Biocatalysis: A Status Report

Andreas S. Bommarius

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, Georgia 30318; email: andreas.bommarius@chbe.gatech.edu

Annu. Rev. Chem. Biomol. Eng. 2015. 6:319–45

The *Annual Review of Chemical and Biomolecular Engineering* is online at chembioeng.annualreviews.org

This article's doi:
10.1146/annurev-chembioeng-061114-123415

Keywords

biocatalyst, enzyme, protein engineering, stabilization

Abstract

This review describes the status of the fields of biocatalysts and enzymes, as well as existing drawbacks, and recent advances in the areas deemed to represent drawbacks. Although biocatalysts are often highly active and extremely selective, there are still drawbacks associated with biocatalysis as a generally applicable technique: the lack of designability of biocatalysts; their limits of stability; and the insufficient number of well-characterized, ready-to-use biocatalysts.

There has been significant progress on the following fronts: (*a*) novel protein engineering tools, both experimental and computational, have significantly enhanced the toolbox for biocatalyst development. (*b*) The deactivation of biocatalysts under various stresses can be described quantitatively via rational models. There are several cases of spectacular leaps of stabilization after accumulating all stabilizing mutations found in earlier rounds. The concept that stabilization against one type of stress commonly also stabilizes against other types of stress is now experimentally considerably better founded than a few years ago. (*c*) A host of developments of novel biocatalysts in the past few years, in part fueled by improved designability and improved methods of stabilization, has considerably broadened the toolbox for synthetic chemistry.

INTRODUCTION AND SCOPE

This review describes the status of the fields of biocatalysts and enzymes, as well as existing drawbacks, and recent advances in the areas deemed to represent drawbacks. Here, biocatalysts are understood to comprise any protein of a length that allows three-dimensional folding, known from statistical mechanics to start at approximately 20 amino acids in length. Therefore, RNA catalysts, cofactor-only catalysts, and other organocatalysts are not covered by this review.

The rapid spread of biocatalytic routes and targets in the past few years has been aided by advances in protein engineering as well as the increasing importance of both short, selective routes to pharmaceuticals and the synthesis of chemicals from renewables. Most of these syntheses are based on the processing of carbohydrates, lignocellulosics, and vegetable oils, rather than coal, oil, or natural gas. Several eminent reviews of the field have been published; as examples, the reader is referred to works by Turner (1), Bornscheuer et al. (2), and Reetz (3).

OVERVIEW OF FIELD

Primer on Proteins

Proteins are a class of macromolecules consisting of one or more folded amino acid chains of overall three-dimensional structure. Proteins consist of one or more linear chains of, with rare exceptions, the common 20 proteinogenic amino acids. The fairly frequent disulfide bonds between cysteines are formed through crosslinking of side-chain interactions but do not constitute branching of the backbone chain. Extremely rarely, an additional genetically coded twenty-first amino acid is found, such as selenocysteine in formate dehydrogenase or glutathione peroxidase (4, 5), or pyrrolysine, the twenty-second amino acid, in methanogens (belonging to the domain of Archaea) (6, 7). This review does not cover the body of excellent work performed to incorporate nonproteinogenic and often even unnatural amino acids into proteins [for reviews, see Ngo & Tirrell (8) and Xiao et al. (9)].

Sequences. The sequence variety of proteins is large, 20^N possible combinations, but not infinite. Monomers of enzymes typically consist of 147 (bovine lysozyme) to 667 (alcohol oxidase from *Aspergillus* or *Pichia*) amino acid residues, most of them clustering between 280 and 400 residues. Protein sequences are searchable with a variety of algorithm-based tools, such as BLAST (**http://blast.ncbi.nlm.nih.gov/Blast.cgi**), ClustalW (**http://www.clustal.org**), or FASTA (**http://blast.ncbi.nlm.nih.gov/blastcgihelp.shtml**), in a variety of databases, such as NCBI (National Center for Biotechnology Information, **http://www.ncbi.nlm.nih.gov**) maintained by the National Institutes of Health (NIH), or ExPaSy (**http://www.expasy.org**), maintained by the EBI (European Bioinformatics Institute, **http://www.ebi.ac.uk**). Recent access to GenBank (January 5, 2015) revealed 184,938,063,614 stored bases. Most sequences deposited in data banks hail from genome sequencing projects and are not actually functionally verified but annotated, i.e., computationally compared to sequences with known function and thus surmised to have the same function as the homologs.

