In Vitro Versus In Vivo: Concepts and Consequences

38

(The biochemist's word) may not be the last in the description of life, but without his help the last word will never be said.

Sir Gowland Hopkins

Biochemists enjoy the freedom to purify and study enzymes in isolation, saturate an enzyme with its substrate, trap/remove the products, and also provide optimal pH, ionic strength, etc. On the other hand, cell extracts are by their very nature "dirty enzymes"; intact cells and organisms are "dirtier" still. The cell by design is greatly constrained to provide a consensus medium to simultaneously support hundreds of diverse enzyme-catalyzed reactions. Only some of these enzymes may be operating under optimal conditions at any time. The context for an enzyme to function in vivo is very different from the well-defined conditions deliberately set up for its study in vitro. And classical biochemistry is founded on several assumptions valid in dilute aqueous solutions. These assumptions are often extended without question to the cellular milieu. But the cell interior is far away from being an ideal solution. The key features that differentiate the state of affairs in vivo from that in vitro are cataloged below.

Organization and Compartmentalization The enzyme study in a test tube presupposes that the solution is homogeneous. This assumption is not valid for intact cells. The cytoplasm may be better described as an aqueous gel than as a homogeneous solution. The presence of supramolecular organizations and membrane-bounded sub-compartments (in eukaryotic cells in particular) confers intrinsic inhomogeneities on the cell interior. Additionally, the organelles themselves are not randomly distributed due to cytoskeleton organization and intracellular transport.

N. S. Punekar, *ENZYMES: Catalysis, Kinetics and Mechanisms*,
https://doi.org/10.1007/978-981-13-0785-0_38

The Dilution Factor Due to the small, finite volume of cells and intracellular compartments, the actual number of a particular molecular species may be surprisingly low. Very few molecules per cell could translate into high, physiologically relevant molar concentrations. These concentrations may be far higher than those at which enzymes are usually assayed. Weak protein associations could well disappear at the far lower protein concentrations with which enzyme kineticists work in vitro.

Concentration of Enzymes, Substrates, and Other Ligands The total concentration of macromolecules inside cells is very high, with proteins being most abundant species. Clearly the aqueous phase of the cytoplasm is crowded rather than dilute. Such crowded solutions are not amenable to the fundamental assumption of the physical chemistry of dilute solutions – interactions between solute molecules cannot be neglected. The concentration of a specific enzyme is generally much higher in cells than in conventional in vitro assays. This may be often much in excess of the K_D as determined in vitro. For instance, the K_D for interaction between calmodulin and myosin light chain kinase in vitro is around 1.0 nM. But the smooth muscle calmodulin concentration is in the order of 40 μM. Similarly, a substantial proportion of substrates may exist as complexes, making the availability of free substrate rate limiting.

Enzymes In Vivo Are Components of an Open System The cell is not a bag of enzymes each working in isolation. The cellular metabolism is a complex web of enzyme pathways, many of which often compete for common substrates. In order to proceed efficiently under these conditions, channeling of substrates from one enzyme to another in a particular metabolic pathway may be necessary. Discrimination among competing interactions may be achieved by sequestering enzymes within an organelle or immobilizing them on a membrane.

We will now elaborate on a few of these key concepts and their consequences on enzyme action in vivo.

38.1 Why Michaelis-Menten Formalism Is Not Suitable In Vivo

Much of the current paradigm to understand enzymes has been extrapolated from studies of dilute solutions containing a single enzyme and its cognate substrate – whose interaction is diffusion limited. Enzymology in the "test tube" is a careful observation and controlled study of an enzyme in isolation from the host of other interactions that otherwise make the understanding difficult. This reductionist approach has led to many valuable insights over the past century. However, accumulated knowledge on measurements of the physical properties of cells indicates that the interior of a cell departs from these ideal conditions in several important ways. Some of these are already listed above.

Table 38.1 Concentration of individual glycolytic enzymes and intermediates

Enzyme	Active site (μM)	Metabolite (substrate)	Concentration (μM)
Phosphofructokinase	24.1	Fructose 6-phosphate	1500
Aldolase	809.3	Fructose 1,6-bisphosphate	80
Triosephosphate isomerase	223.8	Dihydroxyacetone phosphate	160
Glyceraldehyde 3-phosphate dehydrogenase	1398.6	Glyceraldehyde 3-phosphate	80
Phosphoglycerate kinase	133.6	1,3-Bisphospho-glycerate	50
Pyruvate kinase	172.9	Phosphoenolpyruvate	65
Lactate dehydrogenase	296.0	Pyruvate	380

Adapted from Srivastava & Bernhard, Curr Top Cell Regul, 28:1–68. Copyright (1986), with permission from Elsevier

We recall that the Michaelis-Menten rate equation is derived for a well-defined initial conditions and making certain clear assumptions (see Chap. 15). Almost all of these may not be satisfied with respect to enzymes in vivo.

Where Assumptions Fail Enzyme being a catalyst, its assay concentration in vitro is usually held very much lower than that of the substrate. This permits us to approximate $[S_t] \approx [S]$, although $[S_t]$ actually equals "$[S] + [ES]$" while deriving the rate equation. Most often the substrate concentrations are held at least 1000 times higher than that of the enzyme (i.e., $[S_t] >> [E_t]$). The actual $[E_t]/[S_t]$ ratio becomes important in interpreting enzyme kinetic behavior under in vivo conditions. This ratio for various enzymes of glycolysis (in mammalian skeletal muscle) ranges from 0.016 (for phosphofructokinase) to 17.48 (glyceraldehyde 3-phosphate dehydrogenase) (Table 38.1). The assumption $[S_t] \approx [S]$ is valid only when this ratio is very low – such as in the case of phosphofructokinase. But when the ratio approaches/exceeds the value of 0.4, serious deviations from the Michaelis-Menten formalism occur, and the corresponding rate equations are no longer valid.

Reaction Reversibility Strictly the initial rate (velocity "v") has to be recorded in order to apply Michaelis-Menten formalism. This will be the unbiased rate when $[P] \approx 0$, whereas significant product concentrations are often observed in vivo. Products compete with substrates for the available enzyme; by reaction reversal they can also react to form substrate. Increasing $[P]$ can actually generate negative velocity, especially when $[S]$ is not high enough to force the reaction forward. Entire pathways, and most enzymes in such a pathway, are known to work in reverse under certain physiological states. For example, but for two irreversible steps, most enzymes of glycolysis operate in reverse during gluconeogenesis.

Optimized In Vitro But Compromised In Vivo An enzymologist's test tube is optimized for the best measurements of enzyme activity. The parameters like temperature, buffers, pH, ionic strength, cofactors, metal ion activators, etc. are rigorously controlled by the experimenter. A wide range of these conditions can be tested in vitro. However, a given physiological state of the cell provides a single consensus medium common for all its enzymes functioning in that compartment. It is therefore rare that kinetic constants determined in vitro are accurate reflections on the enzymes in vivo. Rate laws determined in vitro with purified, dilute, homogeneous enzyme solutions may not reflect the enzyme–enzyme interactions that are important in vivo. Also, many small molecule regulators may be missed due to clean in vitro (test tube) assays. Optimal behavior for an isolated enzyme may mean nonoptimal behavior for the intact system and *vice versa*. For instance, arginase functions inside a fungal cytoplasm at physiological (near neutral) pH and with in situ Mn [II] concentrations. This is very different from its alkaline pH optimum (of around pH 10.0) and the requirement for incubations with micromolar Mn[II] for optimal activity in the test tube. Regulatory interaction of yeast arginase with ornithine transcarbamylase (the so-called epiarginasic control) is also well documented (Table 37.4, Chap. 37).

In order to make the bottoms-up approach to *Systems Biology* meaningful, standardized in vivo-like conditions are being simulated to study enzyme function. One such standardized assay medium for all yeast cytosolic enzymes was recently described. Potential effects of macromolecular crowding (see section on "Diffusion, Crowding, and Enzyme Efficiency") are still missed and not considered in this standardization effort (Table 38.2). Besides properly defined ionic strength, pH, and buffering, other components of the intracellular milieu should be included in enzyme assays to simulate in vivo conditions. Glutamate, glutathione, and phosphates are the main metabolite pool components that contribute to pH buffering in most cells. Effects of all these factors on k_{cat} as well as on K_M have to be accounted for. These measurements should also include data when [*S*] is well above K_M (effects on k_{cat}) as well as around/below K_M (effects on K_M).

Table 38.2 Standardized assay medium for yeast cytosolic enzymes

Component	Concentration
Potassium (K^+)	300 mM
L-Glutamate	245 mM
Phosphate	50 mM
Sodium (Na^+)	20 mM
Magnesium (Mg^{2+})	2 mM (free)
Sulfate	2.5–10.0 mM (depending on Mg^{2+} levels)
Calcium (Ca^{2+})	0.5 mM
Medium pH	6.8

Adapted with permission from van Eunen et al., FEBS Journal, 277:749–760. Copyright (2010) John Wiley & Sons Inc

Table 38.3 Enzyme systems in vitro and in vivo: a comparison

Nature of	Enzyme in vitro	Enzyme in vivo
Variables	Few	Many
Interactions	Weak	Strong
Connectivity	Linear	Nonlinear
Processes	Additive	Associative
Aggregate behavior	Predictive	Emergent
System	Closed, simple	Open, complex

A study of enzymes in vivo quickly moves into the realm of metabolism and complexity (Table 38.3). One could build an understanding of metabolism (i.e., enzymology in vivo) by studying individual enzymes one at a time and then upwardly integrating this in vitro knowledge (Smallbone et al. 2013). In practice this is very complex – collecting such data for regulatory enzymes with multiple interactions becomes immensely difficult. This "bottoms-up" approach is like describing the behavior of a gas by applying Newton's laws of motion to every individual molecule in the system. More often, enzymology in vivo takes a "top-down" approach – akin to applying simple gas laws of thermodynamics at the macroscopic level. It is important to distinguish between kinetics as the study of molecular mechanisms and the kinetics of system dynamics. The very conditions that made Michaelis-Menten formalism produce the rate law for enzyme reactions in vitro tend to make it invalid for enzyme reactions in vivo. Power-law formalism as an alternative to capture in vivo behavior has been proposed. Interested reader may refer to this literature (Savageau 1992).

