_3 \\
\dot{S}_3 &= v_4 - v_2 \\
S_3 + S_4 &= \text{constant}
\end{aligned}
\quad ,
$$

which has one differential equation less.

Eukaryotic cells contain a variety of organelles like nucleus, mitochondria, or vacuoles, which are separated by membranes. Reaction pathways may cross the compartment boundaries. If a substance S occurs in two different compartments, e.g., in the cytosol and in mitochondria, the respective concentrations can be assigned to two different variables, S^{C1} and S^{C2}. Formally, the transport across the membrane can be considered as a reaction with rate v. It is important to note that both compartments have different volumes V^{C1} and V^{C2}. Thus, transport of a certain amount of S with rate v from compartment $C1$ into the compartment $C2$ changes the concentrations differently:

$$
V^{C1} \cdot \frac{\mathrm{d}}{\mathrm{d}t} S^{C1} = -v \quad \text{and} \quad V^{C2} \cdot \frac{\mathrm{d}}{\mathrm{d}t} S^{C2} = v, \tag{2.81}
$$

where $V \cdot S$ denotes substance amount in moles. Compartmental models are discussed in more detail in Section 3.4.

2.3
Kinetic Models of Biochemical Systems

Summary

An important problem in the modeling of biological systems is to characterize the dependence of certain properties on time and space. One frequently applied strategy is the description of the change of state variables by differential equations. If only temporal changes are considered, ODEs are used. For changes in time and space, partial differential equations (PDEs) are appropriate. In this chapter, we will deal with the solution, analysis, a numerical integration of ODEs, and with basic concepts of dynamical systems theory as state space, trajectory, steady states, and stability.

2.3.1
Describing Dynamics with ODEs

The time behavior of biological systems in a deterministic approach can be described by a set of differential equations

$$
\frac{\mathrm{d}x_i}{\mathrm{d}t} = \dot{x}_i = f_i(x_1, \ldots, x_n, p_1, \ldots, p_l, t) \quad i = 1, \ldots, n, \tag{2.82}
$$

where x_i are the variables, e.g., concentrations, and p_j are the parameters, e.g., enzyme concentrations or kinetic constants, and t is the time. We will use the notions dx/dt and $\dot{x}$ interchangeably. In vector notation, Eq. (2.82) reads

$$\frac{d}{dt}x = \dot{x} = f(x, p, t), \tag{2.83}$$

with $x = (x_1, \ldots, x_n)^T$, $f = (f_1, \ldots, f_n)^T$, and $p = (p_1, \ldots, p_l)^T$. For biochemical reaction systems, the functions f_i are frequently given by the contribution of producing and degrading reactions as described for the balance equations in Section 1.2.

2.3.1.1 Notations

ODEs depend on one variable (e.g., time t). Otherwise, they are called PDEs. PDEs are not considered here.

An implicit ODE

$$F(t, x, x', \ldots, x^{(n)}) = 0 \tag{2.84}$$

includes the variable t, the unknown function x, and its derivatives up to nth order. An explicit ODE of nth order has the form

$$x^{(n)} = f(t, x, x', \ldots, x^{(n-1)}). \tag{2.85}$$

The highest derivative (here n) determines the order of the ODE.

Studying the time behavior of our system, we may be interested in finding solutions of the ODE, i.e., finding an n times differentiable function x fulfilling Eq. (2.85). Such a solution may depend on parameters, so-called integration constants, and represents a set of curves. A solution of an ODE of nth order depending on n integration parameters is a *general* solution. Specifying the integration constants, for example, by specifying n initial conditions (for $n = 1$: $x(t = 0) = x^0$) leads to a special or *particular* solution.

We will not show here all possibilities of solving ODEs, instead we will focus on some cases relevant for the following chapters.

If the right-hand sides of the ODEs are not explicitly dependent on time t ($\dot{x} = f(x, p)$), the system is called autonomous. Otherwise it is nonautonomous. This case will not be considered here.

The system state is a snapshot of the system at a given time that contains enough information to predict the behavior of the system for all future times. The state of the system is described by the set of variables. The set of all possible states is the state space. The number n of independent variables is equal to the dimension of the state space. For $n = 2$, the two-dimensional state space can be called phase plane.

A particular solution of the ODE system $\dot{x} = f(x, p, t)$, determined from the general solution by specifying parameter values p and initial conditions $x(t_0) = x^0$, describes a path through the state space and is called trajectory.

Stationary states or steady states are points $\bar{x}$ in the phase plane, where the condition $\dot{x} = 0 (\dot{x}_1 = 0, \ldots, \dot{x}_n = 0)$ is met. At steady state, the system of n differential equations is represented by a system of n algebraic equations for n variables.

The equation system $\dot{x} = 0$ can have multiple solutions referring to multiple steady states. The change of number or stability of steady states upon changes of parameter values p is called a bifurcation.

Linear systems of ODEs have linear functions of the variables as right-hand sides, such as

$$
\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + z_1 \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + z_2
\end{aligned}
\tag{2.86}
$$

or in general $\dot{x} = Ax + z$. The matrix $A = \{a_{ik}\}$ is the system matrix containing the system coefficients $a_{ik} = a_{ik}(p)$ and the vector $z = (z_1, \ldots, Z_n)^{\mathrm{T}}$ contains inhomogeneities. The linear system is *homogeneous* if $z = 0$ holds. Linear systems can be solved analytically. Although in real-world problems, the functions are usually nonlinear, linear systems are important as linear approximations in the investigation of steady states.

Example 2.8

The simple linear system

$$
\frac{dx_1}{dt} = a_{12}x_2, \qquad \frac{dx_2}{dt} = -x_1
\tag{2.87}
$$

has the general solution

$$
x_1 = \frac{1}{2}e^{-i\sqrt{a_{12}}t}(1 + e^{2i\sqrt{a_{12}}t})C_1 - \frac{1}{2}ie^{-i\sqrt{a_{12}}t}(-1 + e^{2i\sqrt{a_{12}}t})\sqrt{a_{12}}C_2
$$

$$
x_2 = \frac{i}{2\sqrt{a_{12}}}e^{-i\sqrt{a_{12}}t}(1 + e^{2i\sqrt{a_{12}}t})C_1 + \frac{1}{2}e^{-i\sqrt{a_{12}}t}(1 + e^{2i\sqrt{a_{12}}t})C_2
$$

with the integration constants C_1 and C_2. Choosing $a_{12} = 1$ simplifies the system to $x_1 = C_1 \cos t + C_2 \sin t$ and $x_2 = C_2 \cos t - C_1 \sin t$. Specification of the initial conditions to $x_1(0) = 2$, $x_2(0) = 1$ gives the particular solution $x_1 = 2\cos t + \sin t$ and $x_2 = \cos t - 2\sin t$. The solution can be presented in the phase plane or directly as functions of time (Figure 2.8).

2.3.1.2 Linearization of Autonomous Systems

In order to investigate the behavior of a system close to steady state, it may be useful to linearize it. Considering the deviation $\hat{x}(t)$ from steady state with $x(t) = \bar{x} + \hat{x}(t)$, it follows

$$
\dot{x} = f(\bar{x} + \hat{x}(t)) = \frac{d}{dt}(\bar{x} + \hat{x}(t)) = \frac{d}{dt}\hat{x}(t).
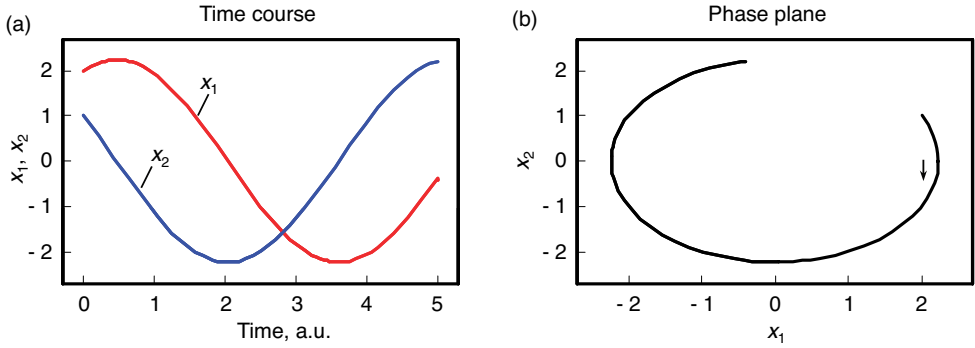\tag{2.88}
$$

Figure 2.8 Phase plane and time course for the linear system of ODEs represented in Eq. (2.87). In time course panel: gray line $x_1(t)$, black line $x_2(t)$. Parameters: $a_{12} = 1$, $x_1(0) = 1$, $x_2(0) = 2$.

Taylor expansion of the temporal change of the deviation, $(\mathrm{d}/\mathrm{d}t)\hat{x}_i = f_i(\bar{x}_1 + \hat{x}_1, \ldots, \bar{x}_n + \hat{x}_n)$, gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_i = f_i(\bar{x}_1, \ldots, \bar{x}_n) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\hat{x}_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 f_i}{\partial x_j \partial x_k}\hat{x}_j\hat{x}_k + \cdots. \qquad (2.89)$$

Since we consider steady state, it holds $f_i(\bar{x}_1, \ldots, \bar{x}_n) = 0$. Neglecting terms of higher order, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x}_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\hat{x}_j = \sum_{j=1}^{n} a_{ij}\hat{x}_j. \qquad (2.90)$$

The coefficients $a_{ij} = \partial f_i/\partial x_j$ are calculated at steady state and are constant. They form the so-called *Jacobian* matrix:

$$J = \{a_{ij}\} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{pmatrix}. \qquad (2.91)$$

For linear systems, it holds $J = A$.

2.3.1.3 Solution of Linear ODE Systems

We are interested in two different types of problems: describing the temporal evolution of the system and finding its steady state. The problem of finding the steady state $\bar{x}$ of a linear ODE system, $\dot{x} = 0$, implies solution of $A\bar{x} + z = 0$. The problem can be solved by inversion of the system matrix A:

$$\bar{x} = -A^{-1}z. \qquad (2.92)$$

The time course solution of homogeneous linear ODEs is described in the following. The systems can be solved with an exponential function as ansatz. In the simplest case $n = 1$, we have

$$\frac{dx_1}{dt} = a_{11} x_1. \tag{2.93}$$

Introducing the ansatz $x_1(t) = b_1 e^{\lambda t}$ with constant b_1 into Eq. (2.93) yields $b_1 \lambda e^{\lambda t} = a_{11} b_1 e^{\lambda t}$, which is true, if $\lambda = a_{11}$. This leads to the general solution $x_1(t) = b_1 e^{a_{11}t}$. To find a particular solution, we must specify the initial conditions $x_1(t = 0) = x_1^{(0)} = b_1 e^{a_{11}t}|_{t=0} = b_1$. Thus, the solution is

$$x_1(t) = x_1^{(0)} e^{a_{11}t}. \tag{2.94}$$

For a linear homogeneous system of n differential equations, $\dot{x} = Ax$, the approach is $x = b e^{\lambda t}$. This gives $\dot{x} = b \lambda e^{\lambda t} = A b e^{\lambda t}$. The scalar factor $e^{\lambda t}$ can be cancelled out, leading to $b\lambda = Ab$ or the characteristic equation

$$(A - \lambda I_n) b = 0. \tag{2.95}$$

For homogeneous linear ODE systems, the *superposition principle* holds: if x_1 and x_2 are solutions of this ODE system, then also their linear combination is a solution. This leads to the general solution of the homogeneous linear ODE system:

$$x(t) = \sum_{i=1}^{n} c_i b^{(i)} e^{\lambda_i t}, \tag{2.96}$$

where $b^{(i)}$ is the eigenvectors of the system matrix A corresponding to the eigenvalues λ_i. A particular solution specifying the coefficients c_i can be found considering the initial conditions $x(t = 0) = x^{(0)} = \sum_{j=1}^{n} c_i b^{(i)}$. This constitutes an inhomogeneous linear equation system to be solved for c_i.

For the solution of inhomogeneous linear ODEs, the system $\dot{x} = Ax + z$ can be transformed into a homogeneous system by the coordination transformation $\hat{x} = x - \bar{x}$. Since $(d/dt)\bar{x} = A\bar{x} + z = 0$, it holds $(d/dt)\hat{x} = A\hat{x}$. Therefore, we can use the solution algorithm for homogeneous systems for the transformed system.

2.3.1.4 Stability of Steady States

If a system is at steady state, it should stay there – until an external perturbation occurs. Depending on the system behavior after perturbation, steady states are either

- *stable* – the system returns to this state
- *unstable* – the system leaves this state
- *metastable* – the system behavior is indifferent

A steady state is *asymptotically* stable, if it is stable and solutions based on nearby initial conditions tend to this state for $t \to \infty$. *Local* stability describes the behavior after small perturbations, *global* stability after any perturbation.

To investigate, whether a steady state $\bar{x}$ of the ODE system $\dot{x} = f(x)$ is asymptotically stable, we consider the linearized system $d\hat{x}/dt = A\hat{x}$ with $\hat{x}(t) = x(t) - \bar{x}$. The steady state $\bar{x}$ is asymptotically stable, if the Jacobian A has n eigenvalues with strictly

negative real parts each. The steady state is unstable, if at least one eigenvalue has a positive real part. This will now be explained in more detail for 1- and 2D systems.

We start with 1D systems, i.e., $n = 1$, and assume without loss of generality $\bar{x}_1 = 0$. The system $\dot{x}_1 = f_1(x_1)$ yields the linearized system $\dot{x}_1 = (\partial f_1 / \partial x_1)|_{\bar{x}_1} x_1 = a_{11} x_1$. The Jacobian matrix $A = \{a_{11}\}$ has only one eigenvalue $\lambda_1 = a_{11}$. The solution is $x_1(t) = x_1^{(0)} e^{\lambda_1 t}$. It is obvious that $e^{\lambda_1 t}$ increases for $\lambda_1 > 0$ and the system runs away from the steady state. For $\lambda_1 < 0$, the deviation from steady state decreases and $x_1(t) \to \bar{x}_1$ for $t \to \infty$. For $\lambda_1 = 0$, consideration of the linearized system allows no conclusion about stability of the original system because higher order terms in Eq. (2.89) play a role.