Structures. Despite the huge number of protein sequences, proteins only fold into comparatively few three-dimensional structures, approximately 1,300 folds (**http://www.proteinstructures.com/Structure/Structure/protein-fold.html**). Protein structures are deposited in databases such as the Protein Data Bank (PDB) of the NIH (**http://www.rcsb.org/pdb/home/home.do**). As of January 5, 2015, there were 105,499 structures deposited in the PDB, among them 34,987 distinct proteins and 7,454 nucleic acid–containing structures. With regard to the structures,

88.9% were obtained via X-ray crystallography, 10.2% through solution NMR spectroscopy, and 0.9% via electron microscopy. Until recently, solution NMR spectroscopy was only able to analyze small (<15–25 kDa) proteins; however, new NMR techniques, in combination with new labeling strategies, now allow solving NMR structures of proteins up to a size of 1 MDa (10).

Protein stability. Stabilizing forces such as hydrophobic effects or hydrogen bonds only slightly dominate destabilizing forces such as Coulombic forces of equal polarity, thus the Gibbs free enthalpy of formation of proteins, $\Delta G_{formation}$, is only weakly negative.

Primer on Biocatalysts

Biocatalysis is achieved by monomeric or multimeric enzymes. Enzymes are a class of macromolecules consisting of one or more folded amino acid chain of overall three-dimensional structure, with the ability to bind small molecules in the substrate binding pocket and to effect reaction through catalysis in the adjacent active site.

Biocatalysts are subject to the same rules of physics and chemistry as other catalysts. Therefore, it is to be expected that a biologically based catalyst exists for basically any chemical reaction. However, as recently as ten years ago, there was a long list of reactions with no known biocatalyst. During the past decade, however, several new biocatalysts were found and/or developed for reactions with no known biocatalyst (for examples, see **Table 1**).

Enzymes do differ from other catalysts in their ability to exclude water from the active site and thus access unusual redox potentials or apparent pKa values. Examples for unusual pKa values, especially of carboxyl or amine groups in proteins, include 6.5 for Glu-35 in lysozyme, 5.9 for the ε-amino group of Lys in acetoacetate decarboxylase, or 3.4 for His-159 in papain (16).

Enzyme kinetics. The kinetics of a one-substrate enzyme reaction typically can be described by the Michaelis-Menten (MM) equation:

$$v = v_{max}[S]/(K_M + [S]) = k_{cat}[E]_{total}[S]/(K_M + [S]) \quad 1.$$

The underlying mechanism assumes three possible elementary reaction steps: (*a*) formation of an enzyme-substrate complex (ES-complex), (*b*) dissociation of the ES-complex into E and S, and (*c*) irreversible reaction to product (P). In this scheme, the product formation step from ES to E + P is assumed to be rate-limiting, thus the ES-complex is modeled to react directly to the free enzyme and the product molecule, which is assumed to dissociate from the enzyme without the

Table 1 Recently developed biocatalysts for reactions with no known prior example in nature

Enzyme	Reaction catalyzed	Method developed	Starting enzyme template	Reference
Diels-Alderase	Diels-Alder condensation	Findings in cell-free extracts	Solanapyrone synthase	(11)
Michael additase	Morita-Baylis-Hillman reaction	Substrate promiscuity	BSA, PLE	(12)
Kemp eliminase	Kemp elimination	Initial target: computation of TS; optimization via DE	None	(13)
Amine dehydrogenase	Reductive amination	Initial target: point mutations; optimization via CASTing/ISM	Amino acid DH	(14)
Cyclopropanase	Cyclopropanation		Cyt P450 oxidase	(15)

Abbreviations: BSA, bovine serum albumin; DE, directed evolution; PLE, porcine liver esterase; TS, transition state.