38.2 Concentration of Enzymes, Substrates, and Their Equilibria

There is always a certain tension between the in vitro and in vivo approaches to enzymology. One aspect that biochemists have struggled with since the beginning is the extent to which in vitro enzyme data are relevant to in vivo metabolism. Cells and biological tissues are disrupted to extract and access enzymes prior to assay. Cell-free extracts are considerably diluted when compared to enzyme/metabolite concentrations in vivo (Albe et al. 1990; Bennett et al. 2009). Invariably, extrapolations from in vitro to in vivo involve assumptions, and this often leads to poor conclusions. All the intracellular compartments are highly concentrated in terms of proteins and metabolites (Srere 1967; 1970). This also has a direct bearing on the physical features of the cytosol such as viscosity, diffusion rates, and excluded volume effects (discussed in detail a little later). We are just beginning to understand the influence of unique intracellular environment on enzyme activities.

A major experimental challenge in doing in vivo enzymology is the precise measurement and manipulation of *absolute concentrations* of enzymes and metabolites inside the cell. There are many issues that influence such numbers.

Table 38.4 Dimensions and content of different cell types

Source	Cell dimensions (volume)	Water content (g/100 g wet cells)	Protein concentration (mg/ml)
E. coli	1 × 3 μm (2 μm^3)	70	235
Yeast	5 μm (66 μm^3)	65	280
Rat liver	10–20 μm (500–4000 μm^3)	69	313
Rat muscle		77	260
Human RBC	6–8 μm (90 μm^3)	65	300
Mitochondrion	1 μm (0.5 μm^3)	50	270–560

A volume of 3.2×10^{-15} l has been directly measured for *E. coli* cells
[1 $\mu m^3 = 10^{-15}$ l (1 femtoliter); 10^3 $\mu m^3 = 10^{-12}$ l (1 picoliter)]

- Direct determination of metabolite concentrations is difficult for two reasons: (a) concentrations may change rapidly during the time required for isolation and (b) low-molecular-weight substances diffuse out of organelles and redistribute during isolation. Based on the total ATP, ADP, creatine, and phosphocreatine concentrations measured in the cell, it was thought that creatine kinase reaction is far from equilibrium (toward phosphocreatine formation). Subsequently, the free concentrations of the phosphorylated compounds were directly measured in living cells using ^{31}P NMR. With this data it was shown that the creatine kinase reaction is close to equilibrium in vivo.
- Cellular enzyme concentrations tend to be underestimated, mainly due to incomplete extraction and associated inactivation during isolation from crude extracts.
- Despite very accurate analytical methods for estimation, serious assumptions/approximations are necessary to compute the final in vivo concentrations. For instance, the water content and the total soluble protein in different cell types (see Table 38.4) are averaged as if the cytoplasm is a homogeneous solution. Typically, in a gram of muscle tissue, all the cytoplasmic soluble components are found in about 0.75 ml of water.
- The metabolite concentrations are usually reported in the literature in units of μmol per g of dry or fresh weight. Such data can be converted to μM using suitable conversion factors (see Table 38.4 above). Arriving at in vivo enzyme concentrations requires some more rough estimations: (i) turnover number (μmol substrate converted × min^{-1} × μmol $enzyme^{-1}$) is calculated from the specific activity (μmol substrate converted min^{-1} × mg $protein^{-1}$) of the most purified enzyme fraction available and the molecular weight of the holoenzyme, assuming that all protein represents active holoenzyme; (ii) the V_{max} value (μmol substrate converted min^{-1} × liter cell $volume^{-1}$), assumed to represent in vivo activity, is calculated from the specific activity of a crude enzyme fraction using the conversion factors (see Table 38.4 above); and (iii) enzyme concentration is now obtained by dividing the V_{max} value (from ii) by the turnover number (from i). The total enzyme site concentration can finally be represented by multiplying enzyme concentration by the number of subunits per holoenzyme, assuming one active site per subunit.

Despite inherent difficulties in measurements, there are some obvious take-home messages from the calculated in vivo concentrations of enzymes and metabolites (Srivastava and Bernhard 1986; Storey 2005). For instance, consider the computed concentrations for glycolytic enzymes and their substrates in the mammalian skeletal muscle (see Table 38.1 above). First, among the enzymes of the same pathway, some are much more abundant than the rest – aldolase and glyceraldehyde 3-phosphate dehydrogenase together constitute 40–50% of the total glycolytic enzymes. Molar concentrations of such enzymes in vivo are quite high. Notwithstanding this range, the sequence of enzymes sustains a common single flux through glycolysis. And further, the muscle glycolytic flux increases 100–1000-fold during "'resting-working" transition without any change in the concentration of its intermediates. Second, the ratios of final calculated enzyme site concentration (expressed as μM) to corresponding free substrate (in μM) concentration(s) are also quite distinct (free metabolite concentrations can be determined by solving simultaneous equations relating the free and bound concentrations of all metabolites; the corresponding Michaelis constants may be used to approximate corresponding equilibrium constants). The calculated ratio (of [enzyme site]/[free substrate]) ranges from very low for phosphofructokinase (0.004) to very high for glyceraldehyde 3-phosphate dehydrogenase (21). It is obvious that $[E_t] \geq [S]$ for at least a few enzymes in vivo (such as aldolase and glyceraldehyde 3-phosphate dehydrogenase) and consequently Michaelis-Menten formalism does not apply. More importantly, in such cases a considerable portion of the substrate may be bound to the enzyme(s). However, calculated free substrate concentration (total minus bound substrate) may be employed in a general Michaelis-Menten analysis to predict the actual velocity of the reaction.

Besides very high total protein concentrations (Table 38.4), the cellular concentrations of some enzymes are also high (Table 38.1). An important consequence of high enzyme active site concentration (particularly when $[E_t] \geq [S]$) in vivo is its effect on the equilibrium position itself. In such cases, one has to consider the equilibrium between enzyme-bound S and P (i.e., $K_{int} = [EP]/[ES]$, the *internal equilibrium constant*) and not the aqueous equilibrium (i.e., $K_{eq} = [P]/[S]$). Table 38.5 lists some K_{eq} values, both for reaction within the enzyme active site (for all reaction components bound) and for the same reaction in aqueous solution. Note that the K_{int} for all reactions are closer to unity despite the wide range in K_{eq} values for the same reaction; this tendency is predicted on thermodynamic grounds in the evolution of "ideal" catalysts. Kinetically "perfect" enzymes (those that have evolved higher rates of turnover) bind substrates and products such that the bound complexes have nearly equal free energies (i.e., the K_{int} are closer to unity implies $\Delta G^\circ = 0$).

We have seen earlier (refer to Chap. 15) that the K_{eq} (of an aqueous reaction chemical equilibrium) is related to corresponding enzyme kinetic constants through Haldane relationship. While this places limits on the overall kinetic behavior, certain features of the reaction catalyzed are still under the control of the enzyme and its properties. This is where the K_{eq} for a reaction within the enzyme active site (K_{int}) becomes very interesting! For example, consider the reaction catalyzed by methionine

Table 38.5 The K_{eq} for aqueous versus enzyme-bound reaction components

Enzyme reaction	K_{eq} when $[S] >> [E_t]$ (aqueous equilibrium) ($S \rightleftarrows P$)	K_{int} when $[E_t] \geq [S]$ (enzyme bound equilibrium) ($ES \rightleftarrows EP$)
Hexokinase	2000	~1.0
Phosphoglucomutase	17	0.4
Triosephosphate isomerase	22	0.6
Pyruvate kinase	0.0003	1.0–2.0
Lactate dehydrogenase	10,000	1.0–2.0
Creatine kinase	0.1	~1.0

Adapted from Srivastava & Bernhard, Curr Top Cell Regul, 28:1–68. Copyright (1986), with permission from Elsevier.

adenosyltransferase. The formation of *S*-adenosylmethionine is thought to be essentially irreversible with a forward rate 2×10^5 times faster than the reverse rate; this is despite the fact that the K_{eq} (for the unbound reactants) is near unity. Clearly, the internal equilibrium constant (i.e., $K_{int} = [EP]/[ES]$) should approximate 2×10^5 and that the enzyme has 2×10^5-fold higher affinity (approximated as inverse of K_M value) for *P* than for *S*.

