# Future of Enzymology: An Appraisal

# 39

*If we wish to catch up with nature we shall need to use the same methods as she does, and I can foresee a time in which physiological chemistry will not only make greater use of natural enzymes but will actually resort to creating synthetic ones.*

*–Emil Fischer, 1902 Nobel Lecture*

There can be no doubt that study of enzymes will continue to occupy the prime position in modern biology (and chemistry!) (Editorial, Closing in on catalysis 2009). This much is amply obvious from each and every context and examples that we have come across in the preceding chapters. The study of enzymes in isolation, most often in purified form, has occupied much of the time in this field. However, as we have seen in the last chapter, the importance of understanding enzyme function in vivo is very much appreciated now; this will form one of the frontiers in enzymology. Present emphasis on systems biology is a pointer in this direction. Enzymes in sequence, in combination with other enzymes and other cellular components, bring in interesting features often not manifested by an enzyme in isolation – coupled reactions, regulatory networks, and distributed control of metabolism are some of them.

Major metabolic pathways and the enzymes that function in them are well documented. But many more novel reactions and corresponding enzyme catalysts are being discovered on an almost regular basis. For instance, let us consider a few recent representative cases:

1. Most bacteria and all archaea synthesize glutaminyl-tRNA indirectly. The glutamate charged onto $tRNA^{Gln}$ is converted to glutamine in the second step by a tRNA-dependent amidotransferase. The two enzymes form the "glutamine transamidosome" that also involves channeling of ammonia (Ito and Yokoyama 2010).



N. S. Punekar, *ENZYMES: Catalysis, Kinetics and Mechanisms*,
https://doi.org/10.1007/978-981-13-0785-0_39

2. The dehydroalanine moieties of the lantibiotic nisin are formed from the Ser residues of the peptide by the action of lantibiotic dehydratase. This enzyme glutamylates Ser side chains using glutamyl-$tRNA^{Glu}$ for activation; subsequent glutamate elimination results in the dehydroalanine formation (Ortega et al. 2015).
3. The biosynthesis of lincomycin A (a sulfur-containing lincosamide antibiotic) recruits two bacterial thiols (mycothiol and ergothioneine) where mycothiol acts as the sulfur donor after thiol exchange (Zhao et al. 2015).
4. Riboflavin (vitamin B2) is a well-known redox cofactor in a wide variety of flavoproteins. Addition of a fourth ring to its existing three-ring system generates a riboflavin derivative (Clarke and Allan 2015). This previously unknown cofactor catalyzes new types of chemistry and is crucial for the decarboxylation of an intermediate in coenzyme Q biosynthesis.
5. A SAM-dependent enzyme-catalyzed pericyclic transformation leads to the formation of the natural product leporin. Such novel roles for SAM (*S*-adenosyl-L-methionine) are likely to be found in other examples of enzyme catalysis (Ohashi, et al. 2017).

These are just a few examples that represent author's personal bias. But certainly, many more novel reactions and enzymes will continue to be discovered and reported.

We will now consider aspects of enzymology where rapid progress is being made. Understanding the rate enhancements of enzymes continues to be a fundamental challenge of mechanistic enzymology (Herschlag and Natarajan 2013). These and other topics are anticipated to attract much attention of enzymologists in the foreseeable future.

## 39.1 Transition–State Analysis and Computational Enzymology

Conventional view of enzyme catalysis treats the transition state (*TS*) in thermodynamic terms. Equilibrium between reactants and the *TS* is assumed here; tight binding to the enzyme active site would sequester the *TS* from solution and increase the reaction rate. Kinetic isotope effects can provide direct information on the enzymatic *TS* (Chap. 27 Isotope effects in enzymology). When the intrinsic kinetic isotope effects (KIEs) are available for an enzyme reaction, then the *TS* structure can be deduced in the usual physical organic chemistry sense. Interpretations of KIEs give detailed bond order and geometric features of the transition state for an enzyme. In turn, molecular electrostatic potential surfaces of these transition-state depictions guide chemical synthesis of transition-state analogs, yielding excellent, high-affinity inhibitors (Schramm 2013). Complexes of such inhibitors with their corresponding enzymes provide structural models for computational analysis of enzymatic transition states. Taken together, KIEs and computational enzymology provide a conceptually complete picture of the *TS* in an enzymatic reaction.

There is a progressive appreciation that the enzymatic transition state is a state of maximum free energy. And this may not be simply captured as an equilibrium between the Michaelis complex, transition state, and products. Dynamic

contributions of protein motion must be incorporated to provide a full understanding of enzyme catalysis. Besides applications of kinetic measurements, isotope effects, time-resolved spectroscopy, NMR, and X-ray crystallography, computational enzymology is making significant contributions in understanding the enzyme *TS*. Most experimental techniques have focused on conformational changes that occur in microsecond to millisecond timescales; these times are often correlated with catalysis. These timescales are much slower than the *TS* lifetimes that occur on a bond-vibrational timescale. Chemical bond-forming/bond-breaking steps of interest (in enzyme catalysis) have vibrational modes in the low femtosecond timescales (Fig. 11.3). They are valuable in understanding *TS* but otherwise are experimentally difficult to access. This is where developments in computational enzymology are making some headway. Bond vibrations during enzyme catalysis can be simulated with some accuracy in the femtosecond to the nanosecond timescale. This provides computational dynamic access to the timescales of the *TS* lifetimes. However, commonly encountered enzymatic catalysis timescales ($10^{-2}$ s) require enormous computational time; and this is far beyond continuous dynamic calculations possible at present. This frontier in computational enzymology and of complex enzyme models was recognized with chemistry Nobel Prize in 2013 (Warshel 2014). The emergence of the quantum mechanical/molecular mechanics (QM/MM) approach allows one to ask what the origins of the catalytic power of enzymes actually are. The enormous increase in computer power makes it virtually certain that computer simulations will increasingly contribute in modeling molecular enzyme catalysis (Garcia-Viloca et al. 2004). This approach also promises to capture the contributions of subtle protein molecular dynamics to enzyme rate accelerations.

## 39.2 Single-Molecule Enzymology

The Michaelis-Menten equation (refer Chap. 15) is a highly satisfactory description of kinetic data involving a very large number of enzyme molecules in the assay. However, all enzyme molecules are not synchronized with each other in an ensemble-averaged kinetic measurement. The extraction of dynamic information from such an asynchronous assembly is complicated and difficult. A single enzyme molecule gives kinetic signals that reflect the dynamic states of that individual catalytic entity. This dynamic information is lost in the average signals of the ensembles. Reactions involving single enzyme molecules can now be examined by the advances in fluorescence and other time-resolved spectroscopic techniques (Smiley and Hammes 2006). These techniques combined with computational approaches allow real-time access to study the dynamic behaviors of individual molecules.

Would the ensemble-averaged enzyme kinetics also hold for a single enzyme molecule? This interesting question has attracted serious attention in the recent times for the following reasons. Single-molecule behavior is a particularly powerful way of uncovering (a) mechanistic pathways and intermediates, (b) how enzyme conformational fluctuations affect catalytic activity, and (c) heterogeneities hidden in the ensemble average. The turnover events of a single enzyme molecule are intrinsically

stochastic. We find the sequence of the time intervals between consecutive turnover events (the waiting times, denoted as $\tau$), as a function of the turnover index number. The average waiting time (denoted by $\langle\tau\rangle$) of a single enzyme molecule plotted against $1/[S]$ mimics the linear Lineweaver–Burk plot (refer Chap. 17) recorded in an ensemble measurement. The rate equation for single enzyme molecule kinetics is shown below.

$$\frac{1}{\langle\tau\rangle} = \frac{\chi^2[S]}{C_M + [S]} \quad \left(\text{compare this with } v = \frac{k_{cat}[E_t][S]}{K_M + [S]}\right)$$

The reciprocal mean waiting time determined for a single enzyme molecule is thus related to enzyme-catalyzed velocity in an ensemble measurement ($1/\langle\tau\rangle = v/[E_t]$). While the Michaelis–Menten relation continues to hold for the single enzyme molecule kinetics, interpretations of kinetic constants are different. Specifically, the $k_{cat}$ (formally represented as $\chi^2$ for single enzyme study) measured at saturating $[S]$ is a weighted harmonic mean of the different catalytic turnover rate constants represented in the single enzyme over time. Similarly, $K_M$ ($= (k_{cat} + k_{-1})/k_1$ or $C_M$ for a single molecule) also acquires an ensemble-averaged meaning (Min et al. 2005).

We notice that single-molecule and ensemble-averaged Michaelis–Menten kinetics can be reconciled (Walter 2006; English et al. 2006). An interesting insight from single-molecule study is that the waiting times measured at high $[S]$ levels follow an asymmetric probability distribution. The long time-span of the catalytic turnover rate constant (i.e., $k_{cat}$) indicates that the single enzyme molecule's catalytic velocity fluctuates over a broad range (from $10^{-3}$ to 10 s) of timescales. These catalytic fluctuations point to conformational isomers of an individual enzyme slowly interconverting over time. Conformational heterogeneity has been experimentally observed in the studies of cholesterol oxidase, staphylococcal nuclease, and a few other enzymes. A single enzyme molecule displays inevitable, stochastic fluctuations in its catalytic activity. The effects of such fluctuations would be less significant for a system comprising of large number of enzyme molecules. However, many processes inside cells rely on the activity of a single enzyme molecule, such as in DNA replication, transcription, translation, and protein transport along the cytoskeleton. Stochastic fluctuations due to low copy number of enzymes have important physiological implications for cells/organelles. These are now being probed on a single-molecule basis in vivo.

## 39.3 Structure-Function Dissection of Enzyme Catalysis

As discussed earlier in this book (Chap. 8), both conformational flexibility and protein motion are very important for enzyme catalysis. A structure–function approach to understand enzyme action uses a combination of kinetic (including rapid, transient kinetics) and structural techniques. The ultimate goal of such a study is to develop an in-depth mechanistic understanding of enzyme function. The field has traditionally made use of protein chemical modifications, spectroscopic tools (like fluorescence, circular dichroism, etc.), and more recently molecular biology

tools (like generation of site-directed mutants, truncated or chimeric enzyme proteins). Whereas solution NMR has provided some information on enzyme molecular dynamics (although at timescales much slower than most events in enzyme catalysis), as of now this structural tool is limited to small-sized proteins. Mass spectrometry of larger polypeptides and oligomeric proteins is also coming of age. Much of direct structural information on enzymes has come from X-ray crystallography – mostly presenting to us a static picture of the enzyme. As aptly stated by the late Jeremy Knowles (biochemist and professor at Harvard University) – "studying the photograph of a racehorse cannot tell you how fast it can run"; and thus there is a limit to what a snapshot protein structure can reveal. Although snapshots of enzyme bound to substrate, product, or transition-state analogs are valuable, they do not capture structural dynamics of catalytic action. Nevertheless, presence of molecular tunnels in enzymes like tryptophan synthase would not have been apparent without detailed structural inputs.