formation of an enzyme-product (EP-) complex:

$$E + S \Leftrightarrow ES \rightarrow E + P \qquad 2.$$

In Equation 1, the reaction rate traditionally is denoted by v [concentration/time] and k_{cat} is the reaction rate constant [time^{-1}]. The equation describes a two-parameter kinetics, with a monotonically rising reaction rate with respect to substrate concentration and saturation at high substrate concentration. The maximum reaction rate at saturation is denoted by v_{max}, with $v_{max} = k_{cat} \cdot [E]$. The K_M value corresponds to the substrate concentration at half saturation ($v_{max}/2$) and is a measure of binding affinity of the substrate to the enzyme: a high K_M value corresponds to loose binding, a low value to tight binding between enzyme and substrate. At low substrate concentration, more precisely, if $[S] \ll K_M$, Equation 1 simplifies to $v = v_{max}[S]/K_M$ and consequently is first order with respect to the substrate concentration [S]. In contrast, at high substrate concentration, more precisely, if $[S] \gg K_M$, Equation 1 simplifies to $v = v_{max}$ and consequently is zeroth order with respect to [S]. In all situations, v is proportional (first-order with respect) to [E].

Acceleration and efficiency. Enzymes have been found to accelerate uncatalyzed reactions very effectively: the most effective biocatalyst, orotate decarboxylase, accelerates the uncatalyzed background reaction 10^{17}-fold, so the acceleration ratio k_{cat}/k_{uncat} is found to be $\leq 1.4 \cdot 10^{17}$ (17). The proficiency of biocatalysts, the ratio $k_{cat}/(k_{uncat} \cdot K_M)$ $(= 1/K_T)$, covers the range of $\leq 10^{23}$ M^{-1} (18). Surprisingly, for a range of enzymatic reactions k_{cat} is within two orders of magnitude, whereas k_{uncat} varies by more than six orders of magnitude; the most effective enzymes were found to accelerate reaction with the slowest uncatalyzed reactions (19).

Strengths of Biocatalysis

The foremost strength of biocatalysts is their often superb specificity: enzymes often far surpass other catalysts with respect to chemical specificity; regioselectivity, such as toluene monooxygenase (20), hydroxyaminobenzene (HAB) mutase (21), or cytochrome P450 BM-3 (22); or enantioselectivity, for which there are legions of examples, such as N-acyl amino acid hydrolase (acylase) (23), lipase (24), epoxide hydrolase (25), amino acid dehydrogenase (26), or amine dehydrogenase (14, 27).

Biocatalysts commonly operate at nearly ambient conditions with respect to temperature and pH value in mostly aqueous solution. Operation at near ambient conditions promises energy efficiency of biocatalytic processes. Typical renewable raw materials, such as carbohydrates or fatty acids, are more hydrophilic than raw materials obtained from petrochemicals, which tend to be hydrocarbons. Such hydrophilic raw materials are advantageously processed in aqueous solvents, so biocatalysis is a good technological fit for renewable raw materials.

Weaknesses of Biocatalysis

Although biocatalysts are often highly active and extremely selective, there are still drawbacks associated with biocatalysis as a generally applicable technique. This review focuses on three:

1. The development of biocatalysts is prone to chance and slow, and does not follow a set of rules. The development of enzymes from nature and improvement through random mutagenesis are perceived this way.
2. A decisive weakness of biocatalysts is their very limited range of stability with respect to temperature, solvents, pH value, ionic strength, and salt type.
3. A continuing weakness of biocatalysis as a problem solution is the small number of well-characterized enzymes in comparison with chemical catalysts. In addition, not many are commercially available and thus hinder the diffusion of this technology.

In the following sections, we address progress associated with each of these three points. Afterward, we discuss a few success stories of biocatalysts.