38.3 Avogadro's Number Is a Very Big Number

One gram mole of a substance contains Avogadro's number (6.023×10^{23}) of molecules. These many molecules dissolved in one liter of solvent make a molar (1.0 M) solution. However, cells and subcellular compartments (organelles) have finite, small volumes. We have already noted that typical volume of a bacterial cell is around 2 μm^3 (2×10^{-15} liter) (Table 38.4); only about 70% of this is aqueous cytoplasmic volume and accessible to solutes. Presence of very few molecules can mean a significant concentration of that compound in such small volumes (Table 38.6). Assuming that free $[H^+]$ contributes to intracellular pH in *E. coli*, just about 60 protons represent a pH of 7.0; addition of another 540 protons brings this intracellular pH down to 6.0. Confined to a limited cellular volume, a single molecule per bacterium (say *lac* operator DNA) implies a concentration of nearly 1.0 nM! At an experimental level, we are now able to manipulate volumes in attolitres (10^{-18} liters) and reach detection limits down to zeptomoles (10^{-21} moles).

In general, for small confined volumes, the number concentration of a particular molecular species (this includes enzymes!) may be more informative than its molar concentration (Halling 1989). Presence of a limited number of molecules within a cell/compartment has interesting implications. Some of these are numerically illustrated below.

Number of Invertase Molecules Per Yeast Cell It is the genius of JBS Haldane that the number of saccharase (popular as 'invertase' now!) molecules per yeast cell was calculated with data available at that time and presented in the book ENZYMES in 1930. A similar calculation is presented here with recent data on invertase.

Invertase is a 60 kDa protein and constitutes 0.9% of total yeast protein.
Yeast cell with 5 μm diameter and density 1.1 weighs 7.2×10^{-11} g.

One gram of yeast consists of 0.65 g water. Considering a protein concentration of 280 mg/ml (Table 38.4), one gram of yeast contains 182 mg protein. Hence,

1 g yeast contains 182×0.9/100 mg, i.e., 1.645 mg invertase.
1 yeast cell (7.2×10^{-11} g) contains $1.645 \times 7.2 \times 10^{-11}$ mg, i.e., 1.2×10^{-14} g invertase.
1.2×10^{-14} g invertase corresponds to $1.2 \times 10^{-14} \times 6.023 \times 10^{23}/60000$ molecules of invertase = 120000 molecules

Thus, one yeast cell contains 120,000 invertase molecules. This number is about ten times lower (12,000/ cell) in uninduced yeast, whereas in filamentous fungus *Aspergillus niger*, it may be only 1% of that number (1200/cell).

Small Number Concentrations Translate Into Discrete Molar Concentrations Few enzymes may be found in very limited copies indeed. There is a single *oriC* locus in *E. coli* genome and just 10–20 copies of DNA polymerase III per cell to initiate replication. On the other hand, with its multiple origins of replication, a mammalian cell contains about 50,000 copies of this enzyme. A single molecule per bacterial cell typically corresponds to a molar concentration of 1.0 nM

Table 38.6 Molar versus number concentrations in confined cellular compartments

Compartment	Volume	Number of molecules	Concentration (approximate)
Vacuole (50 nm diameter)	6×10^{-20} liter	1 H^+ ion	10^{-4} M (pH 4.0)
Vacuole (250 nm diameter)	7.5×10^{-18} liter	1 H^+ ion	10^{-6} M (pH 6.0)
		50 H^+ ions	10^{-5} M (pH 5.0)
E. coli	2×10^{-15} liter	1	10^{-9} M (1 nM)
		20	2×10^{-8} M (20 nM)
		200	2×10^{-7} M (0.2 μM)
Mitochondrion	1×10^{-15} liter	1	10^{-9} M
Yeast	7×10^{-14} liter	15,000	10^{-7} M
Mammalian nucleus	5×10^{-13} liter	300	10^{-9} M
Mammalian cell	2×10^{-12} liter	1	10^{-12} M

As compartmental dimensions vary within a range, typical volumes are presented here. Concentrations are computed by assuming that the solute is uniformly distributed throughout the compartment. Concentration of over-expressed proteins in *E. coli* can be abnormally high and reach up to 5 mM.

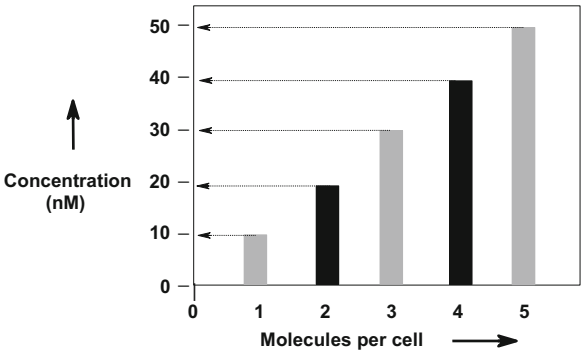

Fig. 38.1 **Molar versus number concentrations in a confined cellular volume of *E. coli***

(Table 38.6); every additional molecule contributes to a discrete increment of 1.0 nM. In such situations, a bacterial cell may experience stepwise concentration increase – in 1.0 nM multiples (and none in between!) (Fig. 38.1). And changes in cellular volume then become important in accessing the full range of molar concentrations.

Dissociation Constants, Equilibrium Binding, and Stochasticity The binding equilibrium between two molecular species depends on the equilibrium constant and the two concentrations. The binding of an enzyme to its substrate is no exception. However, the same binding phenomenon becomes stochastic when small "number concentrations" are involved. Consider a simplistic calculation of DNA polymerase III binding in vivo to *oriC* locus of *E. coli*. A single copy of *oriC* DNA (per chromosome) approximates to 1.0 nM, and ten copies of DNA polymerase III per cell would represent a concentration of 10.0 nM. If DNA polymerase III binds to *oriC* DNA with a K_D of 1.0 μM, then we can evaluate the status of this binding equilibrium as follows:

$$oriC + \text{DNA pol III} \rightleftarrows oriC\text{-DNA pol III complex}$$

$$K_D = \frac{[oriC]\,[\text{DNA pol III}]}{[oriC\text{-DNA pol III complex}]}$$

On rearranging

$$\frac{[oriC]}{[oriC\text{-DNA pol III complex}]} = \frac{[oriC]_{free}}{[oriC]_{bound}} = \frac{K_D}{[\text{DNA pol III}]}$$

Substituting the respective values

$$\frac{[oriC]_{free}}{[oriC]_{bound}} = \frac{10^{-6}\text{M}}{10^{-8}\text{M}} = 100$$

Table 38.7 DNA polymerase III occupying the *oriC* locus in *E. coli*

K_D (*oriC*-DNA pol III complex)	Number of DNA pol III molecules per cell (concentration)	$\frac{[oriC]_{free}}{[oriC]_{bound}}$	*oriC* bound with DNA pol III (probability %)
10^{-3} M (mM)	1 (10^{-9} M)	10^6	0.0009
	10 (10^{-8} M)	10^5	0.009
10^{-6} M (μM)	1 (10^{-9} M)	10^3	0.09
	10 (10^{-8} M)	10^2	0.99
10^{-9} M (nM)	1 (10^{-9} M)	1.0	50.0
	10 (10^{-8} M)	10^{-1}	90.9
10^{-12} M (pM)	1 (10^{-9} M)	10^{-3}	99.9
	10 (10^{-8} M)	10^{-4}	99.99

This implies 0.99 nM of the total of 1.0 nM *oriC* DNA is free (i.e., unbound). Since there is a single copy of *oriC* DNA, presenting the concentration of *oriC* DNA bound to DNA polymerase III as 0.01 nM (or 10^{-11} M) does not make any sense! We now enter the realm of probabilities, and the binding phenomenon is best considered as stochastic. A $[oriC]_{free}/[oriC]_{bound}$ ratio of 100 signifies that, at any given time, there is just 1% chance that DNA polymerase III is occupying the *oriC* locus. A few more representative calculations of this kind are informative (Table 38.7); the lower the K_D (higher the affinity), the greater is the probability of *oriC* DNA bound to the polymerase. At picomolar K_D value, the *oriC* DNA is fully occupied (nearly 100% probability) by DNA polymerase III. While these calculations are useful, actual DNA polymerase III binding to *oriC* locus in vivo is more complex and is supported by additional protein components. The *E. coli* DNA polymerase III (600 kDa protein) is an efficient (about 1000 nucleotides added per second) and highly processive enzyme. It remains bound to DNA template and does not fall off almost until the entire bacterial genome is replicated.

Another way of looking at the phenomenon of stochastic (probabilistic) binding is to compare the "on" and "off" rates. Note that the K_D may also be viewed as the ratio of the off and on rate constants (k_{off}/k_{on}; also see "Tight-Binding Inhibitors" in Chap. 21 Irreversible Inhibitions). The slower the rate of release ($k_{off} < k_{on}$) from the complex, the longer the protein resides on the DNA (since inverse of the first-order rate is time).

Finally, we note that stochastic (probabilistic) binding is encountered in many other cases such as when the tetrameric *lac* repressor (5–10 molecules per cell) binds (K_D, 10 pM) to the operator DNA of lactose operon. The study of single-enzyme molecules, both in vitro and in vivo, is one research frontier in enzymology; we will briefly touch upon this topic in the next chapter.