Overwhelming biochemical, mechanistic, and mutational data for any enzyme study must be supplemented by structural approaches. More recent developments in computational structure prediction (see below) provide complementary inputs. Structure-based drug discovery has also benefited from these approaches. One remarkable success in this venture is the development of HIV protease inhibitors (Wlodawar and Ericson 1993). These inhibitors were rationally designed from the knowledge of the structure and mode of action of aspartyl proteases. The discovery of saquinavir, the first protease inhibitor, made use of the promising transition-state mimic chemistry.

A fine balance between structural rigidity and conformational plasticity results in the unique catalytic power of enzymes. And structural enzymology also aims to address catalytic motions in detail (Ramanathan and Agarwal 2011). How enzymes achieve a catalytically competent state has become approachable only recently through experiments and computation.

**Site-Directed Mutagenesis and Crystal Structures** Application of molecular biology tools to probe enzyme function has matured over the last couple of decades. The reader may refer to standard textbooks of molecular biology and many protocol/recipe books on how to construct site-directed mutants and to genetically engineer enzymes. The relevant cDNA and an expression system to obtain the mutant enzyme are all that are required (Fig. 3.5). The native and various mutant enzyme forms are then subjected to rigorous structural analysis through X-ray crystallography, circular dichroism spectra, and other tools. Excellent insights continue to be gathered on residues critical for catalysis, binding, and structural stability/flexibility for many enzymes.

Site-directed mutagenesis offers a powerful approach to rationally modify an enzyme (Wagner and Benkovic 1990). It enables enzymologists to selectively replace active site residue (or any others) and ask some very interesting mechanistic questions. Yet site-directed mutagenesis does not fully account for enzymatic catalysis, because the effects of individual substitutions on catalysis are neither additive nor independent. At the resolution of amino acid residue level, one can

check whether a given residue is relevant for binding and/or catalysis. This may not give unambiguous answers as, for many enzymes, substrate binding and catalytic residues often overlap; and they may not be clearly demarcated in the enzyme active site. In fact, few examples show that specificity can reside beyond the amino acids directly interacting with the substrate – requiring major structural changes with loop grafting, etc. Nature presents us with 20 naturally occurring amino acids in proteins; an opportunity exists to replace a given residue by any one of the other 19. But still the choice is limited because very few of these substitutions are more conservative than others. For instance, Val, Ile, and Leu are often interchangeably accepted; also, a Glu residue may replace Asp and *vice versa*. Alanine scanning mutagenesis is a mature tool and is often used to determine the contribution of a specific residue to the stability and/or function of a given protein. Alanine is chosen because its R group (methyl) is least disruptive and imitates the secondary structure preferences of many other amino acids. Such Ala replacements are usually done by site-directed mutagenesis or generated randomly by creating a PCR library.

Standard site-directed mutagenesis is largely limited to 20 natural, proteinogenic amino acid residues. Nonnatural side chains often provide mechanistic insights. One approach has been to replace the relevant residue by Cys (through SDM) and then alkylate that Cys-SH by a suitable reagent. Interesting structural variations of imidazole side chains, with subtle pKa changes, could then be tested (Earnhardt et al. 1999). In another example, the catalytic activity of glutamine synthetase R→C mutant was rescued back by chemically modifying the Cys-SH back to an arginine analog by covalent modification with 2-chloroacetamidine (Dhalla et al. 1994). More recently, sophisticated tools and technology are in place to directly incorporate nonprotein amino acids into proteins in a position-specific manner. Nonnatural amino acids may be introduced into proteins by manipulating in vitro protein translation as well as through in vivo strategies by expanding the genetic code (Hendrickson et al. 2004). Pyrrolysine, the twenty-second protein amino acid, was found at the active site of methyltransferases from methane-producing archaea. Like the 20 common amino acids, pyrrolysine is synthesized in the cytoplasm and incorporated at a specific position during the translation of the growing polypeptide chain (Atkins and Gesteland 2002; Ragsdale 2011). The pyrrolysine biosynthetic cassette could be used to incorporate other useful modified amino-acid residues into proteins.

Examples of SDM to probe enzyme function have rapidly grown. We will look at only a few case studies to illustrate main issues and highlight some difficulties with this approach. The first significant effort to change the substrate specificity of an enzyme was with L-lactate dehydrogenase.

- Holbrook's group achieved a highly active, malate-specific dehydrogenase by redesigning *Bacillus stearothermophilus* lactate dehydrogenase framework (Wilks et al. 1988; Clarke et al. 1989; Wagner and Benkovic 1990). This involved three amino acid replacements, namely, D197N, T246G, and Q102R.
- Human arginase II is highly specific and acts on arginine to produce ornithine and urea. A single amino acid replacement (the N149D variant) converts this enzyme into an agmatinase with almost no activity on arginine (Lopez et al. 2005).

- Glucose isomerase site-directed mutagenesis has resulted in enzyme forms with altered pH optima and altered divalent metal ion specificity.
- The rational modification of enzymes to change or extend their coenzyme specificity has also been possible (Moon et al. 2012). The NADP-specific glutathione reductase was altered into a NAD-dependent enzyme. Similar efforts in other dehydrogenases such as glutamate dehydrogenase are reported.
- There are many attempts to rationalize the enzyme stability as a useful parameter through site-directed mutagenesis (Bryan 2000). These have been largely empirical because we do not yet fully understand what contributes to protein stability.
- Site-directed mutagenesis provides enzymologists an opportunity to tinker with active site residues and ask some really interesting mechanistic questions. Almost every position in the subtilisin sequence has been subjected to SDM. The three critical active site catalytic residues S221, H64, and D32 of subtilisin were evaluated through this approach (see Chap. 36 Chymotrypsin). For the subtilisin S221A mutant (which still retains some activity!), the reaction cannot proceed by the usual serine acyl-enzyme intermediate (ping-pong mechanism). While such a mechanistic change from an acyl-enzyme to a direct water attack may be possible, the converse is much more difficult because the essential nucleophile would be missing. It is very easy to lose enzyme function but much more difficult to gain a new function through SDM!
- The H95 residue of triosephosphate isomerase is important for catalyzing the enolization of the substrates. The H95Q mutant is impaired in its ability to stabilize this reaction intermediate. There is an associated change in the proton transfer pathways mediated by the mutant enzyme (Nickbarg et al. 1988).
- The cancer-associated isocitrate dehydrogenase mutant (R132H form of IDH1 isozyme) loses its native function but is able to catalyze the NADPH-dependent reduction of 2-oxoglutarate to *R*(-)-2-hydroxyglutarate (D-2-hydroxyglutarate) (Dang et al. 2009).

Site-specific mutagenesis approach has undoubtedly extended our knowledge of enzyme mechanism and function. The necessary first step in the kinetic analysis of a mutant enzyme is to show that any observed change in catalytic activity is solely because of the targeted alteration. For this, the enzymologist must be prepared to analyze both the structural and functional consequences of the mutation(s) made. To begin with, this should include a thorough evaluation of kinetic parameters like $K_M$, $V_{max}$, $k_{cat}$, and $k_{cat}/K_M$. Gross structural changes in a mutant enzyme can be discerned through techniques like circular dichroism and gel filtration chromatography. X-ray crystallography of the native and mutant enzyme forms (as well as their frozen structures bound to substrate, product, inhibitors, and transition-state analogs) offers valuable information. Of course, the pretty structures should conform to the hard data (Miller 2007). It is a distinct possibility that mutant enzymes might follow a different reaction pathway. This was highlighted above, with examples of triosephosphate isomerase and serine proteases. Therefore, a detailed mechanistic analysis is routinely needed for proper appreciation of the effects from site-directed mutations.

**Changing Landscape of Enzyme Allostery** Historically, understanding the control of enzyme activity through feedback mechanisms led to the simultaneous discovery of enzyme cooperativity and allostery. Feedback inhibition and cooperativity over the years have got connected and appear to be two faces of the same coin. We have seen their implications to metabolic regulation in the previous chapter (Chap. 37 Regulation of enzyme activity). The non-Michaelian kinetic feature allows an enzyme to function as a concentration-dependent metabolic switch. Classical models invoke distinct enzyme conformations associated with allostery. These may involve multiple ligand binding sites (with associated conformational selection and induced fit) or polypeptide oligomerization. In most allosteric enzymes, the allosteric binding site for the ligand lies far away from the active site. Therefore, structural communication paths (via the levers and pulleys of protein structure) must exist between these sites. First mechanistic description for allosteric regulation was proposed more than sixty years ago; other possible mechanisms are being advanced regularly to describe these phenomena (Changeux 2013; Motlagh et al. 2014). Investigations at atomic detail (by high-resolution NMR of specifically labeled side chains) of glucokinase (a monomeric enzyme) for its kinetic cooperativity have provided recent insights. The enzyme molecule samples a number of conformational states in the absence of glucose. However, this population of conformations shifts toward a narrow, well-structured ensemble of states in the presence of glucose (Larion et al. 2012).

Allostery is not always mediated by conformational changes that can be detected by standard techniques like X-ray and NMR relaxation measurements, determination of H/D exchange rates, and isothermal titration calorimetry (ITC) experiments. These methods give a time-averaged snapshot of the protein 3D structure. However, recent advances in spectroscopy (probes to explore time-resolved dynamics of protein conformational changes) and the computational approaches (to study molecular dynamics simulations) indicate that multiple conformational states of the enzyme exist even in kinetically simple Michaelis complexes. Also, the free enzyme itself is a collage of protein conformational states. This protein disorder is clearly observed in the dynamic motions as measured by distance- and time-resolved NMR studies. Present-day structural biology offers unequivocal evidence of multiple conformations in preexisting equilibrium for glucokinase, trypsin-like proteases, maltose-binding protein, etc. The new outlook on allostery incorporates this more dynamic view of the enzyme protein. Accordingly, allosteric control may manifest by a *population shift* in the statistical ensembles of many states, with some regions of low local stability and others of high stability. Ligand binding affects the relative free energies of these states; they in turn differ in their affinities for other ligands and/or their activity. This is different from the earlier concept of a few well-defined static conformational states (such as R and T states). Various mechanisms of allostery described to date are summarized in the Table 39.1.