INSUFFICIENT DESIGNABILITY OF PROTEINS: PROGRESS IN PROTEIN ENGINEERING

The arguably most important reason for the progress of biocatalysis over the past decade is the advance of protein engineering. Progress can be divided into three waves:

1. Rational protein design, the first wave, was developed a few years after the discovery of recombinant DNA technology and flourished in the decade after about 1985. Applying site-directed protein engineering, the community expected to design enzymes with predictable properties (this expectation turned out to be way premature). Even then, site saturation was part of the repertoire, which intellectually can already be attributed to
2. combinatorial protein engineering, the second wave and often also called directed evolution, with its peak between around 1995–2005. Two techniques dominate combinatorial protein design, random mutagenesis and recombination. Random mutagenesis most often operates with a high error rate of DNA duplication and thus introduces mistakes that lead to mutations anywhere in the gene sequence. Recombination digests the DNA into small pieces that are recombined in random fashion; appropriately, its most important protocol is called DNA shuffling. Combinatorial protein engineering is able to access a vastly wider sequence space than rational design but requires increasingly large libraries of mutants, with still low and uncertain chances of a successful outcome, owing to the vast sequence space of proteins. Both the requirement for high-throughput screening equipment and the uncertain hit rates led the community to search for more controllable techniques and protocols, and so
3. data-driven protein engineering, the third wave of protein engineering, sought to improve the hit rate while decreasing library size. Data-driven protein engineering, often also termed semirational protein engineering, routinely utilizes knowledge gained from rational design, such as critical residues from mechanistic studies or inspection of crystal structures, and combinatorial design, such as hot spots found during directed evolution. But it is because of three tools developed over the past 10 years that protein engineering has advanced so much: (*a*) restricted libraries; (*b*) combinatorial active-site saturation testing, or CASTing; and (*c*) iterative site mutagenesis (ISM).
    a. Restricted libraries: The full genetic code operates at the DNA level with four nucleotides [the purines adenine (A) and guanine (G) and the pyrimidines cytosine (C) and thymine (T)] (**Figure 1**). As a triplet codes for an amino acid, there exist $4^3 = 64$ triplet combinations of the nucleotides. Although some amino acids are only coded for by one triplet, such as ATG for methionine (Met, M) or TGG for tryptophan (Trp, W), most amino acids are coded for by several closely related triplets, such as TCN for serine (Ser, S) or CCN for proline (Pro, P) (here, N stands for any nucleotide: A, C, G, or T). Restricted libraries do not offer the full 64 triplet combinations, through restricting the number of nucleotides to fewer than four in one, two, or all three of the triplet positions (28). Although such a restriction very often leads to coding for fewer than the full 20 proteinogenic amino acids, the residues not coded for might not be deemed important for the problem at hand to be solved; if more space in a substrate binding pocket is desired, the inclusion of Trp, the largest of the proteinogenic amino acids, seems less important. However, the number of colonies at constant coverage (such as 95%) required to be screened for an outcome falls dramatically, which enables much faster and more reliable screening (**Figure 2**).

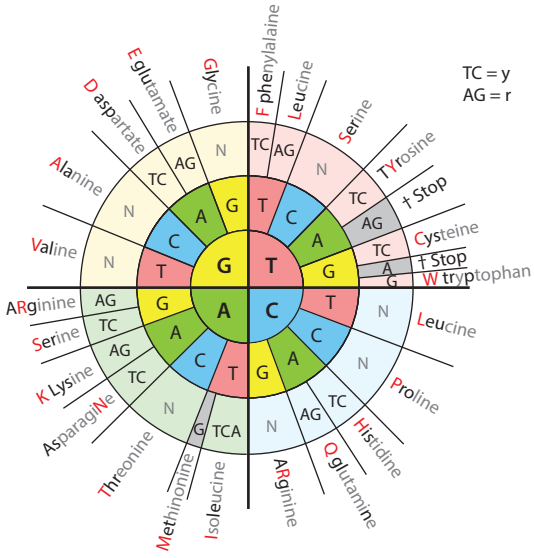


Figure 1

Genetic code.

b. Combinatorial active-site saturation testing, or CASTing, restricts the number of amino acid residues targeted for mutation and then applies a random mutation or a saturation protocol across the targeted sites (29a). Deciding to which residues to restrict the targeting requires information from crystal structures, mechanistic knowledge, sequence matching, or results of prior experiments. Setting coverage at ≥95%, saturation screening of up to two amino acid residues is feasible with standard screening procedures and equipment. With restricted libraries, three and sometimes even four amino acid residues are accessible (**Figure 3**).

c. In almost all cases, more promising amino acid residues are found in an enzyme than can be targeted at a single site in a single round of CASTing. ISM carries out successive CASTing across several sites in a protein (**Figure 4**) (30). Whereas different pathways, say ABCD versus BCDA, often result in intermittent variants with different strength of function, say ABC versus BCD, a variant incorporating a certain number of mutations has properties that are independent of the pathway traveled to arrive at that variant. ISM is very strong, however, at picking up synergistic coupled mutations, termed epistatic interactions.