Consider the 2D case, $n = 2$. For the general (linear or nonlinear) system

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned}, \tag{2.97}$$

we can compute the linearized system

$$\begin{aligned} \dot{x}_1 &= \frac{\partial f_1}{\partial x_1}\bigg|_{\bar{x}} x_1 + \frac{\partial f_1}{\partial x_2}\bigg|_{\bar{x}} x_2 \\ \dot{x}_2 &= \frac{\partial f_2}{\partial x_1}\bigg|_{\bar{x}} x_1 + \frac{\partial f_2}{\partial x_2}\bigg|_{\bar{x}} x_2 \end{aligned} \quad \text{or } \dot{x} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}\bigg|_{\bar{x}} \dfrac{\partial f_1}{\partial x_2}\bigg|_{\bar{x}} \\ \dfrac{\partial f_2}{\partial x_1}\bigg|_{\bar{x}} \dfrac{\partial f_2}{\partial x_2}\bigg|_{\bar{x}} \end{pmatrix} x = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} x = Ax. \tag{2.98}$$

To find the eigenvalues of A, we have to solve the characteristic polynomial

$$\lambda^2 - \underbrace{(a_{11} + a_{22})}_{\text{Tr } A} \lambda + \underbrace{a_{11}a_{22} - a_{12}a_{21}}_{\text{Det } A} = 0, \tag{2.99}$$

with Tr A the trace and Det A the determinant of A, and get

$$\lambda_{1/2} = \frac{\text{Tr } A}{2} \pm \sqrt{\frac{(\text{Tr } A)^2}{4} - \text{Det } A}. \tag{2.100}$$

The eigenvalues are either real for $(\text{Tr } A)^2 / 4 - \text{Det } A \geq 0$ or complex (otherwise). For complex eigenvalues, the solution contains oscillatory parts.

For stability, it is necessary that Tr $A < 0$ and Det $A \geq 0$. Depending on the sign of the eigenvalues, steady states of a 2D system may have the following characteristics:

1. $\lambda_1 < 0$, $\lambda_2 < 0$, both real: stable node
2. $\lambda_1 > 0$, $\lambda_2 > 0$, both real: unstable node
3. $\lambda_1 > 0$, $\lambda_2 < 0$, both real: saddle point, unstable
4. $Re(\lambda_1) < 0$, $Re(\lambda_2) < 0$, both complex with negative real parts: stable focus
5. $Re(\lambda_1) > 0$, $Re(\lambda_2) > 0$, both complex with positive real parts: unstable focus
6. $Re(\lambda_1) = Re(\lambda_2) = 0$, both complex with zero real parts: center, unstable.
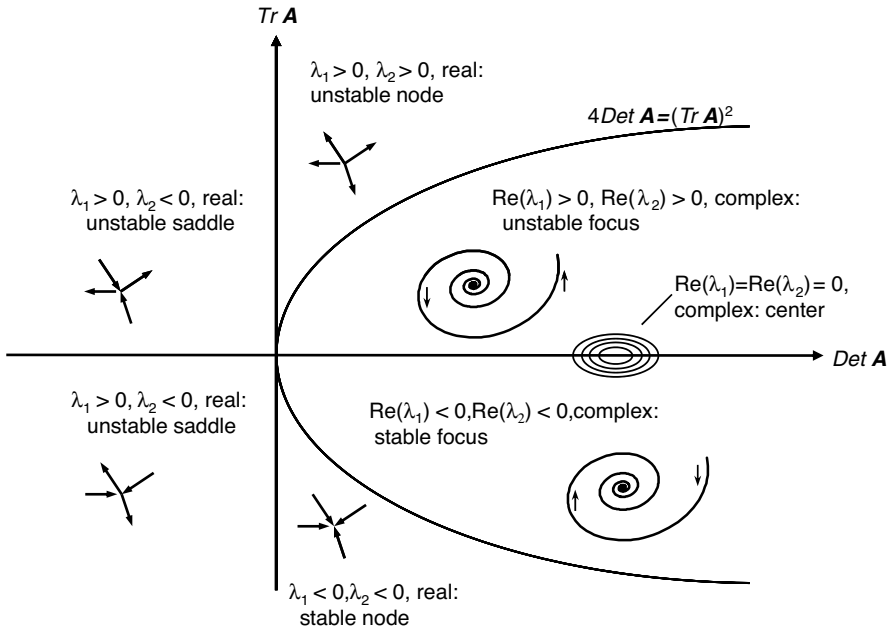
Figure 2.9 Stability of steady states in two-dimensional systems: the character of steady-state solutions is represented depending on the value of the determinant (*x*-axis) and the trace (*y*-axis) of the Jacobian matrix. Phase plane behavior of trajectories in the different cases is schematically represented.

Graphical representation of stability depending on trace and determinant is given in Figure 2.9.

Up to now, we considered only the linearized system. For the stability of the original system, the following holds. If the steady state of the linearized system is asymptotically stable, then the steady state of the complete system is also asymptotically stable. If the steady state of the linearized system is a saddle point, an unstable node or an unstable focus, then the steady state of the complete system is also unstable. This means that statements about the stability remain true, but the character of the steady state is not necessarily kept. For the center, no statement on the stability of the complete system is possible.

Routh–Hurwitz Theorem [31] For systems with $n > 2$ differential equations, we obtain the characteristic polynomial

$$c_n \lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_1 \lambda + c_0 = 0. \tag{2.101}$$

This is a polynomial of degree n, which frequently cannot be solved analytically (at least for $n > 4$). We can use the Hurwitz criterion to test whether the real parts of all eigenvalues are negative. We have to form the Hurwitz matrix H, containing the coefficients of the characteristic polynomial:

$$H = \begin{pmatrix} c_{n-1} & c_{n-3} & c_{n-5} & \dots & 0 \\ c_n & c_{n-2} & c_{n-4} & \dots & 0 \\ 0 & c_{n-1} & c_{n-3} & \dots & 0 \\ 0 & c_n & c_{n-2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_0 \end{pmatrix} = \{h_{ik}\} \text{ with } h_{ik} = \begin{cases} c_{n+i-2k}, & \text{if } 0 \le 2k - i \le n \\ 0, \text{else} \end{cases}.$$

(2.102)

It has been shown that all solutions of the characteristic polynomial have negative real parts, if all coefficients c_i of the polynomial as well as all principal leading minors of H have positive values.

2.3.1.5 Global Stability of Steady States

A state is globally stable, if the trajectories for all initial conditions approach it for $t \to \infty$. The stability of a steady state of an ODE system can be tested with a method proposed by Lyapunov:

Shift the steady state into the point of origin by coordination transformation $\hat{x} = x - \bar{x}$.

Find a function $V_L(x_1, \dots, x_n)$, called Lyapunov function, with the following properties:

(1) $V_L(x_1, \dots, x_n)$ has continuous derivatives with respect to all variables x_i.
(2) $V_L(x_1, \dots, x_n)$ satisfies $V_L(x_1, \dots, x_n) = 0$ for $x_i = 0$ and is positive definite elsewhere, i.e., $V_L(x_1, \dots, x_n) > 0$ for $x_i \ne 0$.
(3) The time derivative of $V_L(x(t))$ is given by

$$\frac{dV_L}{dt} = \sum_{i=1}^{n} \frac{\partial V_L}{\partial x_i} \frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{\partial V_L}{\partial x_i} f_i(x_1, \dots, x_n).$$

(2.103)

A steady state $\bar{x} = 0$ is stable, if the time derivative of $V_L(x(t))$ in a certain region around this state has no positive values. The steady state is asymptotically stable, if the time derivative of $V_L(x(t))$ in this region is negative definite, i.e., $dV_L/dt = 0$ for $x_i = 0$ and $dV_L/dt < 0$ for $x_i \ne 0$.

Example 2.9

The system $\dot{x}_1 = -x_1$, $\dot{x}_2 = -x_2$ has the solution $x_1(t) = x_1^{(0)} e^{-t}$, $x_2(t) = x_2^{(0)} e^{-t}$ and the state $x_1 = x_2 = 0$ is asymptotically stable.

The global stability can also be shown using the positive definite function $V_L = x_1^2 + x_2^2$ as Lyapunov function. It holds $dV_L/dt = (\partial V_L/\partial x_1)\dot{x}_1 + (\partial V_L/\partial x_2)\dot{x}_2 = 2x_1(-x_1) + 2x_2(-x_2)$, which is negative definite.

2.3.1.6 Limit Cycles

Oscillatory behavior is a typical phenomenon in biology. The cause of the oscillation may be different either imposed by external influences or encoded by internal
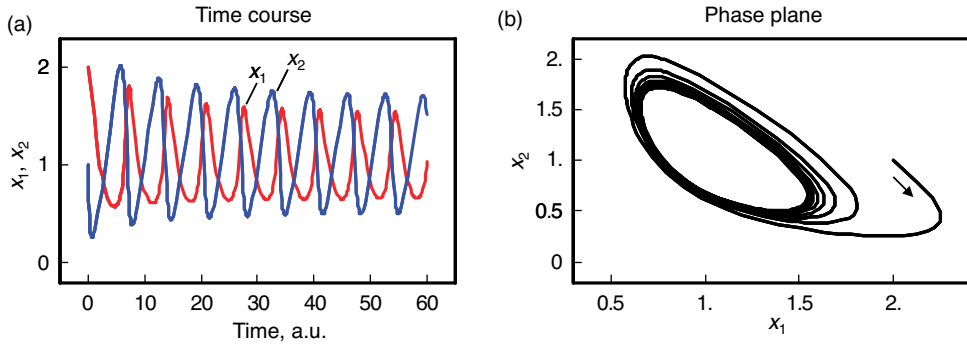
Figure 2.10 Solution of the Equation system in Example 2.10 represented as time course (left panel) and in phase plane (right panel). Initial conditions $x_1(0) = 2$, $x_2(0) = 1$.

structures and parameters. Internally caused stable oscillations can be found if we have a limit cycle in the phase space.

A *limit cycle* is an isolated closed trajectory. All trajectories in its vicinity are periodic solutions winding toward (stable limit cycle) or away from (unstable) the limit cycle for $t \to \infty$.

Example 2.10

The nonlinear system $\dot{x}_1 = x_1^2 x_2 - x_1$, $\dot{x}_2 = p - x_1^2 x_2$ has a steady state at $\bar{x}_1 = p$, $\bar{x}_2 = 1/p$. If we choose, e.g., $p = 0.98$, this steady state is unstable since Tr $\mathbf{A} = 1 - p^2 > 0$ (Figure 2.10).

For 2D systems, there are two criteria to check whether a limit cycle exists. Consider the system of differential equations

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_1, x_2) \end{aligned}. \tag{2.104}$$

The *negative criterion of Bendixson* states: if the expression Tr $\mathbf{A} = \partial f_1/\partial x_1 + \partial f_2/\partial x_2$ does not change its sign in a certain region of the phase plane, then there is no closed trajectory in this area. Hence, a necessary condition for the existence of a limit cycle is the change of the sign of Tr $\mathbf{A}$.

Example 2.11

Example 2.10 holds Tr $\mathbf{A} = (2x_1 x_2 - 1) + (-x_1^2)$. Therefore, Tr $\mathbf{A} = 0$ is fulfilled at $x_2 = (x_1^2 + 1)/(2x_1)$ and Tr $\mathbf{A}$ may assume positive or negative values for varying x_1, x_2, and the necessary condition for the existence of a limit cycle is met.

The criterion of Poincaré–Bendixson states: if a trajectory in the 2D phase plane remains within a finite region without approaching a singular point (a steady state), then this trajectory is either a limit cycle or it approaches a limit cycle. This criterion provides a sufficient condition for the existence of a limit cycle. Nevertheless, the limit cycle trajectory can be computed analytically only in very rare cases.

2.3.2
Metabolic Control Analysis

Metabolic control analysis (MCA) is a powerful quantitative and qualitative framework for studying the relationship between steady-state properties of a network of biochemical reaction and the properties of the individual reactions. It investigates the sensitivity of steady-state properties of the network to small parameter changes. MCA is a useful tool for theoretical and experimental analysis of control and regulation in cellular systems.

MCA was independently founded by two different groups in the 1970s [32, 33] and was further developed by many different groups upon the application to different metabolic systems. A milestone in its formalization was provided by Reder [20]. Originally intended for metabolic networks, MCA has nowadays found applications also for signaling pathways, gene expression models, and hierarchical networks [34–38].

Metabolic networks are very complex systems that are highly regulated and exhibit a lot of interactions such as feedback inhibition or common substrates such as ATP for different reactions. Many mechanisms and regulatory properties of isolated enzymatic reactions are known. The development of MCA was motivated by a series of questions like the following: Can one predict properties or behavior of metabolic networks from the knowledge about their parts, the isolated reactions? Which individual steps control a flux or a steady-state concentration? Is there a rate-limiting step? Which effectors or modifications have the most prominent effect on the reaction rate? In biotechnological production processes, it is of interest which enzyme(s) should be activated in order to increase the rate of synthesis of a desired metabolite. There are also related problems in health care. Concerning metabolic disorders involving overproduction of a metabolite, which reactions should be modified in order to down-regulate this metabolite while perturbing the rest of the metabolism as weakly as possible?

In metabolic networks, the steady-state variables, i.e., the fluxes and the metabolite concentrations, depend on the value of parameters such as enzyme concentrations, kinetic constants (like Michaelis constants and maximal activities), and other model-specific parameters. The relations between steady-state variables and kinetic parameters are usually nonlinear. Up to now, there is no general theory that predicts the effect of large parameter changes in a network. The approach presented here is, basically, restricted to small parameter changes. Mathematically, the system is linearized at steady state, which yields exact results, if the parameter changes are infinitesimally small.

In this section, we will first define a set of mathematical expressions that are useful to quantify control in biochemical reaction networks. Later we will show the relations between these functions and their application for prediction of reaction network behavior.