The new view of allostery encompasses a conformationally dynamic continuum of allosteric phenomena. With recent discoveries, we have moved towards increasing enzyme protein dynamics (disorder or fluctuations) starting from (a) rigid body

**Table 39.1 Molecular mechanisms of allosteric regulation**

| Mechanism | Examples |
|---|---|
| Closure/opening of active site | 3-Phosphoglycerate dehydrogenase (the active site cleft closes upon binding of the end product, Ser) |
| Changes in active site conformation | 3-Deoxy-D-arabinoheptulosonate-7-phosphate synthase (aromatic amino acid (the end product) binding leads to minor conformational modifications and prevents substrate binding) |
| Change electrostatic properties of active site | Chorismate mutase (aromatic amino acid (the end product) binding brings a Glu residue into the active site; causes a major change in its electrostatics and repels the negatively charged substrate) |
| Influence protein–protein complex formation | ATP phosphoribosyltransferase (His (the end product) binding converts an active dimer to an inactive hexamer) |
| Affect protein flexibility | Dihydrodipicolinate synthase (Lys (the allosteric ligand) binding affects distant sites via a change in the protein vibrational modes) |
| Shift in ensemble of conformer population | –Glucokinase (its intrinsically disordered small domain samples a broad conformational ensemble; upon glucose binding, the population shifts toward a narrow, well-ordered ensemble) |
| | –Dimeric catabolite activator protein (binding of the first cAMP molecule lowers the affinity for the second cAMP molecule; first cAMP binding enhances motions within the protein, whereas the binding of second cAMP decreases these motions and flexibility |

movements to (b) side-chain dynamics, (c) backbone dynamics, (d) local unfolding, and finally (e) intrinsically disordered structures.

Allostery, by definition, involves the propagation of signals between different sites in a protein. This may occur in the absence of detectable conformational changes and may be exclusively mediated by transmitted changes in protein motions. The fact that the change in dynamics occurs in the absence of significant structural change suggests that dynamics alone may convey allosteric information. There is much interest to probe the existence of an entire channel/network of amino acids through which allosteric signals are communicated. One approach that could map such paths (perturbations that travel across the structure) and implicate the interacting amino acid residues is through measurements of a double-mutant cycle. Energetics of such residue interactions allows us to infer the degree of functional coupling between different sites of a protein. A complementary approach uses a sequence-based statistical analysis (Statistical Coupling Analysis, SCA) method for elucidating the architecture of functional couplings in proteins (Reynolds et al. 2011). If two residues in a protein are functionally coupled, then they should have coevolved. This coevolution can be scored by statistically comparing homologous protein sequences. A combination of SCA and double-mutant cycles along with functional, structural, and folding analyses can provide insights into the existence of an entire wave, wire, channel, or network of amino acid residues through which allosteric signals are transmitted.

Transplantation of allosteric regulation has been possible in hemoglobins and glycerol kinase (Hardy and Wells 2004). Very few residues may be involved in the manifestation of allosteric regulation. The allosteric features of crocodile hemoglobin could be introduced into human hemoglobin by substitutions at 12 amino acid residues. Changing only 11 of 501 total residues (about 2%) converts an unregulated glycerol kinase to an allosterically regulated glycerol kinase. Small molecules can exert strong effects from unexpected locations, and hence searching for new allosteric sites in enzymes is a challenge. It is hard enough to predict what ligand might bind in a binding pocket, but the presence of allosteric binding sites further complicates the matter. It may become possible to develop algorithms to distinguish allosteric from active sites with better databases in the future. It may also be feasible to predict allosteric regulation from protein structural data (Freire 2000).

The emerging radical view of enzyme function is that each catalytic step corresponds to an ensemble of thermodynamic and structural states. Incidentally, allostery and catalysis no longer appear as distinct phenomena but as the manifestations of the same intrinsic protein dynamics. Important issues for further research in allosteric enzymes include (a) the mechanisms by which an allosteric effect is transmitted via amino acid networks, (b) how the distribution of protein conformations is altered, and (c) the timescales at which the redistribution of these conformations occurs.

**Predicting Enzyme Structure and Function** Computational tools have revolutionized the whole of biology. It is no wonder that substantial progress has taken place in enzymology, one of the more quantitative of the biological sciences. Incorporating new tools and technology to understand enzyme catalysis is a recurring theme in enzymology. Computational enzymology has joined the earlier methods involving kinetic measurements, kinetic isotope effects, crystallography, and distance−/time-resolved NMR. We have noted above that computational enzymology has made significant inroads into transition-state analysis and molecular dynamics of enzyme action. Besides these, there is a rapid move toward enzyme structure and function prediction through computational approaches. This need has arisen as we (a) accumulate a large number of sequenced genomes; (b) come across orphan open reading frames, with no clues of their function; and (c) express sequences into proteins and even crystallize them without their actual functional demonstration (Cuesta-Seijo et al. 2011; Hai et al. 2015).

Assigning valid functions to unknown (putative!) proteins/enzymes identified in genome projects is a challenge (Kuznetsova et al. 2005). While experimental testing remains essential, computational approaches can help guide this experimental design. Bioinformatics approaches are being perfected to (a) identify informative sequence relationships using structure and genome context, (b) allow accurate high-throughput structure prediction through homology modeling, and (c) dock metabolites in silico to provide accurate and testable list of potential enzymes. Microbial metabolic pathways often are encoded by *genome neighborhoods* (synteny and associated gene clusters and/or operons). Such positional information can provide important clues for enzyme function assignment. For instance, pathway

docking is an efficient strategy for predicting in vitro enzyme activities and allocating in vivo physiological functions. It has been possible to identify novel metabolites, enzyme activities, and biochemical pathways through this tactic.

In view of the exponential growth in genome sequence data sets (with significant proportion of sequences with unknown enzyme functions), an integrated strategy for functional assignment was recently proposed (Gerlt et al. 2011; Gerlt 2017). This enzyme function initiative (EFI; the multicentric program under the National Institute of General Medical Sciences, USA) looks to predict the substrate specificities of unknown members of mechanistically diverse enzyme superfamilies – thereby predicting their functions. The approach exploits conserved features within a given superfamily such as known chemistry, identity of active site functional groups, and composition of specificity-determining residues/motifs/structures. Initial enzyme targets chosen for this purpose include members of the amidohydrolase, enolase, glutathione transferase, haloalkanoic acid dehalogenase, and isoprenoid synthase superfamilies. Members of these enzyme superfamilies are functionally diverse (conserved partial reactions or chemical capability but with divergent overall function) which makes functional assignment difficult. Homology inferred from simple sequence comparisons alone cannot guide functional assignment in such situations. Therefore, an integrated approach involving the following components is proposed:

- Perform bioinformatic analysis to cluster sequences into probable isofunctional groups; assign tentative functions for further investigation by structure determination, structural modeling/docking, and biochemical experimentation.
- Carry out homology modeling to expand the use of structural models; thereby guide functional assignment to proteins without experimentally determined structures.
- Employ computational docking methods to leverage structure and guide functional assignment by suggesting substrates/ligands for biochemical experimentation.

That enzyme function needs to be defined not only through its chemical and kinetic competence but also by associated structural features has become increasingly obvious. The change in the title and emphasis of the very well received book on enzymes by Alan Fersht (from "Enzyme Structure and Mechanism" for the first edition to "Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding" in the later version) is a pointer in this direction. Structure is the necessary third leg, along with mechanism and function, of the secure stool to understand enzyme function.

## 39.4 Designing Novel Catalysts

Curiosity and the desire to imitate general principles of biological catalysis have led to many developments in the design and construction of artificial enzymes. Present approaches to create novel catalysts fall into three general categories – (a) de novo design and synthesis of catalysts from polypeptides and nonprotein building blocks

(such as macrocyclic compounds, poly-ethyleneimine, synthetic chemical/genetic polymers, and cyclodextrins); (b) modification/evolution of existing catalysts such as protein enzymes or ribozymes, by genetic or chemical methods; and (c) designer enzymes made to order. While antibodies and RNA as catalysts will be discussed subsequently, we will first focus on chemical models, enzyme mimics, and hybrid catalysts.

**Chemical Models, Enzyme Mimics, and Hybrid Catalysts** Enzyme models of increasing complexity have been designed and discovered. A small chiral molecule-like proline may be considered the simplest enzyme. Recent reports on asymmetric catalysis by proline and its derivatives include activation of carbonyl compounds via nucleophilic enamine intermediates (MacMillan 2008). Several highly enantioselective important carbon–carbon bond-forming reactions (aldol additions and Mannich reaction) have been developed using this approach. Many model systems anticipated their enzyme counterparts much ahead of time. Otto Warburg studied the oxidation of unsaturated fatty acids by combined action of iron and sulfhydryl (-SH) groups in 1925. Lipoxygenases, containing iron and –SH groups essential for their oxidative activity, were discovered much later. Aniline-catalyzed rapid decarboxylation of acetoacetate via the "aniline–acetoacetate complex" is exemplified by JBS Haldane in his 1930 book on enzymes. More recently, chemical hydrogenase mimics with Co and Ni centers are reported for exploiting hydrogen as fuel. Models for enzymes performing free radical chemistry and other redox reactions are being sought to be incorporated in clean energy programs. The study of glutathione peroxidase anticipated the small molecular enzyme model ebselen as well as an antibody enzyme – abzyme (see Table 39.2 in the next section for a full treatment).

As pointed out before, there are two reasons to study enzyme constructs and mimics. One objective of chemists and biologists is to elucidate molecular basis of enzyme function. Secondly, using the available knowledge base (which is still far from complete!), one could attempt to design and build novel catalysts – the so-called tailor-made enzymes. The chemical alteration of an existing enzyme by introducing additional functional groups is one route to rational enzyme design. The *semisynthetic enzyme* so generated can display very different catalytic activities from that of the parent enzyme. For example, papain (a thiol protease) was converted to an effective redox catalyst (an oxidoreductase) by appropriately attaching a flavin to the enzyme sulfhydryl group. In another example, new binding domains are selectively introduced to build/alter enzyme specificity. RNase A is a relatively nonselective enzyme hydrolyzing phosphodiester linkages of RNA. It was made specific to a definite RNA sequence by creating a chimeric RNase A – wherein a covalently attached single-stranded DNA confers specificity by annealing at the complementary RNA sequence (Fig. 39.1).