LACK OF STABILITY OF ENZYMES: RECENT SUCCESSES IN DESCRIPTION OF BIOCATALYST PROCESS STABILITY AND THE STABILIZATION OF BIOCATALYSTS

Process-Relevant Parameters for Enzymes

Several different catalytic technologies (such as homogeneous, heterogeneous, or biological catalysis) dominate their niche of application, as conditions of operation, such as temperature and solvent, are too different across different technologies. Such different operating regimes

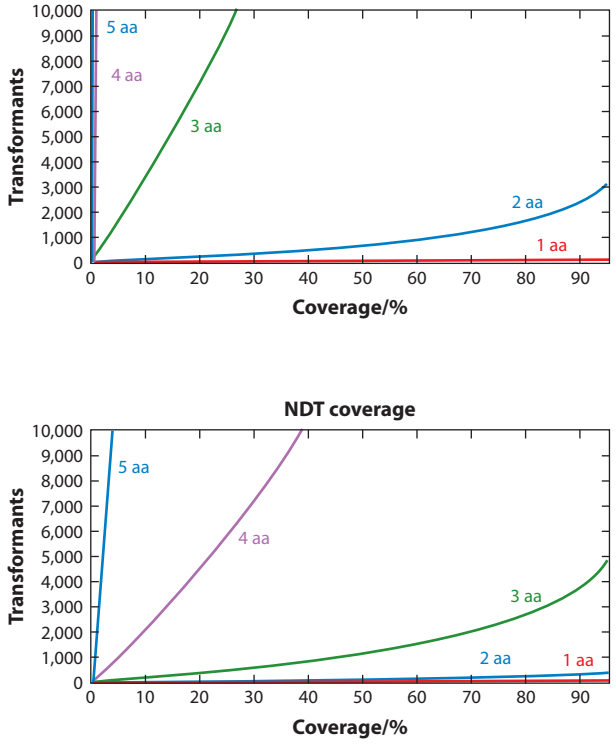


Figure 2

Number of transformants required for level of coverage for different restricted nucleotide libraries. Reproduced from Reetz et al. (28); copyright 2008 Wiley-VCH.

were accepted by process developers, operators, and regulators. As a consequence, most catalytic processes did not have much competition from different catalytic technologies. One important reason for the lack of comparable studies across different catalytic technologies is the absence of accepted catalytic performance criteria.

- Activity in connection with high concentration leads to volumetric productivity, expressed as space-time-yield.
- In contrast to thermodynamic stability of an enzyme, characterized by the melting temperature T_m at the transition between native and unfolded protein state, process or

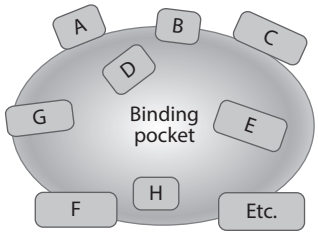


Figure 3

Principle of combinatorial active-site saturation testing, or CASTing. Reproduced from Reference 29b; copyright 2011 Wiley-VCH.

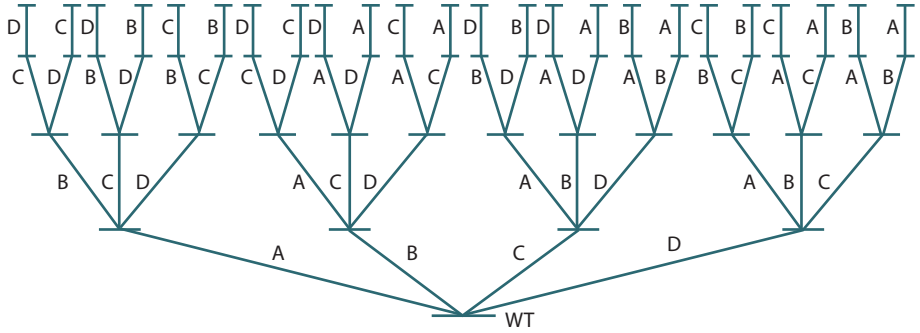


Figure 4

Iterative saturation mutagenesis (ISM). Reproduced from Reference 25, copyright 2006 Wiley-VCH.