2.3.2.1 The Coefficients of Control Analysis

Biochemical reaction systems are networks of metabolites connected by chemical reactions. Their behavior is determined by the properties of their components – the individual reactions and their kinetics – as well as by the network structure – the involvement of compounds in different reaction or in brief: the stoichiometry. Hence, the effect of a perturbation exerted on a reaction in this network will depend on both – the local properties of this reaction and the embedding of this reaction in the global network.

Let $y(x)$ denotes a quantity that depends on another quantity x. The effect of the change Δx on y is expressed in terms of sensitivity coefficients:

$$c_x^y = \left(\frac{x}{y} \frac{\Delta y}{\Delta x} \right)_{\Delta x \to 0}. \tag{2.105}$$

In practical applications, Δx might be, e.g., identified with 1% change of x and Δy with the percentage change of y. The factor x/y is a normalization factor that makes the coefficient independent of units and of the magnitude of x and y. In the limiting case $\Delta x \to 0$, the coefficient defined in Eq. (2.105) can be written as

$$c_x^y = \frac{x}{y} \frac{\partial y}{\partial x} = \frac{\partial \ln y}{\partial \ln x}. \tag{2.106}$$

Both right-hand expressions are mathematically equivalent.

Two distinct types of coefficients, local and global coefficients, reflect the relations among local and global effects of changes. *Elasticity coefficients* are local coefficients pertaining to individual reactions. They can be calculated in any given state. *Control coefficients* and *response coefficients* are global quantities. They refer to a given steady state of the entire system. After a perturbation of x, the relaxation of y to new steady state is considered.

The general form of the coefficients in control analysis as defined in Eq. (2.106) contains the normalization x/y. The normalization has the advantage that we get rid of units and can compare, e.g., fluxes belonging to different branches of a network. The drawback of the normalization is that x/y is not defined as soon as $y = 0$, which may happen for certain parameter combinations. In those cases, it is favorable to work with nonnormalized coefficients. Throughout this chapter, we will consider usually normalized quantities. If we use nonnormalized coefficients, they are flagged as $\tilde{c}$. In general, the use of one or the other type of coefficient is also a matter of personal choice of the modeler.

Changes reflected by the different coefficients are illustrated in Figure 2.11.

2.3.2.2 The Elasticity Coefficients

An elasticity coefficient quantifies the sensitivity of a reaction rate to the change of a concentration or a parameter while all other arguments of the kinetic law are kept
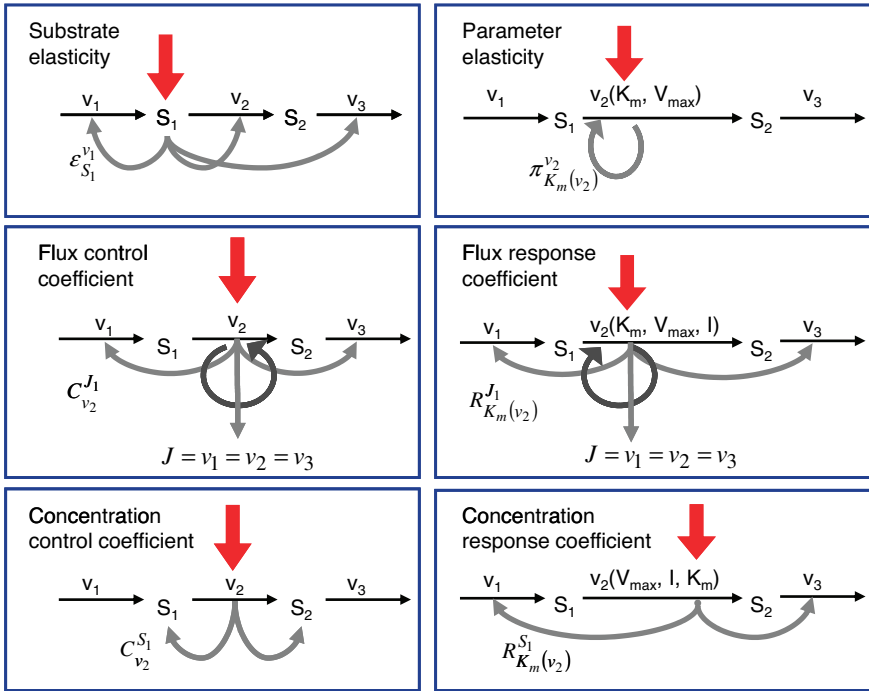
Figure 2.11 Schematic representation of perturbation and effects quantified by different coefficients of metabolic control analysis.

fixed. It measures the direct effect on the reaction velocity, while the rest of the network is not taken into consideration. The sensitivity of the rate v_k of a reaction to the change of the concentration S_i of a metabolite is calculated by the ε-*elasticity*:

$$\varepsilon_i^k = \frac{S_i}{v_k}\frac{\partial v_k}{\partial S_i}. \tag{2.107}$$

The π-*elasticity* is defined with respect to parameters p_m such as kinetic constants, concentrations of enzymes, or concentrations of external metabolites as follows:

$$\pi_m^k = \frac{p_m}{v_k}\frac{\partial v_k}{\partial p_m}. \tag{2.108}$$

Example 2.12

In Michaelis–Menten kinetics, the rate v of a reaction depends on the substrate concentration S in the form $v = V_{max}S/(K_m + S)$ (Eq. (2.22)). The sensitivity is given by the elasticity $\varepsilon_S^v = \partial \ln v / \partial \ln S$. Since the Michaelis–Menten equation defines a mathematical dependency of v on S, it is easy to calculate that

$$\varepsilon_S^\nu = \frac{S}{\nu}\frac{\partial}{\partial S}\left(\frac{V_{max}S}{K_m + S}\right) = \frac{S}{\nu}\frac{V_{max}(K_m + S) - V_{max}S}{(K_m + S)^2} = \frac{S}{K_m + S}. \qquad (2.109)$$

The normalized ε-elasticity in the case of mass action kinetics can be calculated similarly and is always 1. Whenever the rate does not depend directly on a concentration (e.g., for a metabolite of a reaction system that is not involved in the considered reaction), the elasticity is zero.

Example 2.13

Typical values of elasticity coefficients will be explained for an isolated reaction transforming substrate S into product P. The reaction is catalyzed by enzyme E with the inhibitor I, and the activator A as depicted below

$$(2.110)$$

Usually, the elasticity coefficients for metabolite concentrations are in the following range:

$$\varepsilon_S^\nu = \frac{S}{\nu}\frac{\partial \nu}{\partial S} > 0 \quad \text{and} \quad \varepsilon_P^\nu = \frac{P}{\nu}\frac{\partial \nu}{\partial P} \leq 0. \qquad (2.111)$$

In most cases, the rate increases with the concentration of the substrate (compare, e.g., Eq. (2.109)) and decreases with the concentration of the product. An exception from $\varepsilon_S^\nu > 0$ occurs in the case of substrate inhibition (Eq. (2.33)), where the elasticity will become negative for $S > S_{opt}$. The relation $\varepsilon_P^\nu = 0$ holds, if the reaction is irreversible or if the product concentration is kept zero by external mechanisms. The elasticity coefficients with respect to effectors I or A should obey

$$\varepsilon_A^\nu = \frac{A}{\nu}\frac{\partial \nu}{\partial A} > 0 \quad \text{and} \quad \varepsilon_I^\nu = \frac{I}{\nu}\frac{\partial \nu}{\partial I} < 0, \qquad (2.112)$$

since this is essentially what the notions activator and inhibitor mean.

For the most kinetic laws, the reaction rate ν is proportional to the enzyme concentration E. For example, E is a multiplicative factor in the mass action rate law as well as in the maximal rate of the Michaelis–Menten rate law. Therefore, it holds that

$$\varepsilon_E^\nu = \frac{\partial \ln \nu}{\partial \ln E} = 1. \qquad (2.113)$$

More complicated interactions between enzymes and substrates such as metabolic channeling (direct transfer of the metabolite from one enzyme to the next without release to the medium) may lead to exceptions from this rule.

2.3.2.3 Control Coefficients

When defining control coefficients, we refer to a stable steady state of the metabolic system characterized by steady-state concentrations $S^{st} = S^{st}(p)$ and steady-state fluxes $J = v(S^{st}(p), p)$. Any sufficiently small perturbation of an individual reaction rate, $v_k \rightarrow v_k + \Delta v_k$, by a parameter change $p_k \rightarrow p_k + \Delta p_k$ drives the system to a new steady state in close proximity with $J \rightarrow J + \Delta J$ and $S^{st} \rightarrow S^{st} + \Delta S$. A measure for the change of fluxes and concentrations are the control coefficients.

The *flux control coefficient* for the control of rate v_k over flux J_j is defined as

$$C_k^j = \frac{v_k}{J_j} \frac{\partial J_j / \partial p_k}{\partial v_k / \partial p_k}. \tag{2.114}$$

The control coefficients quantify the control that a certain reaction v_k exerts on the steady-state flux J_j. It should be noted that the rate change, Δv_k, is caused by the change of a parameter p_k that has a direct effect solely on v_k. Thus, it holds

$$\frac{\partial v_k}{\partial p_k} \neq 0 \quad \text{and} \quad \frac{\partial v_l}{\partial p_k} = 0 \quad (l \neq k). \tag{2.115}$$

Such a parameter might be the enzyme concentration, a kinetic constant, or the concentration of a specific inhibitor or effector.

In a more compact form the flux control coefficients read

$$C_k^j = \frac{v_k}{J_j} \frac{\partial J_j}{\partial v_k}. \tag{2.116}$$

Equivalently, the *concentration control coefficient* of concentrations S_i^{st} with respect to v_k reads

$$C_k^i = \frac{v_k}{S_i^{st}} \frac{\partial S_i^{st}}{\partial v_k}. \tag{2.117}$$

2.3.2.4 Response Coefficients

The steady state is determined by the values of the parameters. A third type of coefficients expresses the direct dependence of steady-state variables on parameters. The response coefficients are defined as

$$R_m^j = \frac{p_m}{J_j} \frac{\partial J_j}{\partial p_m} \quad \text{and} \quad R_m^i = \frac{p_m}{S_i^{st}} \frac{\partial S_i^{st}}{\partial p_m}, \tag{2.118}$$

where the first coefficient expresses the response of the flux to a parameter perturbation, while the latter describes the response of a steady-state concentration.

2.3.2.5 Matrix Representation of the Coefficients

Control, response, and elasticity coefficients are defined with respect to all rates, steady-state concentrations, fluxes, or parameters in the metabolic system and in the respective model. They can be arranged in matrices:

$$\boldsymbol{C}^J = \{C_k^j\}, \ \boldsymbol{C}^S = \{C_k^i\}, \ \boldsymbol{R}^J = \{R_m^j\}, \ \boldsymbol{R}^S = \{R_m^i\}, \ \boldsymbol{\varepsilon} = \{\varepsilon_i^k\}, \ \boldsymbol{\pi} = \{\pi_m^k\}. \tag{2.119}$$

Matrix representation can also be chosen for all types of nonnormalized coefficients. The arrangement in matrices allows us to apply matrix algebra in control analysis. In particular, the matrices of normalized control coefficients can be calculated from the matrices of nonnormalized control coefficient as follows:

$$
\begin{aligned}
\boldsymbol{C}^J &= (\mathrm{dg}\boldsymbol{J})^{-1} \cdot \tilde{\boldsymbol{C}}^J \cdot \mathrm{dg}\boldsymbol{J} & \boldsymbol{C}^S &= (\mathrm{dg}\boldsymbol{S}^{\mathrm{st}})^{-1} \cdot \tilde{\boldsymbol{C}}^J \cdot \mathrm{dg}\boldsymbol{J} \\
\boldsymbol{R}^J &= (\mathrm{dg}\boldsymbol{J})^{-1} \cdot \tilde{\boldsymbol{R}}^J \cdot \mathrm{dg}\boldsymbol{p} & \boldsymbol{R}^S &= (\mathrm{dg}\boldsymbol{S}^{\mathrm{st}})^{-1} \cdot \tilde{\boldsymbol{R}}^S \cdot \mathrm{dg}\boldsymbol{p} \cdot \\
\boldsymbol{\varepsilon} &= (\mathrm{dg}\boldsymbol{v})^{-1} \cdot \tilde{\boldsymbol{\varepsilon}} \cdot \mathrm{dg}\boldsymbol{S}^{\mathrm{st}} & \boldsymbol{\pi} &= (\mathrm{dg}\boldsymbol{v})^{-1} \cdot \tilde{\boldsymbol{\pi}} \cdot \mathrm{dg}\boldsymbol{p}
\end{aligned}
\tag{2.120}
$$

The symbol "dg" stands for the diagonal matrix, e.g., for a system with three reaction holds

$$
\mathrm{dg}\boldsymbol{J} = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix}.
$$

2.3.2.6 The Theorems of Metabolic Control Theory

Let us assume that we are interested in calculating the control coefficients for a system under investigation. Usually, the steady-state fluxes or concentrations cannot be expressed explicitly as function of the reaction rates. Therefore, flux and concentration control coefficients cannot simply be determined by taking the respective derivatives, as we did for the elasticity coefficients in Example 2.12.

Fortunately, the work with control coefficients is eased by of a set of theorems. The first type of theorems, the *summation theorems*, makes a statement about the total control over a flux or a steady-state concentration. The second type of theorems, the *connectivity theorems*, relates the control coefficients to the elasticity coefficients. Both types of theorems together with network information encoded in the stoichiometric matrix contain enough information to calculate all control coefficients.

Here, we will first introduce the theorems. Then, we will present a hypothetical perturbation experiment (as introduced by Kacser and Burns) to illustrate the summation theorem. Finally, the theorems will be derived mathematically.