The oligonucleotide-tagged RNase A above is an effort in modifying the substrate specificity of an enzyme. However, another attempt was made to build a restriction endonuclease from first principles. Several chemical agents intrinsically possess DNA cleavage activity. If these are incorporated into DNA/RNA binding proteins,

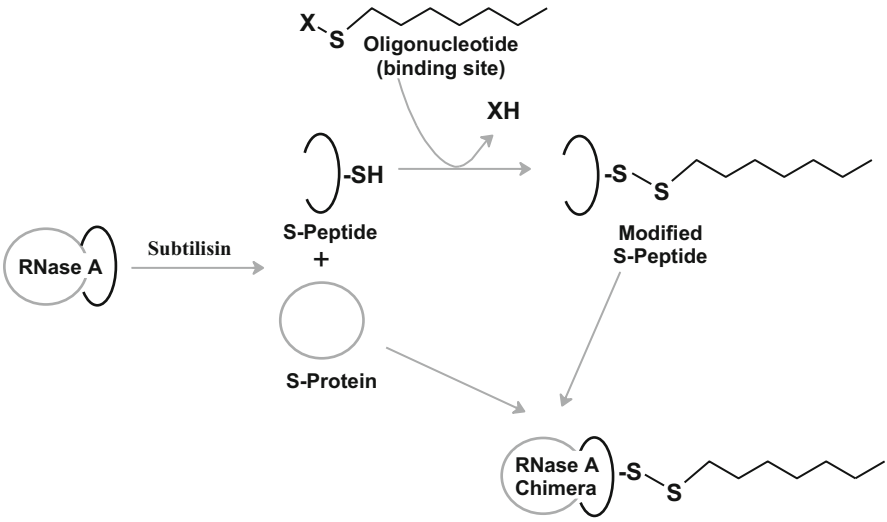


**Fig. 39.1 Construction of a sequence-specific RNase A chimera**. The S-peptide was modified with an oligonucleotide sequence. The functional RNase A was reconstituted by combining the modified S-peptide with the S-protein. (Adapted with permission from Zuckermann and Schultz, J Am Chem Soc, 110:6592–6594. Copyright (1988) American Chemical Society.)

a specific nuclease may be created. Indeed, one could design synthetic hybrid molecules with two components – one for specific DNA sequence recognition and the other to cleave the DNA adjacent to it. The technique called *affinity cleavage* is based on the construction of such bifunctional molecules (Fig. 39.2). Many artificial metalloenzymes have emerged from similar approaches. Chemistry of EDTA-Fe[II] with oxygen is an excellent source of hydroxyl ($HO^{\bullet}$) radicals for DNA cleavage. This reaction can be turned into a catalytic cycle by adding ascorbate so that Fe[III] formed is reduced back to Fe[II]. Several chemical agents possess nuclease activity; rare earth metal ions are active in hydrolyzing phosphodiester linkages in DNA. For instance, DNA scission by Ce[IV] is hydrolytic and not oxidative. The fragments so generated can be religated using a ligase. The *sequence-recognizing moiety* brings the *molecular DNA scissors* close to the target phosphodiester bond. Binding to specific DNA sequences is provided either by a short oligonucleotide or by a specific DNA-binding domain of the protein. Sequence recognition has also been achieved by pcPNA (a pseudo-complementary peptide nucleic acid).

Most restriction enzymes used in molecular biology recognize a stretch of six bases and cut DNA at that site. Assuming random distribution of A, G, T, and Cs, there is a high probability ($4^6$, which is once every 4096 bp of DNA) of finding these restriction sites in DNA. To cut at a single unique position in human genome, we could use a 16mer (or longer) sequence-recognizing moiety ($4^{16}$; for $>10^9$ bp of DNA in human genome!). This site selectivity compares well with the molecular tools like (a) the *Achilles' heel cleavage* (only one of the many restriction sites in a genome is specifically protected from inactivation by a cognate methyltransferase;

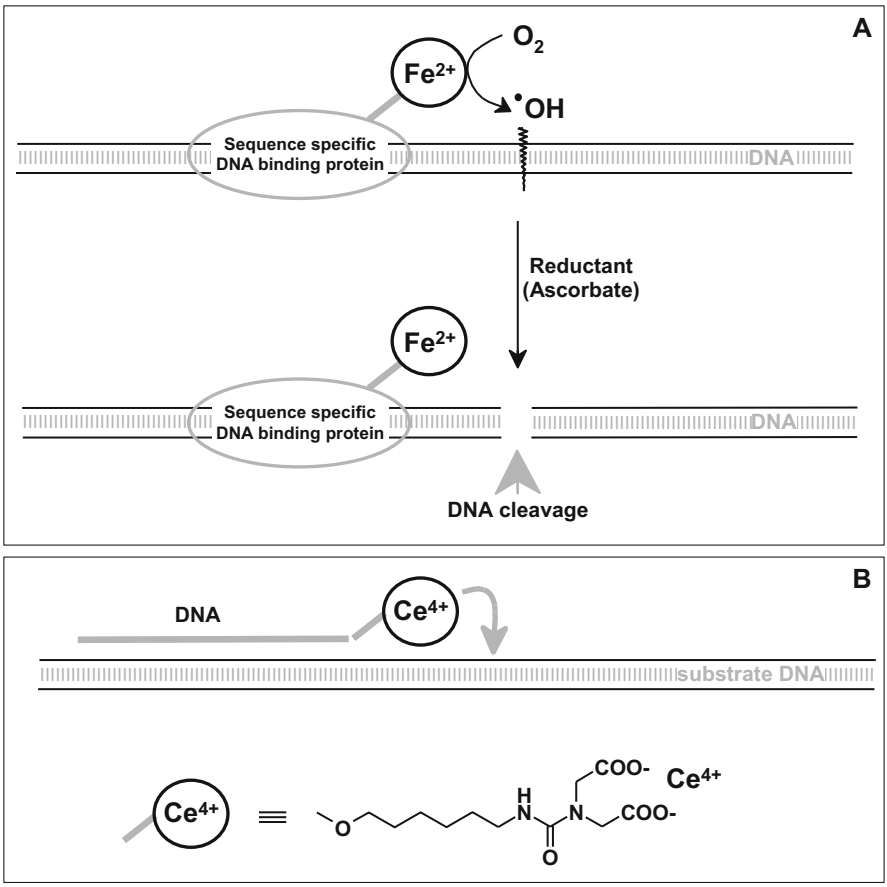


**Fig. 39.2 Construction of artificial restriction enzymes** (**A**) Fenton chemistry of EDTA-Fe[II] with oxygen generates hydroxyl ($HO^{\bullet}$) radicals for local DNA cleavage. Specificity is provided by using a unique sequence-recognizing moiety (e.g., *trp* or *lac* repressor or catabolite activator protein of *E. coli*). (**B**) A restriction enzyme construct of $Ce^{4+}$ complex of iminodiacetate. This $Ce^{4+}$ nuclease is juxtaposed to the specified DNA sequence for cleavage by an oligonucleotide sequence-recognizing moiety

this in turn creates a unique restriction enzyme cleavage site) and (b) the genome editing by the CRISPR/Cas system. An artificial restriction enzyme that can cut only at one position in the human genome will require a 16mer (or longer) sequence-recognizing moiety.

Novel catalysts may also be crafted on nonprotein molecular framework – including synthetic macromolecules of nonbiological origin (Bjerre et al. 2008; Wulff 2002). Poly(ethyleneimine) polymers possess intrinsic acid–base groups; along with pyridoxal phosphate or other suitable cofactors, some degree of catalysis was demonstrated. Breslow's group has made extensive efforts to emulate enzyme catalysis using cyclodextrin scaffolds (Breslow 2005). While the cyclodextrin cavity

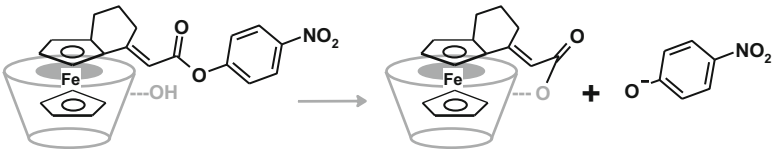


**Fig. 39.3 A β-cyclodextrin mimic of chymotrypsin displaying catalysis of *p*-nitrophenyl ester hydrolysis**. The cyclodextrin cavity binds the hydrophobic ferrocene core of the ester substrate

offers a binding pocket, additional functional groups attached to the cyclodextrin ring offer new enzyme mimics (synzymes). The cavity size can be varied by choosing either β-cyclodextrin (seven glucose units in the ring) or γ-cyclodextrin (eight glucose units in the ring). Mimics of RNase A (β-cyclodextrin bis-imidazole) and chymotrypsin (for its esterase activity!) are demonstrated. A flexible capped cyclodextrin with the well-fitting substrate afforded a rate acceleration of $8 \times 10^7$ fold in *p*-nitrophenyl ester hydrolysis (Fig. 39.3).

Lastly, a word of caution on rate acceleration observed for enzyme mimics. Most protease/esterase models use *p*-nitrophenyl esters as substrates. The *p*-nitrophenyl esters are *not* protease/esterase substrates "in real life." With *p*-nitrophenyl esters as model substrates, partly the better leaving group effect contributes to the observed rate accelerations. Leaving groups whose pKas are above 9.0 lead to much lower acceleration. One should account for such a *p-nitrophenyl ester syndrome* when evaluating the performance of synzymes (Menger and Ladika 1987).

**Antibody Catalysts (Abzymes)** Linus Pauling recognized in the 1940s that the ability of an enzyme to speed up a chemical reaction arises from the "complementarity of its active site structure to the activated complex (i.e., the transition state)." This has given rise to the productive field of catalytic antibodies (or *abzymes*). Antibody molecules represent a class of proteins with high affinity and exquisite selectivity; they could be raised against any small molecule (hapten) of our choice. Antibodies that can bind to transition-state analogs of a substrate should therefore catalyze the conversion of that substrate, through the transition state, to the product. Abzymes production involves the following steps:

- Generating a stable transition-state analog using molecular design and chemical synthesis
- Raising antibodies (monoclonal antibodies to be more precise) with the *TS* analog as the hapten
- Isolating antibodies which bind to the *TS* analog as potential catalysts for that reaction

The seminal prediction of Pauling was verified by Lerner's research group in the 1980s. During an ester hydrolysis, the $sp^2$ hybridized carbonyl carbon is converted to a $sp^3$ hybridized carbon in the intermediate; and the carbonyl oxygen resembles an

oxyanion. The transition state presumably looks more like this unstable intermediate ($sp^3$, oxyanion). A phosphonate ester mimic, with a $sp^3$ hybridized phosphorous replacing the $sp^2$ hybridized carbonyl carbon (Fig. 39.4), was synthesized. While very resistant to hydrolysis, this phosphonate also has negatively charged oxygen similar to the intermediate during ester hydrolysis. The mouse antibodies (the monoclonal – 6D4) against this phosphonate structure catalyzed the corresponding carboxylic acid ester hydrolysis.

More than 100 interesting examples of abzyme catalysis are known. They include many reactions that cannot be achieved by standard chemical methods (Benkovic 1992). Besides the ester hydrolysis mentioned above, these include pericyclic processes, group transfer reactions, additions and eliminations, redox reactions, aldol condensations, and a few cofactor-dependent transformations. Some of these are listed in Table 39.2 below.