38.4 Diffusion, Crowding, and Enzyme Efficiency

The physical event of substrate colliding with the enzyme molecule is diffusion controlled. As a catalyst, an enzyme can do very little to overcome this diffusion limit. The upper bound for catalytic rate acceleration is the prevailing diffusion rate, and an enzyme is considered kinetically perfect when its k_{cat}/K_M approaches that diffusion limit (we have already noted this in Chap. 15) (Hasinoff 1984). The interior aqueous milieu of intact cells is a very crowded place. The diffusion of enzymes and their substrates is accordingly affected by the crowded and concentrated state of cytoplasm (Table 38.4) (Milo and Phillips 2016). Although the concentration of individual enzymes/proteins may not be very high, cells do contain a dense mixture of large and small molecules. This has a direct bearing on the viscosity of the cytoplasm. The *macro-viscosity* (or bulk viscosity as measured, e.g., by Brookfield viscometer) is affected by the polymeric solute molecules (cytoplasmic macromolecules and the synthetic ones like Ficoll or polyethylene glycol). This macroscopic flow property of the system does not necessarily correlate with effects on diffusion of small molecules. However, the small molecular weight solutes (like sucrose and glycerol) do influence diffusion rates of the system at the microscopic scale (the so-called micro-viscosity) (Verkman 2002). Techniques like fluorescent probe diffusion using photobleaching, correlation microscopy, and time-resolved anisotropy have provided a measure of micro-viscosity (fluid-phase viscosity) of cytoplasm. The viscosity of the cytoplasm has been estimated to be about 3–7 times that of water (Luby-Phelps 2000). From such data, the view of the cell interior has evolved from that of a viscous gel to that of a watery but crowded compartment. Cytoplasm has a low micro-viscosity and a high macro-viscosity. The mitochondrial matrix is even more crowded.

Three independent factors affecting solute diffusion in the cytoplasmic compartment are (a) specific binding to intracellular components, (b) slowed diffusion in fluid-phase cytoplasm (micro-viscosity), and (c) collision with intracellular components (macromolecular crowding). We have already noted in Table 38.1 that significant proportion of substrates (metabolites) may be present in their enzyme-bound form. The fluid-phase viscosity of cytoplasm (i.e., micro-viscosity) is not much greater than that of water. Therefore, the diffusion movement of small cytoplasmic solutes is similar to that in a dilute aqueous solution. However, macromolecular crowding leads to steric exclusion and is an important barrier to diffusion in the cytoplasm.

Macromolecular Crowding and Volume Exclusion How much of the intracellular volume is available to macromolecules depends upon the numbers, sizes, and shapes of all the molecules present in that compartment. Cytoplasm is "crowded" rather than "concentrated" because no single macromolecular species occurs at high concentration (Ellis 2001). But, taken together, the macromolecules occupy a significant fraction (typically 20–30%) of the total volume. A simple illustration of this concept is shown in Fig. 38.2. The space occupied by macromolecules is physically unavailable to other molecules, and the resultant volume exclusion has noteworthy

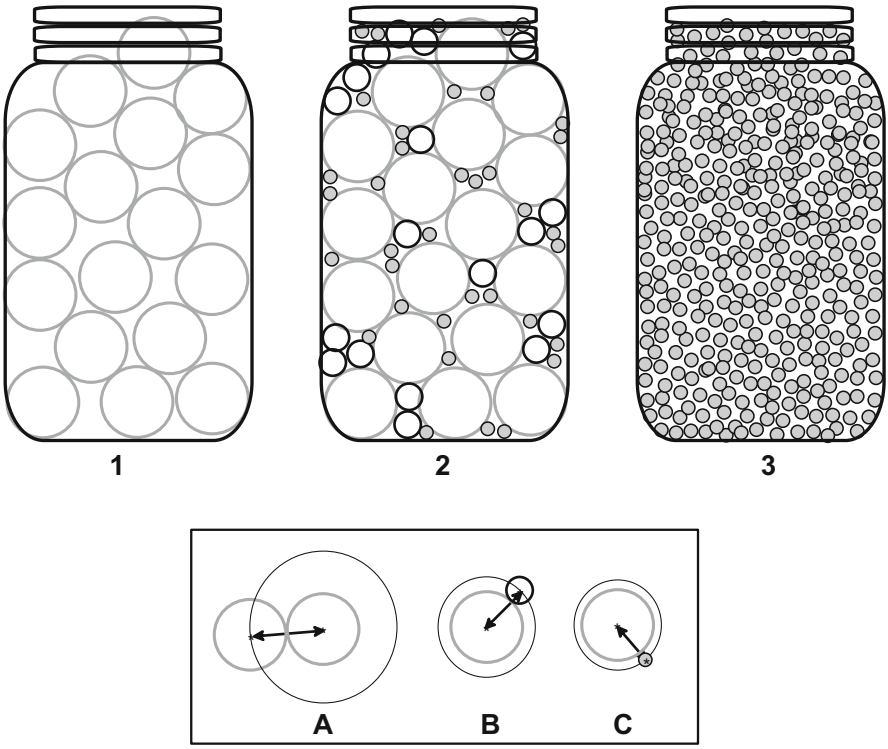

Fig. 38.2 Visualization of crowding effects. A glass jar can be filled by large- (Jar 1), small- (Jar 3), or medium-sized spheres. Jar 3 cannot further accommodate any of these spheres. However, Jar 1 can still accommodate medium−/small-sized spheres in between (Jar 2). This is because large spheres exclude more volume from around them. Similarly, a macromolecule in solution will exclude others from its neighborhood. The position of each molecule is specified completely by the position of its center. The closest any two molecules (assumed as spheres; box below) can approach is a distance equal to the sum of their radii. Around each molecule is a spherical volume from which the centers of all others are excluded. For instance, the volume around a macromolecule from which the center of another similar molecule is excluded is a sphere of radius twice (eight times the volume) that of a single molecule (case A in the box below)

consequences. Macromolecular crowding leads to (1) volume exclusion of reactants, (2) a reduction in the diffusion coefficient of macromolecules, and (3) a reduction in the degree of mixing of molecules with an increase in reactant segregation.

Molecular crowding may alter observed equilibrium constants and/or may profoundly affect the enzyme kinetic parameters. Some examples of the kinetic effects are listed below:

- Concentration of a molecule in the compact form is favored in the presence of a space-filling substance, and the effect increases exponentially with the concentration of inert molecule. If we crowd the solution, the system will change to minimize crowding – molecules will associate, thereby reducing the excluded volume. If the enzyme can associate to an oligomer (with the enzyme activity different from that of the monomer), then crowded solutions will favor the oligomer (and alter the observed enzyme activity). The tetramer of glyceraldehyde 3-phosphate dehydrogenase is less active than the monomer – when crowded by other proteins (as is the case inside the cell), its activity is reduced which is consistent with excluded volume effects.
- Protein diffusion is slowed down by molecular crowding as other macromolecules become obstacles to be avoided (compare a person moving to the exit in a train station that is crowded versus empty). Weak protein associations (with $K_D > 10^{-4}$ M) that are otherwise functionally relevant could well disappear in dilute enzyme assays with which enzyme kineticists work in vitro. Such weak protein–protein interactions may be specifically promoted by molecular crowding. Evolution seems to have conserved not only functional sites of protein molecules but also structural features that might determine the abilities of proteins to associate with one another.
- If the *ES* complex is more compact than E, then crowding will enhance the complex formation; this in turn lowers the K_M. For such enzymes, the in vivo K_M is unlikely to be simulated by kinetic measurements from routine test tube data. DNA replication in vitro requires the inclusion of high concentrations of polymer crowding agents like polyethylene glycol – this enhances the interaction between the DNA and the polymerase and other relevant proteins.
- During catalysis, the enzyme transition state may be expanded or contracted during catalysis. Therefore, in a crowded solution, the activation energy is raised or lowered, respectively; this in turn affects V_{max}. The measured V_{max} for pyruvate reduction by lactate dehydrogenase increases linearly with increasing concentrations of ovalbumin, serum albumin, or dextran – a result consistent with a decrease in the volume of the *TS* upon NADH binding.

After many failed attempts, Arthur Kornberg's group was successful in replicating the *oriC* plasmid in vitro, by including high concentrations of PEG in the incubation mixture (Kornberg 2000, 2003). As Kornberg put it "the PEG occupies most of the aqueous volume and excludes a small volume into which large molecules are crowded. This concentration is essential when several proteins are needed in the consecutive stages of a pathway." The effect of crowding on enzyme activity is reflected as one of his Ten Commandments – *thou shalt correct for extract dilution with molecular crowding*!

Enzyme Size Matters The micro-viscosity of the medium, rather than the macro-viscosity, determines the velocity of a diffusion-limited enzyme reaction. The catalytic perfection of an enzyme can be benchmarked by comparing its k_{cat}/K_M against the prevailing diffusion rate constant (Chap. 15) (Knowles & Albery 1977;

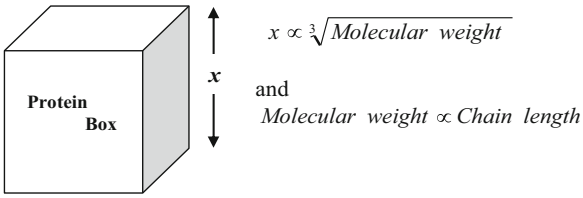

Fig. 38.3 Relation between the three-dimensional size of an enzyme, its molecular weight, and the polypeptide chain length

Burbaum et al. 1989). By this yardstick, most enzymes studied have already achieved "kinetic perfection" (see Table 15.2). Enzymes are some of the smallest phenotypic units on which evolutionary forces act. But then, is catalytic perfection the only feature of the enzyme selected by nature? This quickly brings us to other questions like: Why are enzymes so big? Can we reduce the biosynthetic cost of an enzyme further? Are there tradeoffs while choosing between these features? It appears that other features like cost, stability, and regulation also figure significantly in the evolution of enzyme structure (Benner 1989). As T. Dobzhansky (1973) puts it – *Nothing in biology makes sense except in the light of evolution*. And enzymes are no exception.