2.3.2.7 The Summation Theorems

The summation theorems make a statement about the total control over a certain steady-state flux or concentration. The flux control coefficients and concentration control coefficients fulfill, respectively,

$$
\sum_{k=1}^{r} C_{v_k}^{J_j} = 1 \quad \text{and} \quad \sum_{k=1}^{r} C_{v_k}^{S_i} = 0,
\tag{2.121}
$$

for any flux J_j and any steady-state concentration S_i^{st}. The quantity r is the number of reactions. The flux control coefficients of a metabolic network for one steady-state flux sum up to one. This means that all enzymatic reactions can share the control over this flux. The control coefficients of a metabolic network for one steady-state concentration are balanced. This means again that the enzymatic reactions can share the control over this concentration, but some of them exert a negative control while

others exert a positive control. Both relations can also be expressed in matrix formulation. We get

$$\mathbf{C}^J \cdot \mathbf{1} = \mathbf{1} \quad \text{and} \quad \mathbf{C}^S \cdot \mathbf{1} = \mathbf{0}. \tag{2.122}$$

The symbols $\mathbf{1}$ and $\mathbf{0}$ denote column vectors with r rows containing as entries only ones or zeros, respectively. The summation theorems for the nonnormalized control coefficients read

$$\tilde{\mathbf{C}}^J \cdot \mathbf{K} = \mathbf{K} \quad \text{and} \quad \tilde{\mathbf{C}}^S \cdot \mathbf{K} = \mathbf{0}, \tag{2.123}$$

where $\mathbf{K}$ is the matrix satisfying $\mathbf{N} \cdot \mathbf{K} = \mathbf{0}$ (see Section 2.2). A more intuitive derivation of the summation theorems is given in the following example according to Kacser and Burns [33].

Example 2.14

The summation theorem for flux control coefficients can be derived using a thought experiment.

Consider the following unbranched pathway with fixed concentrations of the external metabolites, S_0 and S_3:

$$S_0 \overset{v_1}{\leftrightarrow} S_1 \overset{v_2}{\leftrightarrow} S_2 \overset{v_3}{\leftrightarrow} S_3 \tag{2.124}$$

What happens to steady-state fluxes and metabolite concentrations, if we perform an experimental manipulation of all three reactions leading to the same fractional change α of all three rates?

$$\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} = \frac{\delta v_3}{v_3} = \alpha. \tag{2.125}$$

The flux must increase to the same extent, $\delta J / J = \alpha$, but, since rates of producing and degrading reactions increase to the same amount, the concentrations of the metabolites remain constant $\delta S_1 / S_1 = \delta S_2 / S_2 = 0$.

The combined effect of all changes in local rates on the system variables S_1^{st}, S_2^{st}, and J can be written as the sum of all individual effects caused by the local rate changes. For the flux holds

$$\frac{\delta J}{J} = C_1^J \frac{\delta v_1}{v_1} + C_2^J \frac{\delta v_2}{v_2} + C_3^J \frac{\delta v_3}{v_3}. \tag{2.126}$$

It follows

$$\alpha = \alpha(C_1^J + C_2^J + C_3^J) \quad \text{or} \quad 1 = C_1^J + C_2^J + C_3^J. \tag{2.127}$$

This is just a special case of Eq. (2.121). In the same way, for the change of concentration S_1^{st}, we obtain

$$\frac{\delta S_1^{st}}{S_1^{st}} = C_1^{S_1} \frac{\delta v_1}{v_1} + C_2^{S_1} \frac{\delta v_2}{v_2} + C_3^{S_1} \frac{\delta v_3}{v_3}. \tag{2.128}$$

Finally, we get

$$0 = C_1^{S_1} + C_2^{S_1} + C_3^{S_1} \quad \text{as well as} \quad 0 = C_1^{S_2} + C_2^{S_2} + C_3^{S_2}. \tag{2.129}$$

Although shown here only for a special case, these properties hold in general for systems without conservation relations. The general derivation is given in Section 2.3.2.9.

2.3.2.8 The Connectivity Theorems

Flux control coefficients and elasticity coefficients are related by the expression

$$\sum_{k=1}^{r} C_{v_k}^{J_j} \varepsilon_{S_i}^{v_k} = 0. \tag{2.130}$$

Note that the sum runs over all rates v_k for any flux J_j. Considering the concentration S_i of a specific metabolite and a certain flux J_j, each term contains the elasticity $\varepsilon_{S_i}^{v_k}$ describing the direct influence of a change of S_i on the rates v_k and the control coefficient expressing the control of v_k over J_j.

The connectivity theorem between concentration control coefficients and elasticity coefficients reads

$$\sum_{k=1}^{r} C_{v_k}^{S_h} \varepsilon_{S_i}^{v_k} = -\delta_{hi}. \tag{2.131}$$

Again, the sum runs over all rates v_k, while S_h and S_i are the concentrations of two fixed metabolites. The symbol $\delta_{hi} = \begin{cases} 0, & \text{if} \quad h \neq i \\ 1, & \text{if} \quad h = i \end{cases}$ is the so-called Kronecker symbol.

In matrix formulation, the connectivity theorems read

$$\mathbf{C}^J \cdot \boldsymbol{\varepsilon} = \mathbf{0} \quad \text{and} \quad \mathbf{C}^S \cdot \boldsymbol{\varepsilon} = -\mathbf{I}, \tag{2.132}$$

where $\mathbf{I}$ denotes the identity matrix of size $n \times n$. For nonnormalized coefficients, it holds

$$\tilde{\mathbf{C}}^J \cdot \tilde{\boldsymbol{\varepsilon}} \cdot \mathbf{L} = 0 \quad \text{and} \quad \tilde{\mathbf{C}}^S \cdot \tilde{\boldsymbol{\varepsilon}} \cdot \mathbf{L} = -\mathbf{L}, \tag{2.133}$$

where $\mathbf{L}$ is the link matrix that expresses the relation between independent and dependent rows in the stoichiometric matrix (Eq. (2.75)) A comprehensive representation of both summation and connectivity theorems for nonnormalized coefficients is given by the following equation:

$$\begin{pmatrix} \tilde{\mathbf{C}}^J \\ \tilde{\mathbf{C}}^S \end{pmatrix} \cdot (\mathbf{K} \quad \tilde{\boldsymbol{\varepsilon}} \mathbf{L}) = \begin{pmatrix} \mathbf{K} & 0 \\ 0 & -\mathbf{L} \end{pmatrix}. \tag{2.134}$$

The summation and connectivity theorem together with the structural information of the stoichiometric matrix are sufficient to calculate the control coefficients

for a metabolic network. This shall be illustrated for a small network in the next example.

Example 2.15

To calculate the control coefficients, we study the following reaction system:

$$P_0 \overset{v_1}{\leftrightarrow} S \overset{v_2}{\leftrightarrow} P_2 \tag{2.135}$$

The flux control coefficients obey the theorems

$$C_1^J + C_2^J = 1 \quad \text{and} \quad C_1^J \varepsilon_S^1 + C_2^J \varepsilon_S^2 = 0, \tag{2.136}$$

which can be solved for the control coefficients to yield

$$C_1^J = \frac{\varepsilon_S^2}{\varepsilon_S^2 - \varepsilon_S^1} \quad \text{and} \quad C_2^J = \frac{-\varepsilon_S^1}{\varepsilon_S^2 - \varepsilon_S^1}. \tag{2.137}$$

Since usually $\varepsilon_S^1 < 0$ and $\varepsilon_S^2 > 0$ (see Example 2.13), both control coefficients assume positive values $C_1^J > 0$ and $C_2^J > 0$. This means that both reactions exert a positive control over the steady-state flux, and acceleration of any of them leads to increase of J, which is in accordance with common intuition.

The concentration control coefficients fulfil

$$C_1^S + C_2^S = 0 \quad \text{and} \quad C_1^S \varepsilon_S^1 + C_2^S \varepsilon_S^2 = -1, \tag{2.138}$$

which yields

$$C_1^S = \frac{1}{\varepsilon_S^2 - \varepsilon_S^1} \quad \text{and} \quad C_2^S = \frac{-1}{\varepsilon_S^2 - \varepsilon_S^1}. \tag{2.139}$$

With $\varepsilon_S^1 < 0$ and $\varepsilon_S^2 > 0$, we get $C_1^S > 0$ and $C_2^S < 0$, i.e., increase of the first reaction causes a raise in the steady-state concentration of S while acceleration of the second reaction leads to the opposite effect.

2.3.2.9 Derivation of Matrix Expressions for Control Coefficients

After having introduced the theorems of MCA, we will derive expressions for the control coefficients in matrix form. These expressions are suited for calculating the coefficients even for large-scale models. We start from the steady-state condition

$$N v(S^{\text{st}}(p), p) = 0. \tag{2.140}$$

Implicit differentiation with respect to the parameter vector p yields

$$N \frac{\partial v}{\partial S} \frac{\partial S^{\text{st}}}{\partial p} + N \frac{\partial v}{\partial p} = 0. \tag{2.141}$$

Since we have chosen reaction-specific parameters for perturbation, the matrix of nonnormalized parameter elasticities contains nonzero entries in the main diagonal and zeros elsewhere (compare Eq. (2.115)).

$$\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} = \begin{pmatrix} \dfrac{\partial v_1}{\partial p_1} & 0 & & 0 \\[2mm] 0 & \dfrac{\partial v_2}{\partial p_2} & & 0 \\ & & \cdots & \\ 0 & 0 & & \dfrac{\partial v_r}{\partial p_r} \end{pmatrix}. \tag{2.142}$$

Therefore, this matrix is regular and has an inverse. Furthermore, we consider the Jacobian matrix

$$\boldsymbol{M} = \boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} = \boldsymbol{N}\tilde{\boldsymbol{\varepsilon}}. \tag{2.143}$$

The Jacobian $\boldsymbol{M}$ is a regular matrix if the system is asymptotically stable and contains no conservation relations. The case with conservation relations is considered below. Here, we may premultiply Eq. (2.141) by the inverse of $\boldsymbol{M}$ and rearrange to get

$$\frac{\partial \boldsymbol{S}^{\text{st}}}{\partial \boldsymbol{p}} = -\left(\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\right)^{-1}\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} = -\boldsymbol{M}^{-1}\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} \equiv \tilde{\boldsymbol{R}}^{S}. \tag{2.144}$$

As indicated, $\partial \boldsymbol{S}^{\text{st}}/\partial \boldsymbol{p}$ is the matrix of nonnormalized response coefficients for concentrations. Postmultiplication by the inverse of the nonnormalized parameter elasticity matrix gives

$$\frac{\partial \boldsymbol{S}^{\text{st}}}{\partial \boldsymbol{p}}\left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}}\right)^{-1} = -\left(\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\right)^{-1}\boldsymbol{N} = \tilde{\boldsymbol{C}}^{S}. \tag{2.145}$$

This is the matrix of nonnormalized concentration control coefficients. The right (middle) site contains no parameters. This means, that the control coefficients do not depend on the particular choice of parameters to exert the perturbation as long as Eq. (2.115) is fulfilled. The control coefficients are only dependent on the structure of the network represented by the stoichiometric matrix $\boldsymbol{N}$, and on the kinetics of the individual reactions, represented by the nonnormalized elasticity matrix $\tilde{\boldsymbol{\varepsilon}} = \partial \boldsymbol{v}/\partial \boldsymbol{S}$.

The implicit differentiation of

$$\boldsymbol{J} = \boldsymbol{v}(\boldsymbol{S}^{\text{st}}(\boldsymbol{p}), \boldsymbol{p}), \tag{2.146}$$

with respect to the parameter vector $\boldsymbol{p}$ leads to

$$\frac{\partial \boldsymbol{J}}{\partial \boldsymbol{p}} = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} + \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\frac{\partial \boldsymbol{S}^{\text{st}}}{\partial \boldsymbol{p}} = \left(\boldsymbol{I} - \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\left(\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\right)^{-1}\boldsymbol{N}\right)\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}} \equiv \tilde{\boldsymbol{R}}^{J}. \tag{2.147}$$

This yields, after some rearrangement, an expression for the nonnormalized flux control coefficients:

$$\frac{\partial \boldsymbol{J}}{\partial \boldsymbol{p}}\left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{p}}\right)^{-1} = \boldsymbol{I} - \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\left(\boldsymbol{N}\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}}\right)^{-1}\boldsymbol{N} = \tilde{\boldsymbol{C}}^{J}. \tag{2.148}$$

The normalized control coefficients are (by use of Eq. (2.120))

$$\boldsymbol{C}^J = \boldsymbol{I} - (\mathrm{dg}\boldsymbol{J})^{-1} \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \left(\boldsymbol{N} \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \right)^{-1} \boldsymbol{N} \right) (\mathrm{dg}\boldsymbol{J}) \quad \text{and}$$

$$\boldsymbol{C}^S = -(\mathrm{dg}\boldsymbol{S}^{\mathrm{st}})^{-1} \left(\left(\boldsymbol{N} \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \right)^{-1} \boldsymbol{N} \right) (\mathrm{dg}\boldsymbol{J}).$$

(2.149)

These equations can easily be implemented for numerical calculation of control coefficients or used for analytical computation. They are also suited for derivation of the theorems of MCA. The summation theorems for the control coefficients follow from Eq. (2.149) by postmultiplication with the vector $\boldsymbol{1}$ (the row vector containing only 1s), and consideration of the relations $(\mathrm{dg}\boldsymbol{J}) \cdot \boldsymbol{1} = \boldsymbol{J}$ and $\boldsymbol{N}\boldsymbol{J} = \boldsymbol{0}$. The connectivity theorems result from postmultiplication of Eq. (2.149) with the elasticity matrix $\boldsymbol{\varepsilon} = (\mathrm{dg}\boldsymbol{J})^{-1} \cdot (\partial \boldsymbol{v}/\partial \boldsymbol{S}) \cdot \mathrm{dg}\boldsymbol{S}^{\mathrm{st}}$, and using that multiplication of a matrix with its inverse yields the identity matrix $\boldsymbol{I}$ of respective type.