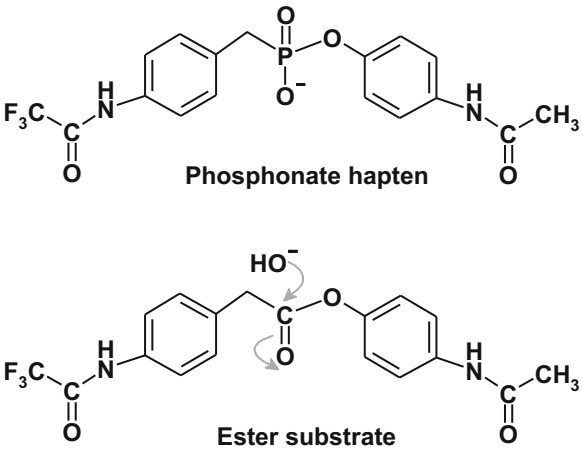


**Fig. 39.4 The phosphonate transition-state analog used as hapten**. The monoclonal antibodies (abzyme 6D4) against this hapten displayed esterase activity on the corresponding ester shown

**Table 39.2 Antibody catalysts generated for different reaction types**

| Abzyme[a] | Reaction catalyzed |
|---|---|
| 6D4 | Carboxylic ester hydrolysis |
| 48G7 | *p*-Nitrophenyl ester hydrolysis |
| 1F7 | Claisen rearrangement (chorismate to prephenate) |
| AZ-28 | Oxy-Cope rearrangement |
| 39-A11 | Diels-Alder reaction |
| 7G12 | Ferrochelatase |
| 33F12 | Aldolase |
| 34E4 | E2 elimination of nitrobenzisoxazole |
| 2F3 (scFv) | Glutathione peroxidase |

[a]All the abzymes (except 2F3) are monoclonal antibodies (Hilvert 2000). The scFv (**s**ingle **c**hain **F**ragment **v**ariable) of abzyme 2F3 was activated by chemical modification of a reactive Ser by attaching Se to it (Ren et al. 2001)

Virtually all experiments with abzymes employ monoclonal antibodies (Table 39.2). The monoclonals are a single homogeneous catalytic species (unlike the polyclonal sera), and their use greatly simplifies kinetic, mechanistic, and structural characterization of an abzyme. While expressing catalytic antibodies (or Fab fragments), folding of the two chains into native state is a challenge. One approach is to express and secrete them into the periplasmic space of *E. coli*, exploiting its diminished protease activity and ability to correctly fold and form disulfide bonds. The **s**ingle **c**hain **F**ragment **v**ariable (**scFv**) version of the antibody exhibits the same catalytic parameters as the parent monoclonal antibody. The scFv is a recombinant protein construct of a VL chain tethered to a VH chain with a polypeptide linker and is expressed efficiently by bacteria.

Fastest enzymes are diffusion limited catalysts. Even the less than *perfect* ones typically have apparent bimolecular rate constants ($k_{cat}/K_M$) between $10^6$ and $10^8$ $M^{-1}\,s^{-1}$. Catalytic antibodies have rate accelerations many orders of magnitude below their enzyme rivals. By definition, abzymes are catalytic antibodies specific to (and raised against) the corresponding transition-state mimic. We note that stabilization of the transition state is *necessary but not sufficient* by itself for achieving good catalysis. Many other factors – like active site functional groups, conformational flexibility, shielding of the reaction intermediates, cofactor needs, etc. – substantially contribute to rate accelerations (refer Chap. 6). Transition states themselves have fleeting lifetimes and cannot be captured. Synthesis of effective *TS* analogs must therefore draw on our chemical intuition about the conformational, stereochemical, and electronic features of the reaction under study. Since no stable molecule can reproduce all the characteristics of the actual *TS*, hapten design strategies have focused on incorporating the salient features of the *TS*. In the sum, limited catalytic ability of abzymes may be attributed to one or more of the following:

- Antibody scaffold is fixed, and this means limited structural space is explored for catalyst building. Real enzymes come in a variety of structural folds.
- The polypeptide sequence space exploited during antibody maturation is also limited.
- Catalytic antibodies may lack structural dynamics necessary for optimal catalysis.
- *TS* analogs are imperfect mimics of the actual transition state. The *TS* mimic designed may not capture the best options for discriminatory binding of the substrate versus *TS*.

Basic strategy to produce catalytic antibodies is indirect. Here the immune system is directed to evolve not for catalysis but toward binding tightly to an imperfect *TS* analog. To date, antibody enzymes display only modest catalytic activity and have not found significant practical utility (though a few have reached the market!). Nonetheless they continue to be of considerable academic interest. Studying abzymes has yielded valuable insights into reaction mechanisms, catalysis, enzyme structure, and function. There is yet the promise of delivering tailored catalysts for difficult reactions for which natural enzymes do not exist. Such catalytic antibodies may be useful even if they do not attain enzyme-like efficiency.

**RNA Catalysts (Ribozymes)** Substantial role for RNA as information molecule is well established in molecular biology. More recently, RNAs have assumed importance as components of (a) gene silencing through double-stranded small interfering RNAs (*siRNA*) generated by Dicer, an RNase III-like enzyme, and (b) genome editing through the CRISPR/Cas system where crRNA-guided interference is exploited. The catalytic role of RNA molecules (the so-called ribozymes) has expanded the realm of biological catalysis beyond proteins. RNA catalysts do satisfy the dual criteria of *catalysis* and *specificity* but are less impressive catalysts when compared to protein-based enzymes. They have had much more impact in understanding the origin of life problem and catalytic evolution. Like polypeptides, RNA molecules can fold into higher-order structures that permit the formation of an active site. On the other hand, DNA is predominantly double helical, cannot fold into complicated shapes, and has limited repertoire of chemical groups (it lacks the 2′-hydroxyl group) for catalysis. Nevertheless, DNA-based catalysts (DNAzymes) were constructed (Baum and Silverman 2008). More recently, elaboration of different catalytic activities from synthetic genetic polymers (XNAs) was demonstrated (Taylor et al. 2015). The XNAs (arabino nucleic acids, hexitol nucleic acids, or cyclohexene nucleic acids) fold into defined structures and bind ligands. Few XNAzymes were elaborated directly from random XNA oligomer pools; some of them exhibited *in trans* RNA endonuclease and ligase activities.

Single-stranded nucleic acid molecules are capable of folding into secondary and tertiary structures. Aptamers are short, single-stranded nucleic acids which bind a variety of ligands with high affinity and specificity. DNA or RNA aptamers can be routinely isolated from synthetic combinatorial nucleic acid libraries by in vitro selection – known as "systematic evolution of ligands by exponential enrichment" (SELEX) (Tuerk and Gold 1990; Weigand and Suess 2009). Riboswitches are natural versions of aptamers discovered subsequently. The aptamer domains of most riboswitch classes are typically fewer than 100 nucleotides. Riboswitches with distinctive ligand recognition capabilities have been found in all domains of life; they occur with highest frequency within the 5′-UTRs of bacterial mRNAs and typically regulate genes involved in metabolism. For instance, glycine riboswitch consists of two different aptamer types that individually bind to a single molecule of glycine (Famulok 2004). Cooperative interaction between the two sites allows better sensing of this metabolite. Aptamers and riboswitches are specific ligand binding RNAs with no catalytic function. They may fold or undergo a conformational change upon binding the cognate ligand. An aptamer may be fused to a ribozyme to generate *aptazymes*. Binding of its cognate ligand to the aptamer displaces it from the stand on which it is bound on the ribozyme, thereby providing for regulation of ribozyme function.

Catalytic RNA was discovered through RNA species capable of auto-cleavage reaction (Abelson 2017). Subsequently, other activities catalyzed by RNA have been reported (Doudna and Lorsch 2005; Wilson and Lilley 2009). While natural ribozymes catalyze mainly self-cleavage or ligation reactions, they can also accelerate other reaction types (Table 39.3). The RNA component of RNase P is responsible for catalytic processing of precursor tRNAs. The self-splicing group I intron from

**Table 39.3 Kinetic constants for some catalytic RNAs**

| Catalyst | Substrate | $k_{cat}$ ($s^{-1}$) | $K_M$ (M) | $k_{cat}/K_M$ |
|---|---|---|---|---|
| RNase A | RNA (cleavage; best substrate) | – | – | $2.8 \times 10^9$ |
| Ribozyme | RNA (cleavage) | $2.0 \times 10^{-1}$ | $2.0 \times 10^{-8}$ | $1.0 \times 10^7$ |
| Self-splicing intron | RNA (splicing) | $1.0 \times 10^{-3}$ | $1.1 \times 10^{-9}$ | $9.0 \times 10^5$ |
| Ribosome | (Peptide bond formation) | $5.0 \times 10^0$ | $5.0 \times 10^{-3}$ | $1.0 \times 10^3$ |
| DNAzyme IV | Ornithine decarboxylase mRNA | $1.0 \times 10^{-3}$ | $3.0 \times 10^{-7}$ | $3.0 \times 10^3$ |
| Abzyme 34E4 (E2 elimination) | Nitrobenzisoxazole | $6.6 \times 10^{-1}$ | $1.2 \times 10^{-4}$ | $5.5 \times 10^3$ |

Catalytic efficiencies of RNase A, a DNAzyme, and an abzyme are also listed for comparison

*Tetrahymena thermophila* was engineered to perform as a multiple turnover RNA enzyme. The peptide bond formation is attributed to the RNA component of the ribosome (Schmeing and Ramakrishnan 2009).

There are similarities as well as differences between RNA and protein catalysts. Greater structural variety (of amino acids) allows better catalytic properties in protein enzymes than in RNA enzymes. Enzymes are superior catalysts, and many of them function at the diffusion limit. Ribozymes, on the other hand, are rather slow with an apparent maximal rate constant of ~1 $min^{-1}$ (Doudna and Lorsch 2005). However, ribozymes might be easier to produce than enzymes. Rigidity allows an enzyme to maximize specific (binding) interactions with the *TS* relative to the ground state and hence maximizes catalysis. Larger size ensures better positioning and rigidity within the active site – features required for most effective catalysis (see section on "Then why are enzymes big?" in Chap. 38). For the same catalytic function, RNA has to be much bigger than an enzyme protein. Protein enzymes are big and RNA enzymes are even bigger. RNA is clumsier than proteins in terms of functional groups, structural variety, and ability to fold. Hammerhead ribozyme and pancreatic RNase A have similar sizes, but RNase A (with best $k_{cat}/K_M$ of $2.8 \times 10^9$ $M^{-1} \times s^{-1}$) achieves $10^5$-fold higher maximal rates. A direct comparison of catalytic strategies (Table 39.4) available to protein enzymes and ribozymes is illustrative (Narlikar and Herschlag 1997).

## 39.5 Enzymes Made to Order

> *Until man duplicates a blade of grass, nature will laugh at his so-called scientific knowledge.*
>
> Thomas Edison

From the general theme of this book and earlier discussion in this chapter (Chap. 39 Future of enzymology – An appraisal), it is obvious that much is known about how enzymes function as catalysts. One measure of how well we understand enzymes is to try and build similar catalysts from first principles. In this sense, de novo enzyme design is an intellectual challenge, and the exercise serves two important objectives.