operational stability of a biocatalyst follows the kinetic stability over time. High process stability maximizes activity over the longest possible lifetime and is expressed by the total turnover number *TTN*, the average turnover per active site over its lifetime.

- Selectivity of a reaction or process can be measured by the enantiomeric or diastereomeric excess of the resulting product (e.e. or d.e. value).

Process efficiency of biocatalysts. Data on volumetric productivities of processes are infrequent, as they often are regarded as confidential; the best collections of productivity data can be found in References 31–33. Space-time-yields can be calculated as the quotient of final product concentration and residence time τ in the reactor ($[P]/\tau$) or the reaction rate r [= moles of product $m_{Product}$ over the residence time τ) per unit of reactor volume V (r/V) or simply the moles of product per unit time and reactor volume ($m_{Product}/(V\cdot\tau)$]. Thus, space-time-yields are influenced by the solubility of the substrate $[S]_{max}$ (via the attainable reaction rate) and product $[P]_{max}$.

Rigorous modeling of total turnover for biocatalysts. Unlike most other types of catalysts, the TTN of biocatalysts can be linked to more accessible observables through simple but rigorous models of unfolding and deactivation (see Rigorous Determination of TTN for Enzymes, sidebar). Presuming the validity of the Lumry-Eyring model of deactivation (34), i.e., first-order irreversible deactivation from the unfolded state U to the chemically deactivated state D, the observed rate constant of deactivation $k_{d,obs}$ can be linked to the intrinsic rate constant of deactivation $k_{d,intr}$ or

RIGOROUS DETERMINATION OF TTN FOR ENZYMES

$$N \overset{K}{\longleftrightarrow} U \overset{k}{\longrightarrow} D; \quad K = [N]/[U]$$

$$k_{d,obs} = k_d \cdot K/(1+K) \qquad 3.$$

$$TTN = k_{cat,obs}/k_{d,obs} \qquad 4.$$

k_d, which differs from the former if part or all of the native conformation N is unfolded to U or deactivated to D. If, in addition, saturation kinetics can be assumed ($[S] \gg K_M$), the TTN can be expressed as the ratio of the catalytic rate constant k_{cat} scaled by the fraction of the remaining active enzyme from $[N]/[N]_0$ ($k_{cat} \cdot [N]/[N]_0 = k_{cat,obs}$) to the observed rate constant of deactivation $k_{d,obs}$ at the same temperature, pH value, ionic strength, and solution composition (35).

Process-relevant biocatalyst activity. What is the most relevant criterion to gauge process-relevant activity of enzymes? Enzyme specificity k_{cat}/K_M, favored by biochemists, has been argued to be misleading at substrate concentrations $[S]_0 > K_M$, as is relevant for most industrial applications, where the substrate concentration is very high, often >1 M (36). Instead, an efficiency function E_f has been proposed, which scales the catalytic constant k_{cat} to the product of the second-order rate constant of diffusion k_{diff} [$\approx 3.2 \cdot 10^9$ ($M \cdot s^{-1}$) in water at 25°C] and a relevant concentration, here the term $K_M + [S]$ (37). In expanded form to accommodate the case of reversible enzyme reactions, the equation for E_f reads:

$$E_f = k_{cat}(1 + 1/K)/\left\{k_{diff}(K_M + [S] + g(k_{cat,f}/k_{cat,r})[S]\right\}, \quad 5.$$

where K is the equilibrium constant; $k_{cat,f}$ and $k_{cat,r}$ are the rate constant for the forward and reverse reaction, respectively; and $g = 1$ for reversible and 0 for irreversible reactions. E_f is useful for comparing different variants of the same enzyme on the same substrate at the same pH value and temperature.