If the reaction system involves conservation relations, we eliminate dependent variables as explained in Section 1.2.4. In this case, the nonnormalized coefficients read

$$\tilde{\boldsymbol{C}}^J = \boldsymbol{I} - \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \boldsymbol{L} \left(\boldsymbol{N}_R \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \right)^{-1} \boldsymbol{N}_R \quad \text{and} \quad \tilde{\boldsymbol{C}}^S = -\boldsymbol{L} \left(\boldsymbol{N}_R \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{S}} \right)^{-1} \boldsymbol{N}_R \qquad (2.150)$$

and the normalized control coefficients are obtained by applying Eq. (2.120).

An example for calculation of flux control coefficients can be found in the web material.

To investigate the implications of control distribution, we will now analyze the control pattern in an unbranched pathway:

$$S_0 \xrightarrow{v_1} S_1 \xrightarrow{v_2} S_2 \cdots S_{r-1} \overset{v_r}{\leftrightarrow} S_r \qquad (2.151)$$

with linear kinetics $v_i = k_i S_{i-1} - k_{-i} S_i$, the equilibrium constants $q_i = k_i/k_{-i}$ and fixed concentrations of the external metabolites, S_0 and S_r. In this case, one can calculate an analytical expression for the steady-state flux,

$$J = \frac{S_0 \prod\limits_{j=1}^{r} q_j - S_r}{\sum\limits_{l=1}^{r} \frac{1}{k_l} \prod\limits_{m=l}^{r} q_m} \qquad (2.152)$$

as well as an analytical expression for the flux control coefficients

$$C_i^J = \left(\frac{1}{k_i} \prod_{j=i}^{r} q_j \right) \cdot \left(\sum_{l=1}^{r} \frac{1}{k_l} \prod_{m=l}^{r} q_m \right)^{-1}. \qquad (2.153)$$

Let us consider two very general cases. First assume that all reactions have the same individual kinetics, $k_i = k_+$, $k_{-i} = k_-$ for $i = 1, \ldots, r$ and that the equilibrium constants, which are also equal, satisfy $q = k_+/k_- > 1$. In this case, the ratio of two subsequent flux control coefficients is

$$\frac{C_i^J}{C_{i+1}^J} = \frac{k_{i+1}}{k_i} q_i = q > 1. \tag{2.154}$$

Hence, the control coefficients of the preceding reactions are larger than the control coefficients of the succeeding reactions and flux control coefficients are higher in the beginning of a chain than in the end. This is in agreement with the frequent observation that flux control is strongest in the upper part of an unbranched reaction pathway.

Now assume that the individual rate constants might be different, but that all equilibrium constants are equal to one, $q_i = 1$ for $i = 1, \ldots, r$. This implies $k_i = k_{-i}$. Equation (2.153) simplifies to

$$C_i^J = \frac{1}{k_i} \cdot \left(\sum_{l=1}^{r} \frac{1}{k_l} \right)^{-1}. \tag{2.155}$$

Consider now the relaxation time $\tau_i = 1/(k_i + k_{-i})$ (see Section 4.3) as a measure for the rate of an enzyme. The flux control coefficient reads

$$C_i^J = \frac{\tau_i}{\tau_1 + \tau_2 + \cdots + \tau_r}. \tag{2.156}$$

This expression helps to elucidate two aspects of metabolic control. First, all enzymes participate in the control since all enzymes have a positive relaxation time. There is no enzyme that has all control, i.e., determines the flux through the pathway alone. Second, slow enzymes with a higher relaxation time exert in general more control than fast enzymes with a short relaxation time.

The predictive power of flux control coefficients for directed changes of flux is illustrated in the following example.

Example 2.16

Assume that we can manipulate the pathway shown in Figure 2.12 by changing the enzyme concentration in a predefined way. We would like to explore the effect of the perturbation of the individual enzymes. For a linear pathway (see Eqs. (2.151)–(2.153)) consisting of four consecutive reactions, we calculate the flux control coefficients. For $i = 1, \ldots, 4$, it shall hold that (i) all enzyme concentrations $E_i = 1$, (ii) the rate constants be $k_i = 2$, $k_{-i} = 1$, and (iii) the concentrations of the external reactants be $S_0 = S_4 = 1$. The resulting flux is $J = 1$ and the flux control coefficients are $\mathbf{C}^J = (\, 0.533 \quad 0.267 \quad 0.133 \quad 0.067 \,)^T$ according to Eq. (2.149).

If we now perturb slightly the first enzyme, lets say perform a percentage change of its concentration, i.e., $E_1 \rightarrow E_1 + 1\%$, then Eq. (2.105) implies that the flux increases as $J \rightarrow J + C_1^J \cdot 1\%$. In fact, the flux in the new steady state is $J^{E_1 \rightarrow 1.01 \cdot E_1} = 1.00531$. Increasing E_2, E_3, or E_4 by 1% leads to flux values of 1.00265, 1.00132, and 1.00066, respectively. A strong perturbation would not yield similar effects. This is illustrated in Figure 2.12.
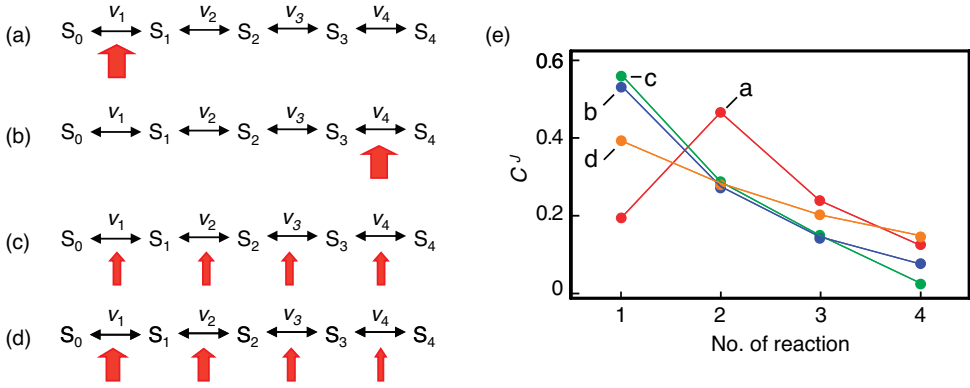
Figure 2.12 Effect of enzyme concentration change on steady-state flux and on flux control coefficients in an unbranched pathway consisting of four reactions. In the reference state, all enzymes have the concentration 1 (in arbitrary units), the control distribution is the same as in case (c), and the steady-state flux is $J = 1$. (a) Change of $E_1 \rightarrow 5E_1$ while keeping the other enzyme concentrations constant results in a remarkable drop of control of the first enzyme.

The resulting flux is $J^{E_1 \rightarrow 5 \cdot E_1} = 1.7741$. (b) The change $E_4 \rightarrow 5E_4$ corresponds to $J^{E_4 \rightarrow 5 \cdot E_4} = 1.0563$. There is only slight change of control distribution. (c) Equal enzyme concentrations with $E_i \rightarrow 2E_i, i = 1, \ldots, 4$ results in $J^{E_i \rightarrow 2 \cdot E_i} = 2$. (d) Optimal distribution of enzyme concentration $E_1 = 3.124$, $E_2 = 2.209$, $E_3 = 1.562$, $E_4 = 1.105$ resulting in the maximal steady-state flux $J^{max} = 2.2871$.

2.4
Tools and Data Formats for Modeling

Summary

This section gives an overview about different simulation techniques and introduces tools, resources, and standard formats used in systems biology. Modeling and simulation functionalities of the tools are presented and common data formats used by these tools and in general in systems biology are introduced. Furthermore, model databases and databases of cellular and biochemical reaction networks are discussed.

The development of models of biological and in particular cellular systems starts by the collection of the model components and its interactions. Usually, in the first step, one formulates the biochemical reaction equations that define the topological structure of the reaction network and the reaction stoichiometries. For this purpose, it is often also useful to draw a diagram that illustrates the network structure either of the whole model or of a particular part. Once the reaction network and its stoichiometry are defined, a more detailed mathematical model can be constructed. For this purpose, often systems of ODEs are applied. Usually, this requires very detailed

information about the kinetics of the individual reactions or appropriate assumptions have to be made.

In this section, databases are presented that provide information on the network structure of cellular processes such as metabolic pathways and signal transduction pathways. Moreover, data formats used for the structural, mathematical, and graphical description of biochemical reaction networks are introduced. We will start this section with an overview of simulation techniques and of software tools that support the user by the development of models.

2.4.1
Simulation Techniques

In systems biology, different simulation techniques are used such as systems of ODEs, stochastic methods, Petri nets, π-calculus, PDEs, cellular automata (CA) methods, agent-based systems, and hybrid approaches. The use of ODEs in biological modeling is widespread and by far the most common simulation approach in computational systems biology [39, 40]. The description of a biological model by a system of ODEs has already been discussed in the earlier sections. Some ODEs are simple enough to be solved analytically and have an exact solution. More complex ODE systems, as they are occurring in most systems biology simulations, must be solved numerically by appropriate algorithms. A first method for the numerical solution of ODEs was derived by Newton and Gauss. Methods that provide more improved computational accuracy are, for instance, Runge–Kutta algorithms and implicit methods that can also handle so-called stiff differential equations. Simulation tools for systems biology have to cope with systems of multiple reactants and multiple reactions. For the numerical integration of such complex ODE systems, they usually make use of more advanced programs such as LSODA [41, 42], CVODE [43], or LIMEX [44]. In the following, Petri nets and CA are described in more detail.

2.4.1.1 Petri Nets
An alternative to ODEs for the simulation of time-dependent processes are Petri nets. A Petri net is a graphical and mathematical modeling tool for discrete and parallel systems. The mathematical concept was developed in the early 1960s by Carl Adam Petri. The basic elements of a Petri net are places, transitions and arcs that connect places and transitions. When represented graphically, places are shown as circles and transitions as rectangles. Places represent objects (e.g., molecules, cars, and machine parts) and transitions describe if and how individual objects are inter-converted. Places can contain zero or more tokens, indicating the number of objects that currently exist. If a transition can take place (can fire) or not depends on the places that are connected to the transition by incoming arcs, to contain enough tokens. If this condition is fulfilled, the transition fires and changes the state of the system by removing tokens from the input places and adding tokens to the output places. The number of tokens that are removed and added depends on the weights of the arcs.

Petri nets are not only an optically pleasing representation of a system but can also be described mathematically in terms of integer arithmetic. For simple types of Petri nets, certain properties can thus be calculated analytically, but often the net has to be run to study the long-term system properties. Over the years many, extensions to the basic Petri net model have been developed for the different simulation purposes [45].

1. Hybrid Petri nets that add the possibility to have places that contain a continuous token number instead of discrete values.
2. Timed Petri nets extend transitions to allow for a specific time delay between the moment when a transition is enabled and the actual firing.
3. Stochastic Petri nets that go one step further and allow a random time delay drawn from a probability distribution.
4. Hierarchical Petri nets, in which modularity is introduced by representing whole nets as a single place or transition of a larger net.
5. Colored Petri nets that introduce different types (colors) or tokens and more complicated firing rules for transitions.

With these extensions, Petri nets are powerful enough to be used for models in systems biology. Biochemical pathways can be modeled with places representing metabolites, transitions representing reactions and stoichiometric coefficients are encoded as different weights of input and output arcs. Consequently, Petri nets have been used to model metabolic networks [46, 47] and signal transduction pathways [48]. Many free and commercial tools are available to explore the behavior of Petri nets. The *Petri Nets World* webpage (http://www.informatik. uni-hamburg.de/TGI/PetriNets/) is an excellent starting point for this purpose.

2.4.1.2 Cellular Automata
Cellular Automata (CA) are tools for the simulation of temporal or spatiotemporal processes using discrete time and/or spatial steps (see Section 3.4.1.3). A cellular automaton consists of a regular grid or lattice of nearly identical components, called cells, where each cell has a certain state of a finite number of states. The states of the cells evolve synchronously in discrete time steps according to a set of rules. Each particular state of cell is determined by the previous states of its neighbors. CA were invented in the late 1940s by von Neumann and Ulam. A well-known CA simulation is Conway's Game of Life [49].

2.4.2
Simulation Tools

In the following, three different simulation tools are presented that essentially make use of ODE systems for simulation, and come along with further functionalities important for modeling, such as graphical visualization of the reaction network, advanced analysis techniques, and interfaces to external model and pathway databases. Further modeling and simulation tools are presented in Chapter 17. Modeling

and simulations tools have also been reviewed by Alves *et al.* [50], Klipp *et al.* [51], Materi and Wishart [52], and Wierling *et al.* [53].

Modeling systems have to accomplish several requirements. They must have a well-defined internal structure for the representation of model components and reactions, and optionally functionalities for the storage of a model in a well-defined structure, standardized format, or database. Further desired aspects are a user-friendly interface for model development, a graphical representation of reaction networks, a detailed description of the mathematical model, integrated simulation engines, e.g., for deterministic or stochastic simulation, along with graphical representations of those simulation results, and functionalities for model analysis and model refinement. This is a very broad spectrum of functionalities. Existing tools cover different aspects of these functionalities. In the following, systems biology tools will be introduced that already accomplish several of the desired aspects. CellDesigner is one of those widely used in the systems biology community [51]. It has a user-friendly process diagram editor, uses the Systems Biology Markup Language (SBML; see Section 2.4.3.1) for model representation and exchange, and provides fundamental simulation and modeling functions. Another program with similar functionalities is COPASI. COPASI has an interface for the model definition and representation and provides several methods for simulation, model analysis, and refinement such as parameter scanning, MCA, optimization, or parameter estimation. Similarly, also PyBioS has rich functionalities for model design, simulation, and analysis. In contrast to the stand-alone programs CellDesigner and Copasi, PyBioS is a web application. A particular feature of PyBioS is its interfaces to pathway databases, like Reactome or KEGG, which can directly be used for model generation.

2.4.2.1 CellDesigner

CellDesigner provides an advanced graphical model representation along with an easy to use user-interface and an integrated simulation engine [54]. The current version of CellDesigner is 4.0.1. The process diagram editor of CellDesigner supports a rich set of graphical elements for the description of biochemical and gene-regulatory networks. Networks can be constructed from compartments, species, and reactions. CellDesigner comes with a large number of predefined shapes that can be used for different types of molecules, such as proteins, receptors, ion channels, small metabolites, etc. It is also possible to modify the symbols to indicate phosphorylations or other modifications. The program also provides several icons for special reaction types like catalysis, transport, inhibition, and activation. For version 4.0, it is announced that the graphical elements are compliant with the Systems Biology Graphical Notation (SBGN; see Section 2.4.3.3).