**Table 39.4 Catalytic strategies of protein and RNA enzymes: A comparison**

| Catalytic strategy | Protein enzyme | RNA enzyme |
|---|---|---|
| Substrate orientation and approximation | Yes | Yes |
| General acid–base catalysis | Always | Deficient[a] |
| Metal ion catalysis | Many | Always |
| Organic cofactors | Many | None |
| Active site electrostatics (dielectric manipulated effectively) | Most proficient | Not proficient |
| Utilization of binding energy from interactions away from active site | Yes | Yes |
| Covalent catalysis | Yes | Yes |

[a]RNA enzymes lack groups with pKas around pH 7.0. The nucleolytic ribozyme GlmS provides an exception to the exclusive use of nucleobases in general acid–base catalysis by ribozymes; a molecule of glucosamine-6-phosphate specifically bound to the RNA structure serves as the general acid in GlmS

It allows experimental validation of the principles of catalysis that we have learned so far. Second, tailor-made catalysts can be built for industrial applications, particularly for those reactions for which natural enzymes do not exist. The design of enzymes with new functions and properties has long been a goal of the protein engineer. However, in enzyme engineering (and de novo design), serendipity continues to outstrip design – a clear sign that our basic understanding of enzyme catalysis to date is far from complete.

Enzyme catalytic power results from a combination of multiple mechanistic strategies (Chap. 6). Therefore, valuable insight into the evolution of catalytic function can be gained through de novo design experiments. However, the methods of kinetic analysis discussed earlier in this book will be applicable to all catalysts (synthetic or natural) regardless of their chemical nature. We may note that foundations of enzyme kinetics were laid much before the chemical nature of enzymes as proteins was established by Sumner (using jack bean urease!).

A major benefit of recombinant DNA technology is the ability to do protein and, hence, enzyme engineering. This includes the skill to precisely replace/delete/add one or more amino acids in a given enzyme. These designed yet specific mutations can be engineered on a desired gene (coding for the desired enzyme) with the help of synthetic oligonucleotide constructs as primers (see *Site-Directed Mutagenesis and Crystal Structures* section, in this chapter). It is very easy to lose an enzyme function through site-directed mutations, but very difficult to gain a new function. Besides point mutations, other functional elements of a protein scaffold could also be replaced/changed through available recombinant DNA tools (Fig. 3.5). The approach is powerful and is anticipated to deliver many tailored enzymes.

Engineering novel enzymes is a rapidly evolving field, and the examples presented here are selective and only representative. A foolproof and robust enzyme activity assay is at the heart of any enzyme engineering and design. While the

methodological details are outside the scope of this book, various strategies to generate novel enzyme designs through recombinant DNA technology are outlined below.

**Enzyme Redesign** One way to generate novel catalysts is to start from an existing enzyme scaffold and rationally alter its structure to effect a predicted change in function. Enzyme redesign may be achieved through simultaneous incorporation and/or adjustment of protein functional elements – through deletion, insertion, loop grafting, and substitution of relevant active site loops, generating chimeras, etc. This could be followed by point mutations to fine-tune the enzyme activity. Using one such approach, a β-lactamase activity was introduced into the scaffold of glyoxalase II. The resulting enzyme completely lost its original activity but catalyzed the hydrolysis of cefotaxime (see Table 39.7 below). Three other interesting examples include the following:

(a) The domain swap in bacterial glutamate dehydrogenases to change their pyridine nucleotide specificity (Sharkey and Engel 2009).
(b) Mixing and matching of different modules of polyketide synthases and non-ribosomal peptide synthases leading to product diversity and the production of hybrid or novel antibiotics (Penning and Jez 2001).
(c) The cancer-associated mutations from isocitrate dehydrogenases which were extrapolated to homologous residues in the active sites of homoisocitrate dehydrogenases, for the catalytic conversion of 2-oxoadipate to (R)-2-hydroxyadipate, a critical step for adipic acid production (Reitman et al. 2012).

Enzyme redesign as an approach is expected to deliver more tailor-made enzymes in the near future.

The protein engineering approach to redesign a known enzyme has the potential to bring about (a) altered substrate/cofactor specificity and improve catalytic efficiency, (b) enhanced enantioselectivity, (c) a change in metal ion specificity, (d) a desired pH optimum of an enzyme, (e) increased enzyme stability, and (f) alteration of an existing site to catalyze a new chemical reaction. Both close and distant mutations appear similarly effective in improving enzymes in terms of thermostability and catalytic activity. The mutations close to the active site are more effective than distant ones for changing enantioselectivity, substrate selectivity, and alternate catalytic activity of an enzyme (Khersonsky et al. 2006). Besides the *chemical mechanism first* view of enzyme redesign, it may be possible to take the more challenging *ligand specificity first* approach. Accordingly, one may introduce catalytic residues into the ligand binding site of a chosen protein to create an active site capable of catalyzing a chemical reaction. After all, the "acid test" of enzyme redesign is to engineer new catalytic activities and aim to change the reaction mechanism itself.

**Directed Enzyme Evolution** Evolution continues to be an excellent teacher of how to evolve/modify biological catalysts. Directed evolution as a tool accelerates the evolutionary process from millions of years to weeks! And natural enzymes are clearly not evolutionary dead ends. The field of directed evolution of enzymes has made rapid progress and provides a much larger scope and canvas than simple enzyme redesign (Bornscheuer et al. 2012). It has moved from being a means for studying the relationship between sequence and function to being an extremely powerful tool for optimizing biocatalysts for industry. This Darwinian approach involves generation of random mutant library (generation of genetic diversity) followed by biological selection (of the fittest) for the desired activity. The greatest advantage of directed evolution is its independence from the prior knowledge of enzyme structure. One need not know the nature of interactions between the enzyme and its substrate either.

The procedure of directed evolution consists of (i) generating a random gene library, (ii) expression of these gene variants in a suitable host, and (iii) screening/selection of library of enzyme forms for the property of interest (Fig. 39.5). It begins with the mutagenesis of the gene encoding the enzyme template of interest (GFE, gene of your favorite enzyme). This important first step should generate a representative yet exhaustive mutant library. The mutated gene library is then inserted into a suitable host (like *E. coli* or yeast) for expression, and the transformants are plated on selective media. Respective protein variants are chosen from the single colonies growing on these plates. In most directed evolution studies, additional cycles of

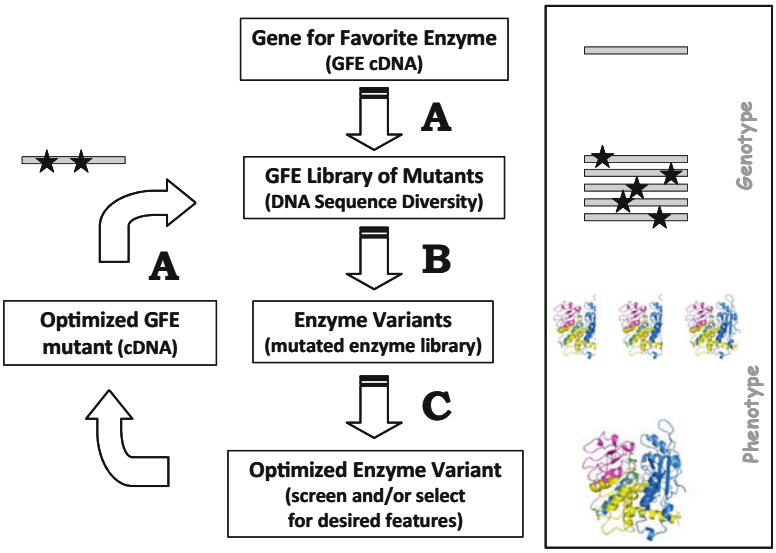


**Fig. 39.5 Iterative steps in the directed evolution of an enzyme**. Starting with the gene (cDNA) for the favorite enzyme (GFE), a random mutant library is generated (step A). These GFE variants are used to express respective enzyme variants (step B). The optimized variants are screened/selected (step C) and subjected to further cycles of directed evolution

**Table 39.5 Methods to create GFE mutant libraries and generate diversity**

| Library diversification strategy | Examples and tools |
|---|---|
| *Random mutagenesis* | |
| Chemical and physical mutagenesis | EMS, MNNG, nitrous acid, UV |
| Error-prone PCR (epPCR) | Supplementing $Mn^{2+}$ and/or unequal dNTP levels |
| Mutator strains | Mutagenesis plasmid (PACE), XL1-red strain of *E. coli* |
| *Focused mutagenesis* | |
| Site-directed saturation mutagenesis | NNK and NNS codons on mutagenic primers[a] |
| Computational strategies for high-quality library design | Rosetta design and computationally guided libraries, ISOR, consensus design, REAP, and SCHEMA |
| *Recombination (gene shuffling)* | |
| Homologous recombination | DNA shuffling, domain swapping, family shuffling, StEP, RACHITT, NExT, heritable recombination, ADO and synthetic shuffling |
| Nonhomologous recombination | ITCHY, SHIPREC, NRR, SISDC and overlap extension PCR (oePCR) |

Abbreviations and acronyms used in this table include – *ADO* assembly of designed oligonucleotides, *EMS* ethyl methanesulfonate, *ISOR* incorporating synthetic oligonucleotides via gene reassembly, *ITCHY* incremental truncation for the creation of hybrid enzymes, *MNNG* N-methyl-N-nitrosoguanidine, *NExT* nucleotide exchange and excision technology, *NRR* non-homologous random recombination, *PACE* phage-assisted continuous evolution, *RACHITT* random chimeragenesis on transient templates, *REAP* reconstructed evolutionary adaptive path, *SCHEMA* a computational algorithm, *SHIPREC* sequence homology-independent protein recombination, *SISDC* sequence-independent site-directed chimeragenesis, *StEP* staggered extension process, *UV* ultraviolet rays

[a]Where *N* can be any of the four nucleotides, K can be G or T, and S can be G or C

Adapted with permission from Packer and Liu, Nat Rev Genet, 16:379–394. Copyright (2015) Springer Nature

mutagenesis are necessary for obtaining an optimal catalyst. While many strategies are available (and continue to be developed) for generating exhaustive libraries, the success of a directed evolution effort depends greatly on the method chosen for finding the best mutant enzyme. And, in directed evolution, "we simply get what we screen for"!

A range of gene mutagenesis methods are available for generating sequence diversity in a directed evolution experiment. A selection of them is listed in Table 39.5. Choice of a method is critical as each one of them has certain advantages as well as limitations. Features such as biased mutational spectrum, use of hazardous chemicals, uneven sampling of codon space, and need for prior structural, biochemical, or phylogenetic knowledge of each method become important considerations. Despite these diverse methods to generate gene diversity, it is impossible to cover the entire mutational space available for a typical protein. For a polypeptide chain of "n" amino acids, $20^n$ combinations are possible as there are 20 different amino acids. Even nature has sampled a tiny fraction of these possible sequences over the huge evolutionary timescale. It is therefore best to begin with an existing enzyme

sequence. The choice of the sequence diversification method will then depend on the nature of evolutionary trajectory required from the initial scaffold to the desired end point. For instance, directed evolution of enzyme robustness (for stability) may require simultaneous changes scattered across the sequence length.