Enzyme active sites most often occupy a small percentage of their total surface area. With the exception of those acting on polymeric substrates (like polysaccharides, polypeptides, DNA or RNA), enzymes are relatively larger than their corresponding substrate(s). Is there no selection pressure to trim them to a smaller size while retaining catalytic function? In other words, are enzymes also evolved to optimize biosynthetic cost to the cell? Enzymes are three-dimensional protein boxes made of linear chain of amino acids linked to each other by peptide bonds. The dimensions of this protein box increase only with the cube root of its molecular weight (Fig. 38.3). A simple calculation will show that to achieve a 5 times increase in the length "x" of the protein box, we need about 100 times increase in the molecular weight of that protein. Taken together with the relative sizes of the protein and the corresponding substrate that occupies the active site on its surface, it seems that enzymes are not really that big after all.

Almost 70% of the cellular energy is spent toward protein synthesis – larger enzymes are more expensive. Considering such high biosynthetic cost, it is not surprising that catalytic efficiency of an enzyme is not the only feature selected for by nature. Improvements in catalytic perfection beyond a point become counterproductive – an incremental increase in efficiency may require substantial increases in biosynthetic investment. This is conceptually illustrated through an imaginary example in Table 38.8.

Larger size of a polypeptide enzyme ensures the proper positioning of the active site residues required for most effective catalysis. Therefore, catalytic efficiency could be improved by accommodating better design through an increase in material input (increase in the number of amino acid residues per catalyst; E1–E3 in Table 38.8). Beyond a point however, the gain in catalytic efficiency does not justify the large investment in the size increase. Natural selection will prefer such a cost-

Table 38.8 Catalyst size and relative biosynthetic cost

Form	Enzyme size (number of amino acid residues)	Notional catalytic efficiency [(k_{cat}/K_M)/Diffusion rate]	Number of enzyme molecules required to achieve same relative flux	Relative biosynthetic cost
E1	10	0.001	10,000	1.000
E2	100	0.100	1000	1.000
E3	1000	1.000	100	1.000
E4	**500**	**0.800**	**125**	**0.625**

efficient catalyst. Yet an increase in the number of less efficient enzyme molecules in principle could compensate for and support the required metabolic flux rate (see the imaginary example E4 whose k_{cat}/K_M is not diffusion controlled; Table 38.8). For example, over-expression of less efficient mutant enzyme forms is known to adequately complement the host defect. A case in point is L-methionine biosynthesis in *E. coli* – two different isoforms can support this pathway. The B_{12}-dependent enzyme is a larger polypeptide (130 kDa) than the corresponding non-B_{12}-dependent isoform (99 kDa). Although the B_{12}-dependent form is larger, this biosynthetic cost is offset by its 100 times superior k_{cat} value. Everything else being equal, in terms of cost, more than 75 non-B_{12}-dependent enzyme molecules are required to replace each B_{12}-dependent enzyme molecule for the same catalytic need (recall that $V_{max} = k_{cat} \times [E_t]$). The organism employs the larger isoform whenever it has access to B_{12} but otherwise resorts to many more molecules of the smaller isoform. Clearly, *an enzyme is also selected by nature for its biosynthetic cost efficiency*.

Then Why Are Enzymes Big? It is obvious from the previous paragraphs that evolutionary pressures do act to optimize enzyme size (biosynthetic cost). Nonetheless enzymes in general are perceived as large molecules compared to their substrates (Srere 1984). This is because we consider size in relation to linear dimensions rather than volume or mass. As noted earlier, however, the dimensions of a protein box increases only with the cube root of its molecular weight (Fig. 38.3). Polypeptides in the size range of 30–50 kDa make up more than 50% of the total cellular proteins; only 3–5% proteins are found above the 80 kDa range. No naturally occurring enzymes with polypeptide chains of less than 50 amino acid residue length are known. Some of the smallest enzymes known include the following: (1) 4-oxalocrotonate tautomerase consists of a 62-residue monomer but functions as a hexamer; (2) acylphosphatase consisting of a 98-residue monomer is one of the smallest enzymes known; it catalyzes the hydrolysis of acylphosphates; and (3) the HIV protease functions as a homodimer; the active site lies between the identical subunits made up of 99 amino acids. Among synthetic organic chemists, proline is considered the smallest chiral catalyst, and it catalyzes asymmetric aldol reactions.

The rate accelerations, however, are nowhere comparable to polypeptide-based enzymes. Lastly, if protein enzymes are considered big, RNA enzymes are even bigger. A short discussion on nonprotein catalysts may be found in the next chapter.

Hexokinase (~50 kDa protein) is about 70 times the combined molecular masses of glucose and Mg-ATP. Moderately sized substrates like glucose make use of as many as 15 or more hydrogen bonds for proper binding to the enzyme. These considerations bring us to question the purpose of relatively larger size of an enzyme. The following points may be considered.

- Enzyme active sites most often occupy a fraction of their total surface area. A typical substrate (like glucose for hexokinase) covers 10–15% of the total enzyme surface area and occupies 2–3% of its total volume. But active site as a rigid entity is not acceptable. They should be flexible enough to bind and release substrate or product but rigid enough to best fit the transition state. The bulk of the enzyme that does not constitute the active site is needed to maintain the active site in geometry faithful to its transition state structure. Enzymes have evolved for conformational flexibility, and this comes with a cost. Their large size ensures that the interaction of the substrate with the active site alters the global conformation of the enzyme in a meaningful way – the active site shifts from an initial substrate-specific geometry to a transition state-specific geometry. In addition, in some cases, interactions (electrostatic!) between the enzyme and its substrate beyond the active site do contribute to rate accelerations.
- Some enzymes, in addition to acting as catalysts, also serve as sensors. To accommodate such regulatory features, the protein box is anticipated to display additional sites to bind the regulatory ligands. Larger proteins have enough surface area for multiple interactions through allosteric sites (such as glutamine synthetase; Chap. 37). Cooperative interaction between different sites requires that the binding information be transmitted across space through the conformational changes in the polypeptide. For example, hemoglobin (long considered an honorary enzyme!) has distinct sites and conformational states to bind oxygen, carbon dioxide, and 2,3-bisphosphoglycerate.
- Some enzymes harbor multiple functions on them and have multiple domains. Domain structure provides for combining catalytic and regulatory properties and protein–protein interactions. Protein–protein interactions are the rule rather than the exception. Through channeling (see below), metabolic advantages accrue to the cell besides the expected kinetic advantage.

We may conclude that the requirements for catalytic perfection, accommodation of regulatory site(s), and/or conformational flexibility are not necessarily congruent properties. One or more of these features may require that others are somewhat compromised. The ultimate design of an enzyme catalyst may be the result of

tradeoffs between catalytic efficiency, protein stability, biosynthetic cost, and inclusion of regulatory features. Fortunately the goal of a *Biochemist* and that of *Natural Selection* is not congruent. This leaves enough scope for enzyme engineering and redesign.

38.5 Consecutive Reactions and Metabolite Channeling

If one focuses on adjacent reaction steps inside a cell, the study of enzymes in vivo quickly becomes the study of a metabolic pathway. In coupled enzyme assays (e.g., see Fig. 12.2; Chap. 12 Principles of enzyme assays), we can deliberately couple almost any two enzymes of our choice, provided they share a substrate–product pair. In the cellular context however, enzymes of metabolism exist in pathways, and they do not function in isolation. The product of the previous enzymatic step feeds into the next enzyme as its substrate.

Consecutive Steps in Metabolism Enzymes catalyzing the consecutive steps of a metabolic sequence provide interesting insights into how metabolism is organized and how it responds. The driver for metabolism is the desire for reactions to reach equilibrium. Nature has exploited this principle to couple reactions of metabolic pathways; reactions are made spontaneous by adjusting the concentration of reactants and products. The direction of an equilibrium reaction is decided by suitably adjusting the mass action ratio (Γ). We have noted earlier (in Chap. 10) that continuous depletion of GA3P (by GA3P dehydrogenase) maintains the ΔG negative for triosephosphate isomerase (DHAP⇄GA3P) reaction and feeds DHAP into glycolysis.

So long as an enzyme is not saturated with the substrate, an increase in [*S*] could stimulate the rate of that reaction – this relationship is typically Michaelian. As a general rule, enzymes will operate with reactant concentrations in the region of their K_M or $S_{0.5}$. This has two implications: (a) the catalytic potential of the enzyme is better utilized and (b) the system tendency to revert to steady state (and stabilize [*S*]) is facilitated. Although individual enzymatic reactions are not at equilibrium in a cell, the metabolic pathways are believed to be at, or close to, steady state. Only when supply is balanced by demand, a steady state concentration of intermediate is obtained. This means the concentrations of metabolic intermediates do not change appreciably while there is a flux (flow of $S \rightarrow P$) through the pathway.