Reading and writing of the models is SBML-based (see Section 2.4.3.1 for more details on SBML) and the models written by CellDesigner pass the online validation at http://sbml.org/tools/htdocs/sbmltools.php and thus are conform with the SBML standard. A nice feature in this respect is the ability to display the SBML model structure as a tree (Figure 2.13, left side). A click on a species or reaction in this tree highlights the corresponding elements in the graphics canvas and in the matching tab on the right side showing further details. This tab is also the place where initial
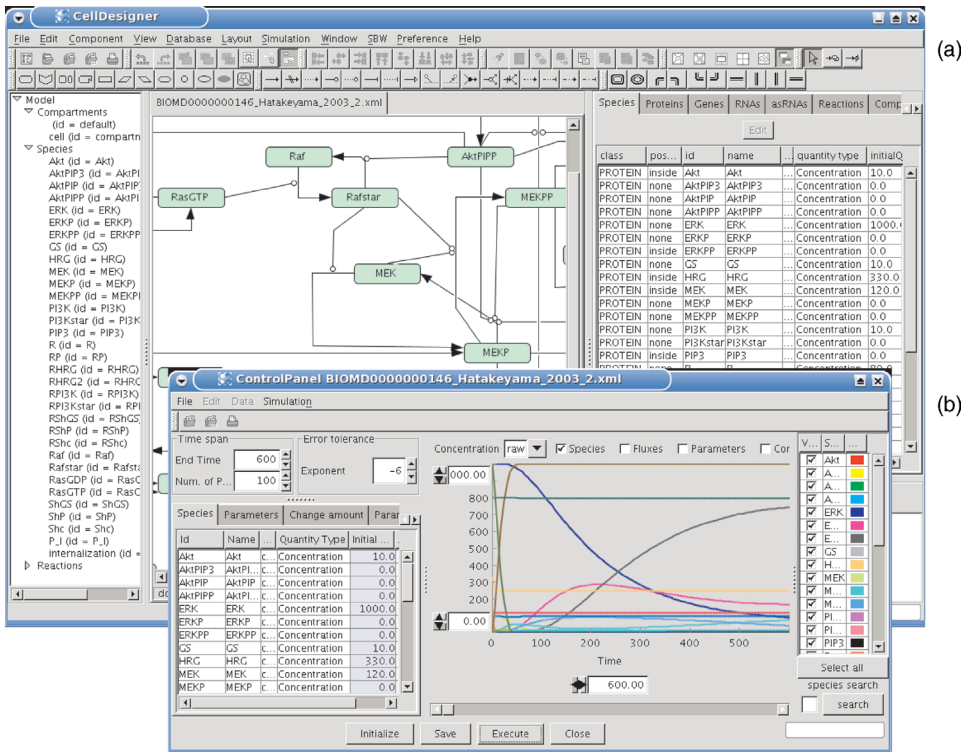
Figure 2.13 CellDesigner's process diagram editor (a) supports a rich set of graphical elements for different cellular species and reaction types. Simulations can be performed in CellDesigner using its integrated simulation engine (b).

concentrations and reaction details are entered. CellDesigner allows entering arbitrary kinetic equations, but has unfortunately no list of standard kinetics (mass action or Michaelis–Menten) that could be applied. For each reaction, the rate law has to be typed in by hand. A connection to the Systems Biology Workbench (SBW, see Section 17.4) is realized via the SBW menu and provides an interface to communicate with other SBW-compliant programs. For a further introduction to CellDesigner, a tutorial can be obtained at its website (http://www.celldesigner.org/). A movie introducing the usage of CellDesigner is availabe from the website of this book.

2.4.2.2 COPASI

Another platform-independent and user-friendly biochemical simulator that offers several unique features is COPASI [55]. COPASI is the successor to Gepasi [56, 57]. Its current version is 4.4 (http://www.copasi.org/). COPASI does not have such a rich visualization of the reaction network as CellDesigner, but it provides advanced functionalities for model simulation and analysis. In contrast to many other tools, it can switch between stochastic and deterministic simulation methods and supports hybrid deterministic-stochastic methods.
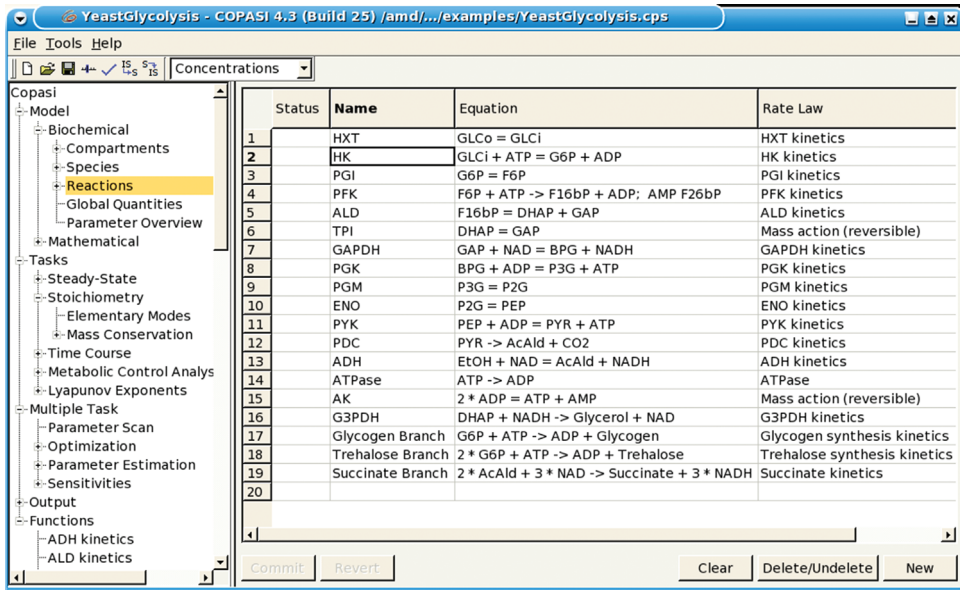
Figure 2.14 The different functionalities of COPASI are arranged in a hierarchical menu at left-hand side of its user interface. Details about the individual methods are listed in the right panel.

The user interface has a hierarchical menu (Figure 2.14, left side) that provides access to all the different functionalities of the tool. The biochemical model can be browsed according to its compartments, metabolites, and reactions including detailed list of the initial concentrations and kinetic parameters of the model. COPASI has a comprehensive set of standard methodologies for model analysis. It comprises the computation of steady states and their stability, supports the analysis of the stoichiometric network, e.g., the computation of elementary modes [25], supports MCA, and has methods for the optimization and parameter estimation. For compatibility with other tools, COPASI also supports the import and export of SBML-based models. For the definition of the kinetics, COPASI provides a copious set of predefined kinetic laws to choose from. A movie that is introducing the usage of COPASI is available from the website of this book.

2.4.2.3 PyBioS

Similarly as CellDesigner and Copasi, also PyBioS is designed for applications in systems biology and supports modeling and simulation [53]. PyBioS is a web-based environment (http://pybios.molgen.mpg.de/) that provides a framework for the conduction of kinetic models of various sizes and levels of granularity. The tool is a modeling platform for editing and analyzing biochemical models in order to predict the time-dependent behavior of the models. The platform has interfaces to external pathway databases (e.g., Reactome and KEGG) that can directly be used during model development for the definition of the structure of the reaction system. Figure 2.15
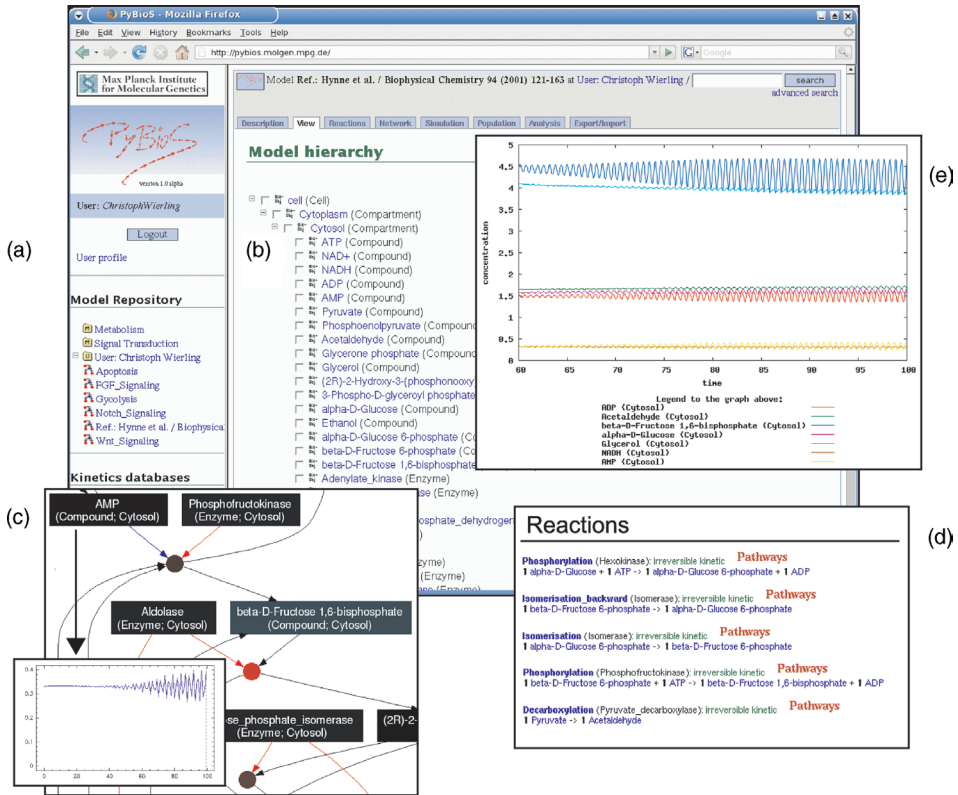
Figure 2.15 The PyBioS simulation environment. A particular model can be selected from the model repository (a) and its hierarchical model structure can be inspected via the View-tab at the top of the browser-window (b). A graphical representation of the model is provided by an automatically generated network diagram (accessible via the Network-tab), for example (c) shows the forward and reverse reaction of the isomerization of glucose-phosphate to fructose-phosphate of a glycolysis model. The Reactions-tab offers an overview of all reactions of the model (d). Simulations can be performed via the Simulation-tab (e). A simulation is based on an automatically generated mathematical model derived from the corresponding object-oriented model that comprises the network of all reactions and their respective kinetics.

shows screenshots of the PyBioS modeling and simulation environment. PyBioS defines a set of object classes (e.g., cell, compartment, compound, protein, complex, gene) for the definition of hierarchical models. Models are stored in a model repository. Support for the export and import of SBML-based models makes the platform compatible with other systems biology tools. Besides time course simulation, PyBioS also provides analysis methods, e.g., for the identification of steady states and their stability or for sensitivity analysis, such as the analysis of the steady-state behavior versus a varying parameter value or the computation of metabolic control coefficients. The reaction network of a model or individual parts of it can be visualized by network diagrams of the model components and their reactions that are

connected via edges. Time course results of simulation experiments can be plotted into the network graphs and used for the interpretation of the model behavior.

2.4.3
Data Formats

The documentation and exchange of models need to be done in a defined way. In the easiest way – as usually found in publications – the biochemical reactions and the mathematical equations that are describing the model can be listed, using common formalism for the representation of biochemical and mathematical equations. These conventions provide a good standard for the documentation and exchange in publications. However, these formats are suitable for humans but not for the direct processing by a computer. This gave rise to the development of standards for the description of models. During the last years, the eXtensible Markup Language (XML, http://www.w3.org/XML) has been proved to be a flexible tool for the definition of standard formats. In the following text, a brief introduction to XML as well as a description of SBML, a standard for model description that is based on XML, is given. Moreover, BioPAX, a standard for the description cellular reaction systems, and SBGN, a standard for the graphical representation of reaction networks, will be described.

2.4.3.1 Systems Biology Markup Language

The Systems Biology Markup Language (SBML, http://www.sbml.org) is a free and open format for the representation of models common to research in many areas of computational biology, including cell signaling pathways, metabolic pathways, gene regulation, and others [58]. It is already supported by many software tools [59]. In January 2009, the SBML homepage listed more than 110 software systems supporting SBML. Currently, there are two SBML specifications denoted Level 1 and Level 2. Level 2 is the most recent specification and therefore it is described in the following text.

SBML is defined as an XML compliant format. XML documents are written as plain text and have a very clear and simple syntax that can easily be read by both humans and computer programs; however, it is generally intended to be written and read by computers, not humans. In XML, information is associated with tags indicating the type or formatting of the information. Tags are used to delimit and denote parts of the document or to add further information to the document structure. Using miscellaneous start tags (e.g., `<tag>`) and end tags (e.g., `</tag>`), information can be structured as text blocks in a hierarchical manner.