There are two broad approaches to finding the best mutant enzyme, namely, screening and selection. Screening of the mutant library can be performed in two different ways: (i) facilitated screening wherein mutants are distinguished based on distinct phenotypes and (ii) random screening in which we pick mutants blindly. Selection is always preferred over screening for its higher efficiency. But selection requires a phenotypic functional link between the target gene and its encoding enzyme product that confers the selective advantage (such as better growth, etc.) (Percival Zhang et al. 2006).

Directed evolution has been successfully applied to achieve favorable changes in enzyme properties like stereo- and region-selectivity, expanding the substrate scope and/or activity, enzyme robustness, pH optimum, and promiscuity as catalysts in organic synthesis (Jeschek et al. 2016; Reetz 2016). Select examples of evolved enzymes reported in the literature are listed in Table 39.6. Industrial-scale biocatalysis applications of optimized enzymes have focused primarily on hydrolases, a few

**Table 39.6 Examples of directed enzyme evolution**

| Enzyme | Feature optimized |
|---|---|
| Monoamine oxidase | Deracemization of racemic amines |
| Ketoreductases (KREDs) | Chiral intermediates for pharmaceuticals |
| Laccase | Catalytic efficiency, neutral pH range |
| Glyphosate N-acetyltransferase | Catalytic efficiency for glyphosate resistance |
| Cephalosporin acylase | Catalytic efficiency toward adipyl-7-ADCA |
| ω-Aminotransferase | Specific activity and thermostability |
| Epoxide hydrolase | Enantioselectivity |
| Hydantoinase | Inverting enantioselectivity (L-Met process) |
| β-Lactamase | Antibiotic resistance against cefotaxime |
| Xylanase, phytase, and lipase | Thermostability |
| Endoglucanase | Thermal stability, alkaline pH range |
| Subtilisin | Stability in organic solvent |
| Halohydrin dehalogenase | Catalytic efficiency |
| Aldolases | Specificity, efficiency, thermostability |
| Non-ribosomal peptide synthase–polyketide synthase (NRPS-PKS) hybrid | Produce broad spectrum antibiotic (andrimid) |

ketoreductases (KREDs), transaminases, oxidative enzymes, aldolases, cofactor regeneration, and enzyme stability in organic solvents (Bornscheuer et al. 2012).

**De Novo Enzyme Design** Designing an enzyme to catalyze the reaction of one's choice is a grand challenge (Nanda 2008). This strategy obviously requires detailed knowledge of the protein structure, structural basis of biological catalysis, and computational tools for enzyme design. In principle, the following steps may be envisaged:

- An appropriate catalytic mechanism is chosen for the target reaction.
- The transition state for this reaction is described.
- An idealized active site to position the catalytic residues and maximize *TS* stabilization be modeled.
- An appropriate protein scaffold is chosen from the available library.
- It is optimized in silico to best accommodate the reaction transition state and catalytic residues.
- The candidate *theozyme* polypeptides are actually created/produced.
- They are tested for catalytic activity.
- Best candidate designer enzyme may be fine-tuned by further sculpting around the transition-state model.

Although the initial activities of de novo enzyme constructs are typically low, they can be substantially improved through directed evolution approaches. Computational enzyme design has emerged as a promising tool to generate custom-built biocatalysts. The Rosetta de novo enzyme design protocol may be used to tailor enzyme catalysts for a variety of chemical reactions. There are constant efforts to improve the reliability of the Rosetta design cycle (Richter et al. 2011). Both the pre- and post-design analysis of protein structures promises to play an increasingly important role here.

Genetic engineering is relatively easy and routinely accessible now. But to attempt rational enzyme engineering, a vastly improved understanding of protein structure–sequence relationship is required. Our appreciation of protein dynamics is still very limited, and this makes structure–function correlation very hard. Subtle changes in the active site geometry have tremendous unexpected consequences for enzyme function. Hence, rational de novo design of a reasonably efficient enzyme continues to elude us. Few pioneering and brave attempts to build designer enzymes have been made. Relevant examples listed in Table 39.7 highlight the successes and limitations encountered in the de novo design of enzymes.

As can be seen from the representative cases listed above, most designer enzymes have not reached the expected catalytic performance (Bar-Even et al. 2011). This is because of our limited mastery over proper protein folding, structure, stability, dynamics, and catalysis. Indeed, subtle changes in the active site geometry are enough to generate remarkable unpredicted consequences for enzyme function. Poor catalytic performance of the de novo designs comes with its own caveats.

**Table 39.7 Examples of de novo enzyme design**

| Designer enzyme | Comments on the design |
|---|---|
| Oxaloacetate decarboxylase (metal-free) | A rationally designed synthetic 14 amino acid residue cyclic peptide – Oxaldie; decarboxylates oxaloacetate via an imine intermediate on its Lys $NH_2$. Catalytic efficiency comparable with abzymes (Johnsson et al. 1993) |
| Dihydrodipicolinate synthase | The N-acetylneuraminate lyase scaffold was rationally redesigned to switch the activity toward dihydrodipicolinate synthase. The designed activity showed 19-fold increased specificity for the new substrate (Joerger et al. 2003) |
| β-Lactamase (from glyoxalase II) | Several loop grafting steps at the active site to achieve major switch in function. Evolved new activity on an existing glyoxalase II scaffold (Park et al. 2006) |
| Sesquiterpene synthase | A promiscuous sesquiterpene synthase scaffold was used to build seven novel terpene synthases, catalyzing the synthesis of different sesquiterpenes (Yoshikuni et al. 2006) |
| Retro-aldolase | Retro-aldol cleavage of a carbon–carbon bond. The mechanism involves enamine catalysis by lysine (via a Schiff base or imine intermediate) giving a catalytic proficiency of $10^4$, which is far from natural enzymes (Jiang et al. 2008) |
| Diels–Alderase (intermolecular) | Organic bimolecular reaction forming two carbon–carbon bonds and up to four new stereogenic centers in one step. Naturally occurring enzymes are not known for this reaction. The designed enzyme is 20 times better than the corresponding abzymes (see Table 39.2) (Siegel et al. 2010) |
| Triosephosphate isomerase | The Rosetta enzyme design protocol demonstrated for the triose phosphate isomerase reaction as an example (Richter et al. 2011) |
| Kemp eliminase | A well-studied organic model system for proton transfer from carbon. In Kemp elimination, the deprotonation of substrate (5-nitrobenzisoxazole) leads to electronic rearrangements that break the C−H and N−O bonds while forming a C≡N triple bond. The designer enzyme accelerates the elementary chemical reaction ($6 \times 10^8$ fold), nearly as efficient as natural enzyme like triose phosphate isomerase (Blomberg et al. 2013) |

Ensuring that the observed activity is really due to the designed enzyme, and not a contaminating activity from the expression host, is critical. Attempts to convert catalytically inert ribose-binding protein into an active triose phosphate isomerase ran into such difficulties and elicited the response – "It is a bush-league error not to purify your proteins well, especially in such work" (Hayden 2008).

De novo protein design allows us to explore the full sequence space. Computational methodology has progressed well for a wide range of structures to be designed from scratch and with atomic-level accuracy (Huang et al. 2016). However, obtaining more active catalysts will require improved control over substrate binding and better pre-organization of the active site. Modifying existing protein scaffolds through rational redesign has been a more fruitful option so far – where the emphasis is on finding what works rather than predicting what works (Khersonsky et al. 2006). The fine-tuning of engineered enzymes can only be fulfilled today by combinatorial

approaches. The marriage of rational design and directed evolution (combinatorial redesign!) seems to be the way to go at present.

## 39.6 Summing Up

Our knowledge on and databases of genome sequences, metabolic pathways, protein sequences and their three-dimensional folds, enzyme active sites, and chemical reactions is expanding very fast. Sophisticated computational methods are expected to rationalize this vast information and aid in predicting the changes required to alter one enzyme into another. Efficiently introducing a new enzymatic activity in a chosen protein scaffold may not be too far off in the future. Meanwhile, rational de novo design of an enzyme continues to be a grand challenge (Editorial 2009).

This final section is demanding and difficult to cover – the subject matter of research is very current, and many reviews and new developments are reported on a very frequent basis! Almost every chapter, especially the last part (Part V) of this book, becomes outdated in a short span of months. Therefore, the relevant literature after the year 2016 has received limited attention. It is remarkable that for a subject so much undervalued, and displaced away from the mainstream biology of today, very-high-quality research gets added continuously to the literature.

## References

### General

Editorial (2009) closing in on catalysis. Nat Chem Biol 5:515
Clarke CF, Allan CM (2015) Unexpected role for vitamin B2. Nature 522:427–428
Herschlag D, Natarajan A (2013) Fundamental challenges in mechanistic enzymology: progress toward understanding the rate enhancements of enzymes. Biochemistry 52:2050–2067
Ito T, Yokoyama S (2010) Two enzymes bound to one transfer RNA assume alternative conformations for consecutive reactions. Nature 467:612–616
Ohashi M, Liu F, Hai Y, Chen M, Tang M, Yang Z, Sato M, Watanabe K, Houk KN, Tang Y (2017) SAM-dependent enzyme-catalyzed pericyclic reactions in natural product biosynthesis. Nature 549:502–506
Ortega MA, Hao Y, Zhang Q, Walker MC, van der Donk WA, Nair SK (2015) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. Nature 517:509–512
Zhao Q, Wang M, Xu D, Zhang Q, Liu W (2015) Metabolic coupling of two small-molecule thiols programs the biosynthesis of lincomycin A. Nature 518:115–119

### Transition State Analysis and Computational Enzymology

Garcia-Viloca M, Gao J, Karplus M, Truhlar DG (2004) How enzymes work: analysis by modern rate theory and computer simulations. Science 303:186–195
Schramm VL (2013) Transition states, analogues, and drug development. ACS Chem Biol 8:71–81
Warshel A (2014) Multiscale modeling of biological functions: from enzymes to molecular machines (Nobel Lecture). Angew Chem Int Ed 53:10020–10031

## Single Molecule Enzymology

English BP, Min W, van Oijen AM, Lee KT, Luo G, Sun H, Cherayil BJ, Kou SC, Xie XS (2006) Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. Nat Chem Biol 2:87–94

Min W, English BP, Luo G, Cherayil BJ, Kou SC, Xie XS (2005) Fluctuating enzymes: lessons from single-molecule studies. Acc Chem Res 38:923–931

Smiley RD, Hammes GG (2006) Single molecule studies of enzyme mechanisms. Chem Rev 106:3080–3094