Consider a simple two-enzyme system with two consecutive irreversible reactions:

$$A \xrightarrow{E_1} B \xrightarrow{E_2} C$$

Assuming that both E_1 and E_2 display Michaelian behavior, the velocity of the first step alone is given by

$$v_1 = \frac{V_{\max 1}\,[\mathrm{A}]}{K_\mathrm{A} + [\mathrm{A}]}$$

and of the second by

$$v_2 = \frac{V_{\max 2}\,[\mathrm{B}]}{K_\mathrm{B} + [\mathrm{B}]}$$

However, rate of change in [*B*] is given by

$$\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = (v_1 - v_2) = \frac{V_{\max 1}\,[\mathrm{A}]}{K_\mathrm{A} + [\mathrm{A}]} - \frac{V_{\max 2}\,[\mathrm{B}]}{K_\mathrm{B} + [\mathrm{B}]}$$

We can now consider three distinct cases. If $v_1 > v_2$, then *B* accumulates (increase in [*B*]). Second, if $v_2 > v_1$, then [*B*] tends to zero; at the extreme, *B* may simply be transferred from the active site of E_1 to that of E_2 without any buildup of [*B*]. We will have more to say about this phenomenon of *channeling* in the next section. The third situation is where $v_1 = v_2$ and the so-called steady-state levels of [*B*] are attained. This $[B]_{\text{steady state}}$ is given by

$$[\mathrm{B}]_{\text{steady state}} = v_1 \left(\frac{K_\mathrm{B}}{V_{\max 2} - v_1} \right)$$

The reader may look up the related treatment of a sequential two-step process (Fig. 10.4 and the dish washing analogy!) and the accumulation of intermediate presented in Chap. 10.

Coupled (natural or artificial) enzyme assays are often used in biochemical analysis. From the above equation, it follows that E_2 should ideally have a smaller K_B and a larger $V_{\max 2}$, for efficient coupling. Enzymologists use much higher activity of the second enzyme (E_2 in the above case) to achieve very little or no lag time (the time before the system enters steady state).

Substrate Channeling In the consecutive steps of a metabolic sequence ($A \rightarrow B \rightarrow C$, as above), the steady-state level of *B* is also determined by its *transit time*. This is the time required for the product B of the first enzyme (E_1) catalyzed reaction to diffuse to the active site of the next enzyme (E_2). And it depends on (a) the distance between the two sequential enzymes (E_1 and E_2) of the pathway and (b) the exact diffusion coefficient of *B* in the medium between them. A range of possibilities exist: a metabolic intermediate like *B* may completely equilibrate with its pool in the surrounding medium, only a certain fraction of it may equilibrate, or it may be directly transferred (channeled) to the next enzyme active site. The last possibility – limiting case of direct transfer of *B* between active sites without any release into the bulk phase – with the shortest transit time is called *channeling*.

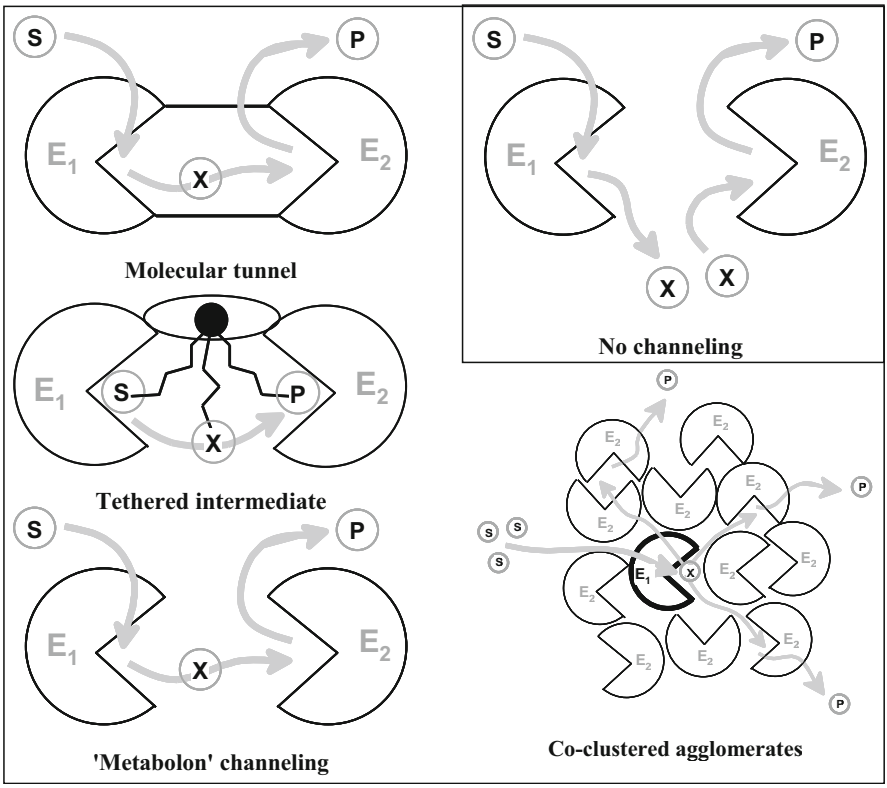

Fig. 38.4 Possible interactions involving consecutive active sites of sequential reactions. The product (X) from the first active site (E_1) may be taken to the next one (E_2) by a molecular tunnel or a covalent tether. Channeling of X may also occur from E_1 to E_2 by their proximity (forming a metabolon) or simple co-clustering. X equilibrates with the bulk metabolic pool only when there is no channeling

Enzymes exhibiting multiple activities (see Table 14.3; Chap. 14) are obvious candidates to look for this phenomenon. Metabolite channeling in vivo may be achieved through a range of sequential active site interactions (Fig. 38.4) enumerated below.

- Direct channeling relies on the formation of protein tunnels that connect consecutive active sites, preventing metabolic intermediates from diffusing away. The first *molecular tunnel* within tryptophan synthase was discovered in 1988. Indole derived from the cleavage of indole-3-glycerol phosphate at one active site traverses 25 Å through a protein tunnel to the other site where it condenses with L-serine. The steady-state concentration of indole is extremely low as very little leaves the enzyme into the surrounding medium. The translocation of

ammonia, derived from the hydrolysis of glutamine, also occurs frequently through protein tunnels (Weeks et al. 2006). Channeling ammonia has been discovered within carbamoyl phosphate synthetase, asparagine synthetase, glutamine phosphoribosylpyrophosphate amidotransferase, glutamate synthase, imidazoleglycerol phosphate synthase, and glucosamine 6-phosphate synthase. Another example is the carbon monoxide channeling demonstrated in carbon monoxide dehydrogenase/acetyl-CoA synthase. It is interesting to note that in all these examples, *the intermediate is not covalently bound to the enzyme but is simply shepherded to another active site*. The presence of molecular tunnels is becoming a recurring theme in structural enzymology (we will have more to say about this in the next chapter).

- Yet another strategy to ensure that metabolic intermediates are directed from one active site to the next is by tethering (Fischbach and Walsh 2006; Perham 2000). Both multifunctional polypeptides and multienzyme complexes belong to this category. The *tethered intermediates* are held on a chemical leash, for example, through lipoate, biotin, pantothenate, etc. In this assembly-line strategy, *swinging arms carry the reactant from one site to the other*. Examples include all α-keto acid dehydrogenase complexes, both type I and type II fatty acid synthases, polyketide synthases, and nonribosomal peptide synthases.
- Most high-affinity protein–protein complexes (typical $K_D < 10^{-6}$ M) are readily detectable. But when their interaction is weak or very weak (e.g., $K_D > 10^{-4}$ M), many conventional approaches fail to detect them. At high enzyme concentrations (and with crowded environment!) in vivo, ultra-weak interactions also become important, despite their transient nature and low stability. Enzyme–enzyme interactions can become more likely, and the preference for direct metabolite transfer becomes accordingly far more favorable. Paul Srere coined the term *metabolon* to denote a complex of sequential enzymes, which may involve loosely or transiently associated proteins. Such a metabolon may have the ability to channel a metabolic pathway, and it involves the preferential transfer of an intermediate from one enzyme to a physically adjacent enzyme, with restricted diffusion into the surrounding milieu. Carbamyl phosphate – dedicated for pyrimidine biosynthesis versus arginine biosynthesis – in *N. crassa* is an example. More recently, simple enzyme clustering was shown to accelerate the processing of intermediates through proximity channeling. Co-clustering multiple enzymes into compact agglomerates yields the same efficiency benefits as direct channeling (Castellana et al. 2014; Banani et al. 2017). However, simply fusing two enzymes together will not cause productive channeling (Sanyal et al. 2015).

Channeling is an example of first-level metabolic organization, indicating that cell is not a simple bag of enzymes (Mathews 1993). Regardless of how it is achieved, channeling does provide certain advantages to the cell. Namely, it could (1) serve to protect toxic, unstable, or scarce metabolites by maintaining them in the protein-bound state, (2) provide a metabolic advantage by maintaining concentration gradients, (3) protect the solvation capacity of cell water and reduce solute burden, (4) provide kinetic advantages in terms of rate accelerations beyond bulk diffusion

rate limitation, (5) facilitate a quick response of the pathway to inhibitors and activators, and (6) provide a regulatory feature through the dynamic formation/destruction of the metabolon complex. Channeling in a limiting case (through molecular tunnels) leads to one-dimensional diffusion of reactants. This reduction of dimensionality increases the speed and economy of diffusion-controlled reactions (as expected intuitively!) (Hardt 1979).