Example 2.17

The following example of the phosphorylation reaction of aspartate catalyzed by the aspartate kinase illustrates the general structure of an SBML file.

$$\text{Aspartate} + \text{ATP} \xrightarrow{\text{Aspartate kinase}} \text{Aspartyl phosphate} + \text{ADP}$$
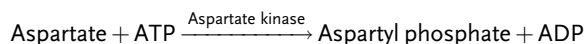
```
(1) <?xml version="1.0" encoding="UTF-8"?>
(2)  <sbml level="2" version="1" xmlns="http://www.
sbml.org/sbml/level2">
(3)  <model id="AK_reaction">
(4)   <listOfUnitDefinitions>
(5)    <unitDefinition id="mmol">
(6)     <listOfUnits>
(7)      <unit kind="mole" scale="-3" />
(8)     </listOfUnits>
(9)    </unitDefinition>
(10)    <unitDefinition id="mmol_per_litre_per_sec">
(11)     <listOfUnits>
(12)      <unit kind="mole" scale="-3" />
(13)      <unit kind="litre" exponent="-1" />
(14)      <unit kind="second" exponent="-1" />
(15)     </listOfUnits>
(16)    </unitDefinition>
(17)   </listOfUnitDefinitions>
(18)   <listOfCompartments>
(19)     <compartment id="cell" name="Cell" size="1"
units="volume" />
(20)   </listOfCompartments>
(21)   <listOfSpecies>
(22)     <species id="asp" name="Aspartate"
compartment="cell" initialConcentration="2"
substanceUnits="mmol" />
(23)     <species id="aspp" name="Aspartyl phosphate"
compartment="cell" initialConcentration="0"
substanceUnits="mmol" />
(24)   <species id="atp" name="ATP" compartment="cell"
initialConcentration="0" substanceUnits="mmol" />
(25)   <species id="adp" name="ADP" compartment="cell"
initialConcentration="0" substanceUnits="mmol" />
(26)   </listOfSpecies>
(27)   <listOfReactions>
(28)    <reaction id="AK" reversible="false">
(29)     <listOfReactants>
(30)      <speciesReference species="asp"
stoichiometry="1" />
(31)      <speciesReference species="atp"
stoichiometry="1" />
(32)     </listOfReactants>
(33)     <listOfProducts>
(34)      <speciesReference species="aspp"
stoichiometry="1" />
```

```
(35)     <speciesReference species="adp" stoichiometry=
"1" />
(36)     </listOfProducts>
(37)     <kineticLaw>
(38)      <math xmlns="http://www.w3.org/1998/Math/
MathML">
(39)       <apply>
(40)        <times />
(41)         <ci> k </ci>
(42)         <ci> asp </ci>
(43)         <ci> atp </ci>
(44)         <ci> cell </ci> <ci> cell </ci>
(45)        </apply>
(46)       </math>
(47)       <listOfParameters>
(48)       <parameter id="k" value="2.25" units="per_mM_
and_min" />
(49)       </listOfParameters>
(50)     </kineticLaw>
(51)    </reaction>
(52)   </listOfReactions>
(53)  </model>
(54) </sbml>
```

Line 1 in the above example defines the document as a XML document. The SBML model is coded in lines 2–54. It is structured into several lists that define different properties of the model. Most important lists that are usually used are the definition of units (lines 4–17), of compartments (lines 18–20), of species (lines 21–26), and finally of the reactions themselves (lines 27–52). Most entries in SBML have one required attribute, id, to give the instance a unique identifier by which other parts of the SBML model definition can refer to it. Some base units, like gram, meter, liter, mole, second, etc., are already predefined in SBML. More complex units derived from the base units are defined in the list of units. For instance, mM/s that is equal to $\text{mmol} \cdot l^{-1} \sec^{-1}$ can be defined as shown in lines 10–16 and used by its id in the subsequent definition of parameters and initial concentrations. Compartments are used in SBML as a construct for the grouping of model species. They are defined in the list of compartments (lines 18–20) and can be used not only for the definition of cellular compartments but also for grouping in general. Each compartment can have a name attribute and defines a compartment size. Model species are defined in the list of species. Each species has a recommended id attribute that can be used to refer it and can define its name and initial value with its respective unit. Species identifiers are used in the list of reactions (lines 27–52) for the definition of the individual biochemical reactions. Reversibility of a reaction is indicated by an attribute of the reaction tag

(lines 28). Reactants and products of a specific reaction along with their respective stoichiometry are specified in separate lists (lines 29–36).

The kinetic law of an individual reaction (lines 37–50) is specified in MathML for SBML Level 2. MathML is an XML-based markup language especially created for the representation of complicated mathematical expressions. In the above example, the rate law reads $k \cdot [asp] \cdot [atp] \cdot cell^2$, where k is a kinetic parameter [asp] and [atp] are the concentrations of aspartate and ATP, respectively, and cell is the volume of the cell. The consideration of the cell volume is needed, since rate laws in SBML are expressed in terms of amount of substance abundance per time instead of the traditional expression in terms of amount of substance concentration per time. The formulation of the rate law in the traditional way embodies the tacit assumption that the participating reaction species are located in the same, constant volume. This is done because attempting to describe reactions between species located in different compartments that differ in volume by the expression in terms of concentration per time quickly leads to difficulties.

2.4.3.2 BioPAX

Another standard format that is used in systems biology and designed for handling information on pathways and topologies of biochemical reaction networks is BioPAX (http://www.biopax.org). While SBML is tuned toward the simulation of models of molecular pathways, BioPAX is a more general and expressive format for the description of biological reaction systems even it is lacking definitions for the representation of dynamic data such as kinetic laws and parameters. BioPAX is defined by the BioPAX working group (http://www.biopax.org/). The BioPax Ontology defines a large set of classes for the description of pathways, interactions, and biological entities as well as their relations. Reaction networks described by BioPAX can be represented by the use XML. Many systems biology tools and databases make use of BioPAX for the exchange of data.

2.4.3.3 Systems Biology Graphical Notation

Graphical representations of reaction networks prove as very helpful tools for the work in systems biology. The graphical representation of a reaction system is not only helpful during the design of a new model and as a representation of the model topology, it is also helpful for the analysis and interpretation for instance of simulation results. Traditionally, diagrams of interacting enzymes and compounds have been written in an informal manner of simple unconstrained shapes and arrows. Several diagrammatic notations have been proposed for the graphical representation (e.g., [60–64]). As a consequence of the different proposals, the Systems Biology Graphical Notation (SBGN) has been set up recently. It provides a common graphical notation for the representation of biochemical and cellular reaction networks. SBGN defines a comprehensive set of symbols, with precise semantics, together with detailed syntactic rules defining their usage. Furthermore, SBGN defines how such graphical information is represented in a machine-readable form to ensure its proper storage, exchange, and reproduction of the graphical representation.

SBGN defines three different diagram types: (i) State Transition diagrams that are depicting all molecular interactions taking place, (ii) Activity Flow diagrams that are representing only the flux of information going from one entity to another, and (iii) Entity Relationship diagrams that are representing the relationships between different molecular species. In a State Transition diagram, each node represents a given state of a species, and therefore a given species may appear multiple times. State Transition diagrams are suitable for following the temporal process of interactions. A drawback of State Transition diagrams, however, is that the representation of each individual state of a species results quickly in very large diagram and due to this, it becomes difficult to understand what interactions actually exist for the species in question. In such a case, an Entity Relation diagram is more suitable. In an Entity Relation diagram, a biological entity appears only once.

SBGN defines several kinds of symbols, whereas two types of symbols are distinguished: nodes and arcs. There are different kinds of nodes defined. Reacting state or entity nodes represent, e.g., macromolecules, such as protein, RNA, DNA, polysaccharide, or simple chemicals, such as a radical, an ion, or a small molecule. Container nodes are defined for the representation of a complex, compartment, or module. Different transition nodes are defined for the representation of transitions like biochemical reactions, associations, like protein-complex formation, or dissociations, like the dissociation of a protein complex. The influence of a node onto another is visualized by different types of arcs representing, e.g., consumption, production, modulation, stimulation, catalysis, inhibition, or trigger effect. Not all node and arc symbols are defined for each of the three diagram types. A detailed description of the different nodes, arcs, and the syntax of their usage by the different diagram types is given in the specification of SBGN (see http://sbgn.org/).

Examples of a State Transition and an Entity Relationship diagram is given in Figure 2.16.

2.4.3.4 Standards for Systems Biology

With the increasing amount of data in modern biology the requirement of standards used for data integration became more and more important. For example, in the course of a microarray experiment, a lot of different information accumulates, as information about the samples, the type of microarray that is used, the experimental procedure including the hybridization experiment, the data normalization, and the expression data itself. It turns out that an important part of systems biology is data integration. This requires a conceptual design and the development of common standards.

The development of a standard involves four steps: an informal design of a conceptual model, a formalization, the development of a data exchange format, and the implementation of supporting tools [65]. For micorarray experiments, a conceptual model about the minimum information that is required for the description of such an experiment is specified by MIAME (Minimum Information About a Microarray Experiment [65]). Similar specifications have also been done for, e.g., proteomics data with MIAPE (Minimum Information About a Proteomics Experiment [66]), or systems biology models with MIRIAM (Minimum information

Figure 2.16 Systems Biology Graphical Notation (SBGN). (a) State transition diagram. (b) Entity relation diagram describing gene regulation and transcription of a gene. The two transcription factors TF A and TF A′ compete for the same transcription factor-binding site. If one of the transcription factors is bound, the binding site is blocked for the other one, but only TF A can activate the transcription of the gene. The abbreviation "ct" indicates conceptual types of the respective entity.

requested in the annotation of biochemical models). MIRIAM specifies a set of rules for curating quantitative models of biological systems that define procedures for encoding and annotating models represented in machine-readable form [67].

2.4.4
Data Resources

The development of models of biological systems requires diverse kind of data. This is, for instance, information about the different model components (e.g., metabolites, proteins, and genes) and their different functions and interactions. Such information can be extracted from literature or dedicated data resources, like pathway databases. Two pathway databases that are well known are KEGG und Reactome. Both are described below in more detail. Another important data for modeling are information about reaction kinetics. Database dealing with such data are described in

more detail in Sections 2.4.4.2 and 3.1. Further information about databases providing primary data is given in Chapter 16.

2.4.4.1 Pathway Databases

Kyoto Enzyclopedia of Genes and Genomes   Kyoto Enzyclopedia of Genes and Genomes (KEGG; http://www.genome.ad.jp/kegg/) is a reference knowledge base offering information about genes and proteins, biochemical compounds and reactions, and pathways. The data is organized in three parts: the gene universe (consisting of the GENES, SSDB, and KO database), the chemical universe (with the COMPOUND, GLYCAN, REACTION, and ENZYME databases which are merged as LIGAND database), and the protein network consisting of the PATHWAY database [68]. Besides this, the KEGG database is hierarchically classified into categories and subcategories at four levels. The five topmost categories are metabolism, genetic information processing, environmental information processing, cellular processes, and human diseases. Subcategories of metabolism are, e.g., carbohydrate, energy, lipid, nucleotide, or amino acid metabolism. These are subdivided into the different pathways, like glycolysis, citrate cycle, purine metabolism, etc. Finally, the fourth level corresponds to the KO (KEGG Orthology) entries. A KO entry (internally identified by a K number, e.g., K00001 for the alcohol dehydrogenase) corresponds to a group of orthologous genes that have identical functions.

The gene universe offers information about genes and proteins generated by genome sequencing projects. Information about individual genes is stored in the GENES database, which is semiautomatically generated from the submissions to GenBank, the NCBI RefSeq database, the EMBL database, and other publicly available organism-specific databases. K numbers are further assigned to entries of the GENES database. The SSDB database contains information about amino acid sequence similarities between protein-coding genes computationally generated from the GENES database. This is carried out for many complete genomes and results in a huge graph depicting protein similarities with clusters of orthologous and paralogous genes.

The chemical universe offers information about chemical compounds and reactions relevant to cellular processes. It includes more than 11,000 compounds (internally represented by C numbers, e.g., C00001 denotes water), a separate database for carbohydrates (nearly 11,000 entries; represented by a number preceded by G, e.g., G10481 for cellulose), more than 6000 reactions (with R numbers, e.g., R00275 for the reaction of the superoxide radical into hydrogen peroxide), and more than 4000 enzymes (denoted by EC numbers as well as K numbers for orthologous entries). All these data are merged as LIGAND database [69]. Thus, the chemical universe offers comprehensive information about metabolites with their respective chemical structures and biochemical reactions.

KEGG's protein network provides information about protein interactions comprising pathways and protein complexes. The 235 KEGG reference pathway diagrams (maps), offered on the website, give clear overviews of important pathways. Organism-specific pathway maps are automatically generated by coloring of organism-specific genes in the reference pathways.

The KEGG database can be queried via the web interface, e.g., for genes, proteins, compounds, etc. Access to the data via FTP (http://www.genome.ad.jp/anonftp) as well as access to it via a SOAP server (http://www.genome.ad.jp/kegg/soap) is possible for academic users, too.

Reactome   Reactome (formerly known as Genome Knowledgebase [70–72]) is an open, online database of fundamental human biological processes. The Reactome project is managed as a collaboration of the Cold Spring Harbor Laboratory, the European Bioinformatics Institute (EBI), and the Gene Ontology Consortium. The database is divided into several modules of fundamental biological processes that are thought to operate in humans. Each module of the database has one or more primary authors and is further peer reviewed by experts of the specific field. Each module can also be referenced by its revision date and thus can be cited like a publication.

On one hand, the Reactome database is intended to offer valuable information for the wet-lab scientist, who wants to know, e.g., more about a specific gene product she or he is unfamiliar with. On the other hand, the Reactome database can be used by the computational biologist to draw conclusions from large data sets like expression data gained by cDNA chip experiments.

Another tool offered by Reactome is the "Pathfinder." This utility enables the user to find the shortest path between two physical entities, e.g., the shortest path between the metabolites D-fructose and pyruvate, or the steps from the primary mRNA to its processed form. The computed path can be shown graphically. The pathfinder offers also the possibility to exclude specific entities, like the metabolites ATP or NADH that show high connectivity and thus their input might lead to a path that is not the one intended to be found.

Data from Reactome can be exported in various formats upon which are SBML and BioPAX.

2.4.4.2   Databases of Kinetic Data

High-throughput projects, such as the international genome sequencing efforts, accumulate large amounts of data at an amazing rate. These data are essential for the reconstruction of phylogenetic trees and gene-finding projects. However, for kinetic modeling, which is at the heart of systems biology, kinetic data of proteins and enzymes are needed. Unfortunately, this type of data is notoriously difficult and time-consuming to obtain since proteins often need individually tuned purification and reaction conditions. Furthermore, the results of such studies are published in a large variety of journals from different fields. In this situation, the databases BRENDA and SABIO-RK aim to be comprehensive resources of kinetic data. They are discussed in more detail in Section 4.1.1.