Walter NG (2006) Michaelis-Menten is dead, long live Michelis-Menten! Nat Chem Biol 2:66–67

## Structure-Function Dissection of Enzyme Catalysis

Atkins JF, Gesteland R (2002) The 22nd amino acid. Science 296:1409–1410

Bryan PN (2000) Protein engineering of subtilisin. Biochim Biophys Acta 1543:203–222

Changeux J-P (2013) 50 years of allosteric interactions: the twists and turns of the models. Nat Rev Mol Cell Biol 14:819–829

Clarke AR, Atkinson T, Holbrook JJ (1989) From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part I. Trends Biochem Sci 14:101–105 Part II, Trends Biochem Sci, 14:145-148 (1989)

Cuesta-Seijo JA, Borchert MS, Navarro-Poulsen J-C, Schnorr KM, Mortensen SB, Leggio LL (2011) Structure of a dimeric fungal α-type carbonic anhydrase. FEBS Lett 585:1042–1048

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

Dhalla AM, Li B, Alibhai ME, Yost KJ, Hemmingsen JM, Atkins WM, Schineller J, Villafranca JJ (1994) Regeneration of catalytic activity of glutamine synthetase mutants by chemical activation: exploration of the role of arginines 339 and 359 in activity. Protein Sci 3:476–481

Earnhardt JN, Wright SK, Qian M, Tu C, Laipis PJ, Viola RE, Silverman DN (1999) Introduction of histidine analogs leads to enhanced proton transfer in carbonic anhydrase V. Arch Biochem Biophys 361:264–270

Freire E (2000) Can allosteric regulation be predicted from structure? Proc Natl Acad Sci U S A 97:11680–11682

Gerlt JA (2017) Genomic enzymology: web tools for leveraging protein family sequence-function space and genome context to discover novel functions. Biochemistry 56:4293–4308

Gerlt JA, Allen KN, Almo SC, Armstrong RN, Babbitt PC, Cronan JE, Dunaway-Mariano D, Imker HJ, Jacobson MP, Minor W, C., Poulter D, Raushel FM, Sali A, Shoichet BK, Sweedler JV (2011) The enzyme function initiative. Biochemistry 50:9950–9962

Hai Y, Kerkhoven EJ, Barrett MP, Christianson DW (2015) Crystal structure of an arginase-like protein from *Trypanosoma brucei* that evolved without a binuclear manganese cluster. Biochemistry 54:458–471

Hardy JA, Wells JA (2004) Searching for new allosteric sites in enzymes. Curr Opin Struct Biol 14:706–715

Hendrickson TL, de Crecy-Lagard V, Schimmel P (2004) Incorporation of nonnatural amino acids into proteins. Annu Rev Biochem 73:147–176

Kuznetsova E, Proudfoot M, Sanders SA, Reinking J, Savchenko A, Arrowsmith CH, Edwards AM, Yakunin AF (2005) Enzyme genomics: application of general enzymatic screens to discover new enzymes. FEMS Microbiol Rev 29:263–279

Larion M, Salinas RK, Bruschweiler-Li L, Miller BG, Bruschweiler R (2012) Order–disorder transitions govern kinetic cooperativity and allostery of monomeric human glucokinase. PLoS Biol 10(12):e1001452

Lopez V, Alarcon R, Orellana MS, Enrıquez P, Uribe E, Martınez J, Carvajal N (2005) Insights into the interaction of human arginase II with substrate and manganese ions by site-directed mutagenesis and kinetic studies: alteration of substrate specificity by replacement of Asn149 with Asp. FEBS J 272:4540–4548
Miller C (2007) Pretty structures, but what about the data? Science 315:459
Moon H-J, Tiwari MK, Singh R, Kang YC, Lee J-K (2012) Molecular determinants of the cofactor specificity of ribitol dehydrogenase, a short-chain dehydrogenase/reductase. Appl Environ Microbiol 78:3079–3086
Motlagh HN, Wrabl JO, Li J, Hilser VJ (2014) The ensemble nature of allostery. Nature 508:331–339
Nickbarg EB, Davenport RC, Petsko GA, Knowles JR (1988) Triosephosphate isomerase: removal of a putatively electrophilic histidine residue results in a subtle change in catalytic mechanism. Biochemistry 27:5948–5960
Ragsdale SW (2011) How two amino acids become one. Nature 471:583–584
Ramanathan A, Agarwal PK (2011) Evolutionarily conserved linkage between Enzyme Fold, Flexibility, and Catalysis. PLoS Biol 9:e1001193
Reynolds KA, McLaughlin RN, Ranganathan R (2011) Hot spots for allosteric regulation on protein surfaces. Cell 147:1564–1575
Wagner CR, Benkovic SJ (1990) Site directed mutagenesis: a tool for enzyme mechanism dissection. Trends Biotechnol 8:263–270
Wilks HM, Hart KW, Feeney R, Dunn CR, Muirhead H, Chia WN, Barstow DA, Atkinson AR, Holbrook JJ (1988) A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. Science 242:1541–1544
Wlodawer A, Ericson JW (1993) Structure-based inhibitors of HIV-1 protease. Annu Rev Biochem 62:543–586

## Designing Novel Catalysts

Abelson J (2017) The discovery of catalytic RNA. Nat Rev Mol Cell Biol 18(11):653–653
Baum DA, Silverman SK (2008) Deoxyribozymes: useful DNA catalysts in vitro and in vivo. Cell Mol Life Sci 65:2156–2174
Benkovic SJ (1992) Catalytic antibodies. Annu Rev Biochem 61:29–54
Bjerre J, Rousseau C, Marinescu L, Bols M (2008) Artificial enzymes, “chemzymes”: current state and perspectives. Appl Microbiol Biotechnol 81:1–11
Breslow R (2005) Artificial enzymes. Wiley-VCH Wienheim
Doudna JA, Lorsch JR (2005) Ribozyme catalysis: not different, just worse. Nat Struct Mol Biol 12:395–402
Famulok M (2004) RNAs turn on in tandem. Science 306:233–234
Hilvert D (2000) Critical analysis of antibody catalysis. Annu Rev Biochem 69:751–793
MacMillan DWC (2008) The advent and development of organocatalysis. Nature 455:304–308
Narlikar G, Herschlag D (1997) Mechanistic aspects of enzyme catalysis: lessons from comparison of RNA and protein enzymes. Annu Rev Biochem 66:19–59
Penning TM, Jez JM (2001) Enzyme redesign. Chem Rev 101:3027–3046
Reitman ZJ, Choi BD, Spasojevic I, Bigner DD, Sampson JH, Yan H (2012) Enzyme redesign guided by cancer-derived IDH1 mutations. Nat Chem Biol 8:887–889
Ren X, Gao S, You D, Huang H, Liu Z, Mu Y, Liu J, Zhang Y, Yan G, Luo G, Yang T, Shen J (2001) Cloning and expression of a single-chain catalytic antibody that acts as a glutathione peroxidase mimic with high catalytic efficiency. Biochem J 359:369–374
Schmeing TM, Ramakrishnan V (2009) What recent ribosome structures have revealed about the mechanism of translation. Nature 461:1234–1242
Sharkey MA, Engel PC (2009) Modular coenzyme specificity; a domain swapped chimera of glutamate dehydrogenase. Proteins 77:268–278

Taylor AI, Pinheiro VB, Smola MJ, Morgunov AS, Peak-Chew S, Cozens C, Weeks KM, Herdewijn P, Holliger P (2015) Catalysts from synthetic genetic polymers. Nature 518:427-430
Traut TW (2007) Allosteric Regulatory Enzymes. Springer Science & Business Media, Boston
Tuerk C, Gold L (1990) Systemic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510
Weigand JE, Suess B (2009) Aptamers and riboswitches: perspectives in biotechnology. Appl Microbiol Biotechnol 85:229–236
Wilson TJ, Lilley DMJ (2009) The evolution of ribozyme chemistry. Science 323:1436–1438
Wulff G (2002) Enzyme-like catalysis by molecularly imprinted polymers. Chem Rev 102:1–28
Zuckermann RN, Schultz PG (1988) Hybrid sequence-selective ribonuclease S. J Am Chem Soc 110:6592–6594

## Enzymes Made to Order

Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50:4402–4410
Blomberg R, Kries H, Pinkas DM, Mittl PRE, Grutter MG, Privett HK, Mayo SL, Hilvert D (2013) Precision is essential for efficient catalysis in an evolved Kemp eliminase. Nature 503:418–421
Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K (2012) Engineering the third wave of biocatalysis. Nature 485:185–194
Hayden EC (2008) Designer Debacle – News Feature. Nature 453:275–278
Huang P-S, Boyken SE, Baker D (2016) The coming of age of de novo protein design. Nature 537:320–327
Jeschek M, Reuter R, Heinisch T, Trindler C, Klehr J, Panke S, Ward TR (2016) Directed evolution of artificial metalloenzymes for in vivo metathesis. Nature 537:661–665
Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF III, Hilvert D, Houk KN, Stoddard BL, Baker D (2008) De novo computational design of retro-aldol enzymes. Science 319:1387–1391
Joerger AC, Mayer S, Fersht AR (2003) Mimicking natural evolution in vitro: an N-acetylneuraminate lyase mutant with an increased dihydrodipicolinate synthase activity. Proc Natl Acad Sci U S A 100:5694–5699
Johnsson K, Allemann RK, Widmer H, Benner SA (1993) Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. Nature 365:530–532
Khersonsky O, Roodveldt C, Tawfik DS (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. Curr Opin Chem Biol 10:498–508
Menger FM, Ladika M (1987) Origin of rate accelerations in an enzyme model: the p-nitrophenyl ester syndrome. J Am Chem Soc 109:3145–3146
Nanda V (2008) Do-it-yourself enzymes. Nat Chem Biol 4:273–275
Packer MS, Liu DR (2015) Methods for the directed evolution of proteins. Nat Rev Genet 16:379–394
Park H-S, Nam S-H, Lee JK, Yoon CN, Mannervik B, Benkovic SJ, Kim H-S (2006) Design and evolution of new catalytic activity with an existing protein scaffold. Science 311:535–538
Percival Zhang Y-H, Himmel ME, Mielenz JR (2006) Outlook for cellulase improvement: screening and selection strategies. Biotechnol Adv 24:452–481
Reetz M (2016) Directed evolution of selective enzymes: catalysts for organic chemistry and biotechnology. Wiley, Weinheim
Richter F, Leaver-Fay A, Khare SD, Bjelic S, Baker D (2011) De novo enzyme design using Rosetta3. PLoS One 6(5):e19230

Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, St JL, Clair JLG, Hilvert D, Gelb MH, Stoddard BL, Houk KN, Michael FE, Baker D (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. Science 329:309–313

Yoshikuni Y, Ferrin TE, Keasling JD (2006) Designed divergent evolution of enzyme function. Nature 440:1078–1082