Channeling is an attractive concept that makes much in vivo metabolic sense. But to provide experimental proof of channeling is a challenge. Channeling is suspected if an endogenous intermediate produced in a pathway fails to mix (either partially or completely) with the same intermediate produced exogenously by an enzyme located elsewhere in the cell. Experimentally, channeling can be observed by providing a radiolabeled precursor to the pathway and monitoring either the product or an intermediate; label mixing with pools of nonradioactive intermediates (and subsequent dilution) leads to diminished specific radioactivity. Channeling implies facilitated transfer of channeled intermediates. Conversely, restricted access by exogenous intermediates to the pathway is expected. Confirmation of metabolite channeling demands multiple experimental approaches, with each approach failing to disprove it (rather than proving it!). Besides analysis of its advantageous kinetic features (like reduction in transient times, enhanced in vivo reaction flux rates, demonstration of direct transfer of intermediates, metabolite compartmentation in the absence of organelles, etc.), several approaches could be used to establish the physical proximity of enzymes in a metabolon. These tools include co-fractionation/co-localization of enzyme activities, use of bifunctional cross-linking reagents, and immunoprecipitation and pull-down assays. Partial or complete reconstitution of a functional complex from purified protein components also offers reasonable support for metabolite channeling.

Metabolic Branch Points – Enzymes Competing for a Metabolite When different substrates compete for the active site of the same enzyme, it can be used to glean useful kinetic insight to enzyme mechanism (see Chap. 23 "Alternate Substrate Interactions"). Then again, two different enzymes may compete for the same substrate. For example, L-arginine is a substrate for both nitric oxide synthase and arginase in mammalian cells. There are effectively two kinds of enzyme–substrate competitions in vivo. A single metabolite may be substrate for more than one enzyme or one enzyme may accept different metabolites as its substrate.

Many enzymes vying for the same substrate (metabolite) are found at metabolic branch points. Consider the oxaloacetate node. Intracellular oxaloacetate concentration is very low, and much of it is enzyme bound. The mitochondrial oxaloacetate is available to citrate synthase, malate dehydrogenase, phosphoenolpyruvate carboxykinase, and aspartate aminotransferase. However, distinct cellular metabolic states decide the flow of oxaloacetate either into Krebs cycle (citrate synthase) or toward gluconeogenesis (phosphoenolpyruvate carboxykinase). A host of intensive and extensive properties influence the overall rate of an enzyme-catalyzed reaction in cellular metabolism (Fig. 37.1; Chap. 37 Regulation of enzyme activity). Of these,

relative affinities for a common substrate and concentrations of competing enzymes (V_{max} values) significantly decide the fate of a branch point metabolite. The enzyme with a lower K_M will win this competition and operates closer to its full capacity. Such an enzyme binds more substrate because of its higher affinity; also, enzyme present at higher concentration hogs larger share of the common substrate. This in turn generates a steeper substrate gradient near its vicinity. In general, relative K_M values of the competing enzymes dictate the fate of that branch point metabolite. For instance, phosphofructokinase and glucose-1-phosphate uridyltransferase drain glucose-6-phosphate as their respective hexose phosphate substrates – fructose-6-phosphate and glucose-1-phosphate. But with a much lower K_M for its substrate, phosphofructokinase has a higher preference for hexose phosphate; significant glycogen synthesis begins only after phosphofructokinase is fully saturated. The following examples will illustrate this point further. Whereas many more equally interesting cases do exist, the three examples presented below carry a personal bias:

1. Fungal arginine metabolism: Filamentous fungi are capable of biosynthesizing arginine starting from glutamate and via ornithine. They can also utilize arginine supplied from outside. Most of the cellular arginine (and ornithine) is located in the vacuoles, and the cytosolic arginine concentrations are much lower. Arginase, vacuolar arginine transporter, and arginyl-tRNA synthetase compete for this arginine pool (Fig. 38.5, panel A). At most cellular concentrations of arginine, the vacuolar transporter would be more saturated than arginase. This ensures that significant arginine catabolism ensues only after the vacuolar reserves are filled up. Low cytosolic concentration of arginine enables arginyl-tRNA synthetase to compete successfully with arginase; on the basis of their relative K_M values, arginyl-tRNA synthetase will be operating at >90% maximum while arginase is hardly active. When the pool of arginine in the cytosol increases rapidly (due to externally supplied arginine), arginase comes into action.
2. Metabolic fate of 2-oxoglutarate: Cellular 2-oxoglutarate occupies a branch point that connects carbon and nitrogen metabolism. It is either oxidized to succinyl-CoA (by 2-oxoglutarate dehydrogenase complex and taken through further steps of Krebs cycle) or gets reductively aminated to glutamate (by NADP-glutamate dehydrogenase and leads to biogenesis of glutamate family amino acids). Respective 2-oxoglutarate K_M values for the two enzymes dictate how it partitions between the two routes – cellular energy needs versus biomass (Fig. 38.5, panel B). The 2-oxoglutarate dehydrogenase complex is effectively fully saturated at high [2-oxoglutarate], and only then is significant glutamate synthesis expected. In another example, this split ratio is an important determinant in the glutamate fermentation by *C. glutamicum*.
3. Pyruvate branch point: Pyruvate largely originates from phosphoenolpyruvate in glycolysis. The flux away from pyruvate is defined by four enzyme activities in *A. niger*. Except malic enzyme, others (pyruvate carboxylase, alanine aminotransferase and pyruvate dehydrogenase complex) have sub-millimolar affinity for pyruvate (Fig. 38.5, panel C). Scrutiny of their relative K_M values and concentrations of competing enzymes is a prerequisite to attempt diverting

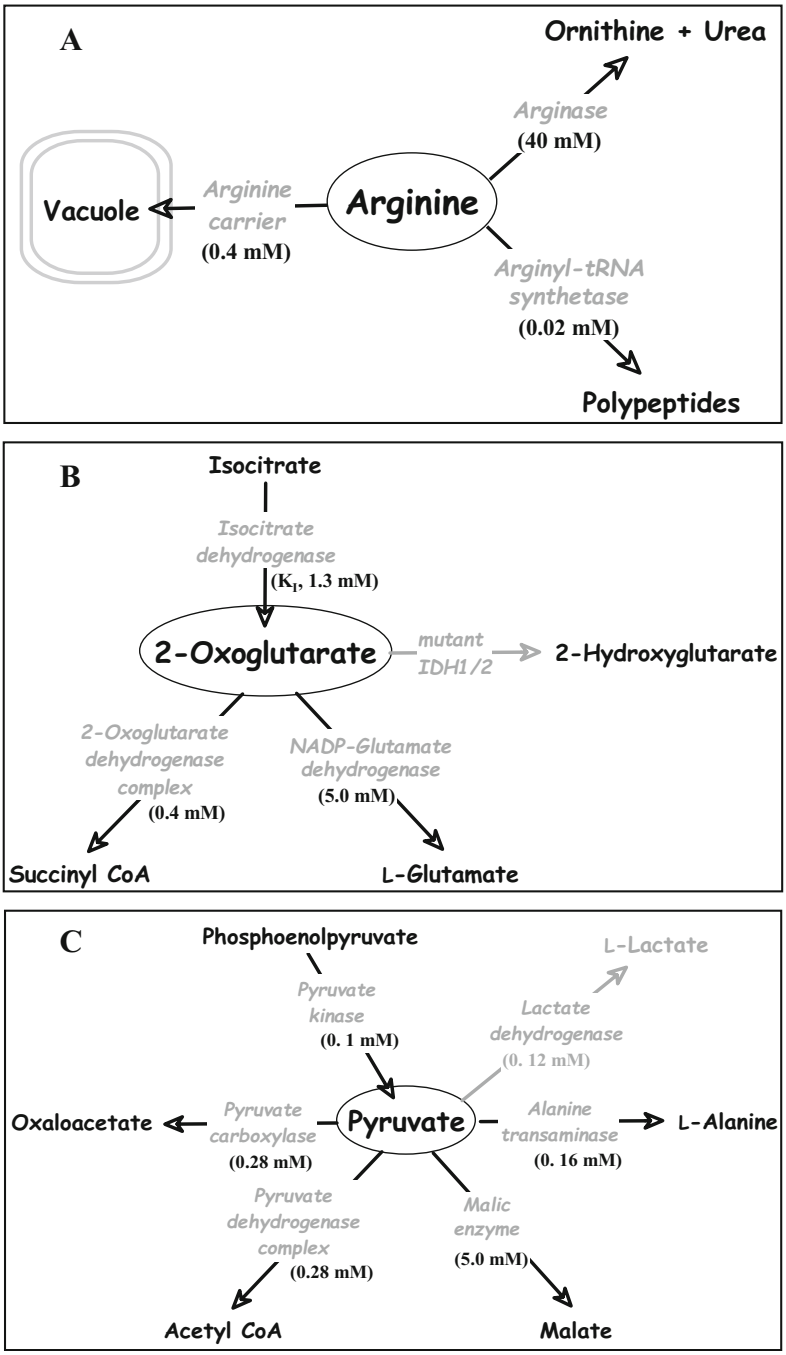

Fig. 38.5 Enzymes compete in vivo for a common substrate at the metabolic branch point. Competing reactions at L-arginine (**A**), 2-oxoglutarate (**B**), and pyruvate (**C**) nodes are illustrated. The K_M values shown for the respective enzymes (in gray) are from fungal literature (mostly from aspergilli)

pyruvate flux to lactate. This fungus does not produce lactate. Introducing a lactate dehydrogenase with appropriate kinetic features could in principle facilitate lactate formation (Dave and Punekar 2015).