2.4.4.3   Model Databases

A lot of different mathematical models of biological systems have already been developed in the past and are described in the literature. However, these models are usually not available in a computer-amenable format. During the last years, big efforts have been done on the gathering and implementation of existing models in

databases. Two well-known databases on this are BioModels and JWS, which are described in more detail in the following.

BioModels   The BioModels.net project (http://biomodels.net) is an international effort to (i) define agreed-upon standards for model curation, (ii) define agreed-upon vocabularies for annotating models with connections to biological data resources, and (iii) provide a free, centralized, publicly accessible database of annotated, computational models in SBML, and other structured formats. The ninth release of the databases has 192 models, of which 150 are in the curated and 42 are in the noncurated branch. Models can be browsed in the web interface, online simulations can be performed via the external simulation engine of JWS online (see below), or they can be exported in several prominent file formats (e.g., SBML, CellML, BioPAX) for external usage by other programs.

JWS   Another model repository that is providing kinetic models of biochemical systems is JWS online [73]. As of February 2008, this model repository provides 84 models (http://jjj.biochem.sun.ac.za). Models in JWS online can be interactively run and interrogated over the internet.

Exercises and Problems

1. A canonical view of the upper part of glycolysis starts with glucose and comprises the following reactions (in brackets: possible abbreviations): The enzyme hexokinase (HK, E_1) phosphorylates glucose (Gluc, S_1) to glucose-6-phosphate (G6P, S_2) under consumption of ATP (S_5) and production of ADP (S_6). The enzyme phosphoglucoisomerase (PGI, E_2) converts glucose-6-phosphate to fructose-6-phosphate (F6P, S_3). The enzyme phosphofructokinase (PFK, E_3) phosphorylates F6P a second time to yield fructose-1,6-bisphosphate (F1,6,BP, S_4). The enzyme fructosebisphosphatase catalyzes the reverse reaction (E_4).

   (a) Sketch the reaction network and formulate a set of differential equations (without specifying the kinetics of the individual reactions).
   (b) Formulate the stoichiometric matrix **N**. What is the rank of **N**?
   (c) Calculate steady-state fluxes (matrix **K**) and conservation relations (matrix **G**).
   (d) Compare your results with Example 2.6.

2. (a) Write down the sets of differential equations for the networks N1–N6 given in Table 2.4 without specifying their kinetics.
   (b) Determine the rank of the stoichiometric matrices, independent steady-state fluxes, and conservation relations.
      Do all systems have a (nontrivial) steady state?

3. Inspect networks N3 and N4 in Table 2.4. Can you find elementary flux modes? Use an available tool (e.g., Metatool) to check out.

4. Assign the following kinetics to network N3 in Table 2.4: $v_1 = k_1$, $v_2 = (V_{max2} \cdot S_1)/(K_{m2} + S_1)$, $v_3 = (V_{max3} \cdot S_1)/(K_{m3} + S_1)$ with $k_1 = 10$, $V_{max2} = 3$, $K_{m2} = 0.2$, V_{max2} 5, and $K_{m2} = 0.4$. Compute the steady-state concentration of S_1 and calculate the flux control coefficients.

5. For the reaction system $A \xrightarrow{v_1} B$, $B \xrightarrow{v_2} C$, $C \xrightarrow{v_3} A$ with $v_1 = k_1 \cdot A$, $v_2 = k_2 \cdot B$, $v_3 = k_3 \cdot C$, and $k_1 = 2$, $k_2 = 2$, $k_3 = 1$, write down the set of systems equations.

   (a) Compute the Jacobian $\mathbf{J}$!
   (b) Determine the eigenvalues and eigenvectors of the Jacobian $\mathbf{J}$!
   (c) What is the general solution of the ODE system?
   (d) Compute the solution with the initial condition $A(0) = 1$, $B(0) = 1$, $C(0) = 0$!

6. The Jacobian $\mathbf{A}_a$ of the following ODE system depends on the parameter a:

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 10 + a & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

   (a) To every specific choice of parameter a belongs a point $(\text{Tr}\,\mathbf{A}_a, \text{Det}\,\mathbf{A}_a)$ in the plane spaned by trace and determinate of $\mathbf{A}_a$. Draw the curve $(\text{Tr}\,\mathbf{A}_a, \text{Det}\,\mathbf{A}_a)$ in this space for a as a changing parameter.
   (b) For which values of a is $(x, y) = (0, 0)$ a saddle point, node or focus?

7. What is the use of standards important for the development of new systems biology tools?

References

1 Guldberg, C.M. and Waage, P. (1879) Über die chemische Affinität. *Journal für Praktische Chemie*, **19**, 69.

2 Guldberg, C.M. and Waage, P. (1867) *Études sur les affinités chimiques*, Christiana.

3 Waage, P. and Guldberg, C.M. (1864) *Studies concerning affinity*, Forhandlinger: Videnskabs-Selskabet, Christiana, pp 35.

4 Brown, A.J. (1902) Enzyme action. *Journal of the Chemical Society*, **81**, 373–386.

5 Michaelis, L. and Menten, M.L. (1913) Kinetik der invertinwirkung. *Biochemistry Zeitung*, **49**, 333–369.

6 Briggs, G.E. and Haldane, J.B.S. (1925) A note on the kinetics of enzyme action. *The Biochemical Journal*, **19**, 338–339.

7 Lineweaver, H. and Burk, D. (1934) The determination of enzyme dissocation constants. *Journal of the American Chemical Society*, **56**, 658–660.

8 Eadie, G.S. (1942) The inhibition of cholinesterase by physostigmine and prostigmine. *The Journal of Biological Chemistry*, **146**, 85–93.

9 Hanes, C.S. (1932) Studies on plant amylases. I. The effect of starch concentratio upon the velocity of hydrolysis by the amylase of germinated barley. *The Biochemical Journal*, **26**, 1406–1421.

10 Haldane, J.B.S. (1930) *Enzymes*, Longmans, Green and Co., London.

11 Schellenberger, A. (1989) *Enzymkatalyse*, Fischer Verlag, Jena.

12 Wegscheider, R. (1902) Über simultane Gleichgewichte und die Beziehungen

zwischen Thermodynamik und Reaktionskinetik homogener Systeme. *Zeitschrift für Physikalische Chemie*, **39**, 257–303.

13 Hill, A.V. (1910) The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *The Journal of Physiology*, **40**, iv–vii.

14 Hill, A.V. (1913) The combinations of hemoglobin with oxygen and with carbon monoxide. *The Biochemical Journal*, **7**, 471–480.

15 Monod, J. *et al.* (1965) On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*, **12**, 88–118.

16 Savageau, M. (1985) Mathematics of organizationally complex systems. *Biomed Biochim Acta*, **44**, 839–884.

17 Heijnen, J.J. (2005) Approximative kinetic formats used in metabolic network modelling. *Biotechnology and Bioengineering*, **91**, 534–545.

18 Liebermeister, W. and Klipp, E. (2006) Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical Biology and Medical Modelling*, **3**, 42.

19 Glansdorff, P. and Prigogine, I. (1971) *Thermodynamic Theory of Structure, Stability and Fluctuations*, Wiley-Interscience, London.

20 Reder, C. (1988) Metabolic control theory: a structural approach. *Journal of Theoretical Biology*, **135**, 175–201.

21 Heinrich, R. and Schuster, S. (1996) *The Regulation of Cellular Systems*, Chapman & Hall, New York.

22 Michal, G. (1999) *Biochemical Pathways*, Spektrum Akademischer, Heidelberg.

23 Pfeiffer, T. *et al.* (1999) METATOOL: for studying metabolic networks. *Bioinformatics (Oxford, England)*, **15**, 251–257.

24 Schilling, C.H. *et al.* (1999) Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnology Progress*, **15**, 296–303.

25 Schuster, S. *et al.* (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, **17**, 53–60.

26 Schuster, S. *et al.* (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, **18**, 326–332.

27 Schuster, S. *et al.* (2002) Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *Journal of Mathematical Biology*, **45**, 153–181.

28 Schilling, C.H. and Palsson, B.O. (2000) Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *Journal of Theoretical Biology*, **203**, 249–283.

29 Schilling, C.H. *et al.* (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, **203**, 229–248.

30 Wiback, S.J. and Palsson, B.O. (2002) Extreme pathway analysis of human red blood cell metabolism. *Biophysical Journal*, **83**, 808–818.

31 Bronstein, I.N. and Semendjajew, K.A. (1987) Taschenbuch der Mathematik, 23rd edition, Nauka, Moscow.

32 Heinrich, R. and Rapoport, T.A. (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *European Journal of Biochemistry*, **42**, 89–95.

33 Kacser, H. and Burns, J.A. (1973) The control of flux. *Symposia of the Society for Experimental Biology*, **27**, 65–104.

34 Bruggeman, F.J. *et al.* (2002) Modular response analysis of cellular regulatory networks. *Journal of Theoretical Biology*, **218**, 507–520.

35 Hofmeyr J.H. and Westerhoff, H.V. (2001) Building the cellular puzzle: control in multi-level reaction networks. *Journal of Theoretical Biology*, **208**, 261–285.

36 Kholodenko, B.N. *et al.* (2000) Diffusion control of protein phosphorylation in

signal transduction pathways. *The Biochemical Journal*, **350** (Pt 3), 901–907.

37 Liebermeister, W. *et al.* (2004) A theory of optimal differential gene expression. *Bio Systems*, **76**, 261–278.

38 Westerhoff, H.V. *et al.* (2002) ECA: control in ecosystems. *Molecular Biology Reports*, **29**, 113–117.

39 de Jong, H. (2002) Modeling and simulation of genetic regulatory systems: a literature review. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, **9**, 67–103.

40 Kitano, H. (2002) Computational systems biology. *Nature*, **420**, 206–210.

41 Hindmarsh, A.C. (1983) ODEPACK, A systematized collection of ODE solvers, in *Scientific Computing* (eds R.S. Stepleman *et al.*) North-Holland, Amsterdam, pp. 55–64.

42 Petzold, L. (1983) Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM Journal on Scientific and Statistical Computing*, **4**, 136–148.

43 Cohen, S.D. and Hindmarsh, A.C. (1996) CVODE, a stiff/nonstiff ODE solver in C. *Computers in Physics*, **10**, 138–143.

44 Deuflhard, P. and Nowak, U. (1987) Extrapolation integrators for quasilinear implicit ODEs, in *Large Scale Scientific Computing* (eds P. Deuflhard and B. Engquist) Birkkäuser, Basel, pp. 37–50.

45 Bernardinello, L. and de Cindio, F. (1992) *A Survey of Basic Net Models and Modular Net Classes*, Springer, Berlin.

46 Küffner, R. *et al.* (2000) Pathway analysis in metabolic databases via differential metabolic display (DMD). *Bioinformatics*, **16**, 825–836.

47 Reddy, V.N. *et al.* (1996) Qualitative analysis of biochemical reaction systems. *Computers in Biology and Medicine*, **26**, 9–24.

48 Matsuno, H. *et al.* (2003) Biopathways representation and simulation on hybrid functional Petri net. *In Silico Biology*, **3**, 389–404.

49 Gardner, M. (1970) Mathematical games. *Scientific American*, **223**, 120–123.

50 Alves, R. *et al.* (2006) Tools for kinetic modeling of biochemical networks. *Nature Biotechnology*, **24**, 667–672.

51 Klipp, E. *et al.* (2007b) Systems biology standards – the community speaks. *Nature Biotechnology*, **25**, 390–391.

52 Materi, W. and Wishart, D.S. (2007) Computational systems biology in drug discovery and development: methods and applications. *Drug Discovery Today*, **12**, 295–303.

53 Wierling, C. *et al.* (2007) Resources, standards and tools for systems biology. *Briefings in Functional Genomics and Proteomics*, **6**, 240–251.

54 Funahashi, A. *et al.* (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico*, **1**, 159–162.

55 Hoops, S. *et al.* (2006) COPASI – a COmplex PAthway SImulator. *Bioinformatics*, **22**, 3067–3074.

56 Mendes, P. (1993) GEPASI: a software package for modelling the dynamics steady states and control of biochemical and other systems. *Computer Applications in the Biosciences*, **9**, 563–571.

57 Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends in Biochemical Sciences*, **22**, 361–363.

58 Hucka, M. *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**, 524–531.

59 Hucka, M. *et al.* (2004) Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project. *Systematic Biology (Stevenage)*, **1**, 41–53.

60 Kitano, H. (2003) A graphical notation for biochemical networks. *BIOSILICO*, **1**, 169–176.

61 Kitano, H. *et al.* (2005) Using process diagrams for the graphical representation of biological networks. *Nature Biotechnology*, **23**, 961–966.

62 Kohn, K.W. (1999) Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Molecular Biology of the Cell*, **10**, 2703–2734.

63 Moodie, S.L. *et al.* (2006) A graphical notation to describe the logical interactions of biological pathways. *Journal of Integrative Bioinformatics*, **3**, 36.

64 Pirson, I. *et al.* (2000) The visual display of regulatory information and networks. *Trends in Cell Biology*, **10**, 404–408.

65 Brazma, A. *et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics*, **29**, 365–371.

66 Taylor, C.F. *et al.* (2007) The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnology*, **25**, 887–893.

67 Le Novere, N. *et al.* (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotechnology*, **23**, 1509–1515.

68 Kanehisa, M. *et al.* (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Research*, **32**, D277–D280.

69 Goto, S. *et al.* (2002) LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Research*, **30**, 402–404.

70 Joshi-Tope, G. *et al.* (2003) The genome knowledgebase: a resource for biologists and bioinformaticists. *Cold Spring Harbor Symposia on Quantitative Biology*, **68**, 237–243.

71 Joshi-Tope, G. *et al.* (2005) Reactome: a knowledgebase of biological pathways. *Nucleic Acids Research*, **33**, D428–D432.

72 Vastrik, I. *et al.* (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biology*, **8**, R39.

73 Olivier, B.G. and Snoep, J.L. (2004) Web-based kinetic modelling using JWS Online. *Bioinformatics*, **20**, 2143–2144.