Underground Metabolism Most enzymes recognize certain substrate analogs and are able to use them as alternative substrates. In some cases these analogs are endogenous, natural metabolites. Under normal conditions they are not accessible or acted upon by such enzymes. Such reactions catalyzed by normal enzymes acting on substrate analogs which themselves are endogenous metabolites constitute *underground metabolism*. For example, the biosynthetic pathways of arginine and proline involve analogous reactions on different substrates. *N. crassa* strains with deletions of the first two enzymes of proline synthesis (γ-glutamyl kinase and γ-glutamyl phosphate dehydrogenase) are viable because the deacetylase, which normally deacylates N-acetylornithine, promiscuously deacetylates N-acetylglutamate semialdehyde. Thus, early part of arginine biosynthesis gets short-circuited to generate proline using arginine biosynthetic enzymes.

Yet another example of underground metabolism is the reduction of 2-oxoglutarate to 2-hydroxyglutarate. 3-Phosphoglycerate dehydrogenase has an interesting dual function. The *S. cerevisiae* enzyme functions in the anabolic pathway of serine synthesis and may also reduce 2-oxoglutarate. The latter activity is manifest under physiological states when excess 2-oxoglutarate and reducing power coexist in vivo. Consistent with this idea, anaerobic growth of *E. coli* is also inhibited by external supply of 2-oxoglutarate, and this inhibition is reversed by serine addition. Accumulation of 2-hydroxyglutarate (considered an *oncometabolite*) can modulate the activities of 2-oxoglutarate-utilizing dioxygenases. Wild-type isocitrate dehydrogenase (IDH1 and IDH2 isoforms) catalyzes the NADP-dependent reversible conversion of isocitrate to 2-oxoglutarate. But cancer-associated gain-of-function mutations (at substrate-binding residues namely, Arg 132 in IDH1 and Arg 140 and Arg 172 in IDH2) enable mutant IDH1/2 to catalyze the NADPH-dependent reduction of 2-oxoglutarate to *R*(-)-2-hydroxyglutarate (D-2-hydroxyglutarate) (Fig. 38.5, panel B) (Dang et al. 2009).

Broader substrate specificity (enzyme promiscuity and moonlighting activities of some enzyme proteins; Chap. 14) of some enzymes manifests itself as underground metabolism. And it has been well argued that underground metabolism is a testing ground for evolution of metabolic pathways.

38.6 Summing Up

A tiger in the cage is not the same as a tiger in the jungle! Enzyme study in vivo is much different from that in vitro due to subcellular organization and compartmentation. The aqueous phase of the cytoplasm is crowded and often has high enzyme concentrations. The Michaelis-Menten formalism is not suitable in vivo for many such enzymes (where $[S_t] > [E_t]$), is not true). Further, the cell provides a single

consensus medium common for all its enzymes functioning in that compartment, whereas an enzymologist's test tube is optimized for the enzyme of his choice. Many interactions and regulatory features are simply not there in such a clean system. Rate laws determined in vitro with purified dilute homogeneous enzyme solutions may not reflect the enzyme–enzyme interactions that are important in vivo. Molar concentrations of many enzymes in vivo are quite high because of which a considerable portion of the substrate may be enzyme bound. For small confined volumes (as we see in cellular compartments such as lysosomes, mitochondria, and peroxisomes), few molecules of enzyme or substrate may mean a significantly high molar concentration. Cytoplasm has a low micro-viscosity and a high macro-viscosity because of macromolecular crowding. Molecular crowding in turn may alter observed equilibrium constants and/or may profoundly affect the enzyme kinetic parameters. The ultimate design of an enzyme catalyst in vivo may therefore be the result of tradeoffs between catalytic efficiency, protein stability, biosynthetic cost, and inclusion of regulatory features.

Erwin Chargaff held that *But the cell is certainly more than a chemical slum*. What is this more? The study of enzymes in vivo quickly becomes the study of a metabolic pathway. The product of a previous enzymatic step feeds into the next enzyme as its substrate. This may occur with or without the intermediates freely equilibrating with the cellular metabolite pool. Metabolite channeling may occur through a range of sequential active site interactions. Many enzymes may compete for a single metabolite (at the branch point) or a single enzyme may accept more than one metabolite as its substrate. Enzymology in vivo thus merges into the complexity of cellular metabolism. A study in *systems biology* will not be complete without incorporating enzymes and their various properties.

References

Albe KR, Butler MH, Wright BE (1990) Cellular concentrations of enzymes and their substrates. J Theor Biol 143:163–195

Banani SF, Lee HO, Hyman AA, Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Molec Cell Biol 18:285–298

Benner SA (1989) Enzyme kinetics and molecular evolution. Chem Rev 89:789–806

Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. Nat Chem Biol 5:593–599

Burbaum JJ, Raines RT, Albery WJ, Knowles JR (1989) Evolutionary optimization of the catalytic effectiveness of an enzyme. Biochemistry 28:9293–9305

Castellana M, Wilson MZ, Xu Y, Joshi P, Cristea IM, Rabinowitz JD, Gitai Z, Wingreen NS (2014) Enzyme clustering accelerates processing of intermediates through metabolic channeling. Nat Biotechnol 32:1011–1018

Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, Fantin VR, Jang HG, Jin S, Keenan MC, Marks KM, Prins RM, Ward PS, Yen KE, Liau LM, Rabinowitz JD, Cantley LC, Thompson CB, Vander Heiden MG, Su SM (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature 462:739–744

Dave K, Punekar NS (2015) Expression of lactate dehydrogenase in *Aspergillus niger* for L-lactic acid production. PLoS One 10(12):e0145459

Dobzhansky T (1973) Nothing in biology makes sense except in the light of evolution. Am Biol Teach 35:125–129
Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597–604
Fischbach MA, Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem Rev 106:3468–3496
Halling PJ (1989) Do the laws of chemistry apply to living cells? Trends Biochem Sci 14:313–353
Hardt SL (1979) Rates of diffusion controlled reactions in one, two and three dimensions. Biophys Chem 10:239–243
Hasinoff BB (1984) Kinetics of carbonic anhydrase catalysis in solvents of increased viscosity: a partially diffusion controlled reaction. Arch Biochem Biophys 233:676–681
Knowles JR, Albery WJ (1977) Perfection in enzyme catalysis: the energetics of trisoephosphate isomerase. Acc Chem Res 10:105–111
Kornberg A (2000) Ten commandments: lessons from the enzymology of DNA replication. J Bacteriol 182:3613–3618
Kornberg A (2003) Ten Commandments of enzymology, amended. Trends Biochem Sci 28:515–517
Luby-Phelps K (2000) Cryo-architecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int Rev Cytol 192:189–221
Mathews CK (1993) The cell – a bag of enzymes or a network of channels? J Bacteriol 175:6377–6381
Milo R, Phillips R (2016) Cell biology by the numbers. Garland Science, Taylor & Francis Group, LLC, New York
Perham RN (2000) Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. Annu Rev Biochem 69:961–1004
Sanyal N, Arentson BW, Luo M, Tanner JJ, Becker DF (2015) First evidence for substrate channeling between proline catabolic enzymes. J Biol Chem 290:2225–2234
Savageau MA (1992) Critique of the enzymologist's test tube. In: Fundamentals of medical cell biology, vol 3A, chemistry of the living cell. Elsevier, Amsterdam
Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E, Swainston N, Dada JO, Khan F, Pir P, Simeonidis E, Spasic' I, Wishart J, Weichart D, Hayes NW, Jameson D, Broomhead DS, Oliver SG, Gaskell SJ, McCarthy JEG, Paton NW, Westerhoff HV, Kell DB, Mendes P (2013) A model of yeast glycolysis based on a consistent kinetic characterization of all its enzymes. FEBS Lett 587:2832–2841
Srere PA (1967) Enzyme concentrations in tissues. Science 158:936–937
Srere PA (1970) Enzyme concentrations in tissue II. Biochem Med 4:43–46
Srere PA (1984) Why are enzymes so big? Trends Biochem Sci 9:387–390
Srivastava DK, Bernhard SA (1986) Enzyme-enzyme interactions and the regulation of metabolic reaction pathways. Curr Top Cell Regul 28:1–68
Storey KB (2005) Functional metabolism: regulation and adaptation. Wiley, Hoboken (First three chapters, in particular)
van Eunen K, Bouwman J, Daran-Lapujade P, Postmus J, Canelas AB, Mensonides FIC, Orij R, Tuzun I, van den Brink J, Smits GJ, van Gulik WM, Brul S, Heijnen JJ, de Winde JH, Teixeira de Mattos MJ, Kettner C, Nielsen J, Westerhoff HV, Bakker BM (2010) Measuring enzyme activities under standardized in vivo-like conditions for systems biology. FEBS J 277:749–760
Verkman AS (2002) Solute and macromolecule diffusion in cellular aqueous compartments. Trends Biochem Sci 27:27–33
Weeks A, Lund L, Raushel FM (2006) Tunneling of intermediates in enzyme-catalyzed reactions. Curr Opin Chem Biol 10:465–472