For applications in synthesis, the time to reach a certain, often very high, degree of conversion (such as 99%) is crucial. Several common measures of enzyme performance are all inadequate in that regard. Specificity, k_{cat}/K_M, does not capture conditions at substrate saturation, i.e., $K_M >(>) [S]_0$, conditions that are very common in synthesis, as high starting concentration $[S]_0$ allows to make more product in one batch. The catalytic constant k_{cat} by itself does not take into account uncompetitive and noncompetitive product inhibitions, which both reduce the effective maximum rate constant. Both k_{cat}/K_M and k_{cat} commonly do not take into account strongly varying substrate concentration during synthesis runs, so different degrees of conversion result in very different residence times.

Recently, the average velocity v_{avg} over the course of the reaction to a specified final degree of conversion, or catalytic effectiveness, has been proposed (38). For irreversible enzyme reactions with competitive product inhibition (inhibition constant K_{IC}) at 99% conversion, so that $[P] \approx [S]_0$, v_{avg} is calculated to be

$$v_{avg} \approx [P]/t = v_{max}/(1 + K_M^{eff}/[S]_0), \quad 6.$$

with

$$K_M^{eff} = K_M(1 + 0.78[S]_0/K_{IC})(\ln(1/(1 - 0.99)) \quad 7.$$

or

$$K_M^{eff} = K_M(1 + 0.78[S]_0/K_{IC})4.61. \quad 8.$$

In cases where both v_{max} and K_M are lower in a variant enzyme, this variant often is termed to be adequate in performance if v_{max}/K_M is used as a criterion. However, the reaction employing the variant as a catalyst then will take longer to reach a specified high degree of conversion. Likewise, if the product inhibition constant K_{IC} differs between two enzymes or two variants of the same enzyme, estimates of performance based on initial rate measurements cannot be extrapolated to higher degrees of conversions and thus not to the prediction of average velocity v_{avg} or catalytic effectiveness.

Stabilization of Enzymes Through Protein Engineering

Xylanases are important in improving digestion of hemicelluloses from feed in farm animals. As animal feed is heat-sterilized before use, good thermal stability is key to successful utilization of xylanases. Thermal stabilization through the application of two different methods of combinatorial protein engineering, gene site saturation mutagenesis (GSSM) and exhaustive mutation among nine sites with a xylanase gene, Palackal et al. (39) improved the melting temperature T_m of xylanase from 61°C to as high as 96°C!

SCHEMA recombines modules of highly related proteins, where the modules have been designed with the help of crystal structures to dissect the wild-type enzyme at points of least disruption, most of them [but by no means all of them (40)] in loops (41). Frances Arnold (Caltech, Pasadena, CA) succeeded in thermally stabilizing various cellulases, important for utilizing cellulosic biomass for fuels and chemicals. The group not only found stabilized catalytic domain variants when recombining three related fungal cellobiohydrolases 2 (CBH2s) but also discerned additivity of the stabilizing mutations by treating the sequence-stability data via a linear regression algorithm (42). Almost coincidentally, the group also found a single, rather conservative, point mutation (C313S) that contributed markedly to stabilization (43). The approach of combining SCHEMA with linear regression treatment was extended to cellobiohydrolase 1 (CBH1), the stabilized chimeras of which hydrolyzed solid cellulose at 70°C while featuring almost a 40% hit rate among all chimeras investigated (44). Up-to-date protein engineering protocols for the stabilization of cellulases have been compiled recently (45).

The consensus method suggests that mutations toward the most frequently represented amino acid at that position in an alignment stabilizes the protein, as protein stability follows Boltzmann kinetics, which poses that the more frequently populated states are at lower energy. Application of the consensus method to protein stabilization was pioneered by Steipe et al. (46) on an IgG-type monoclonal antibody and by Lehmann et al. (47, 48) and Amin et al. (49) on enzymes. Application of the consensus method in combination with inspection of a crystal structure (structure-guided consensus method) to glucose dehydrogenase (GDH) from *Bacillus megaterium* led to a 10^6-fold increase of the half-life $\tau_{1/2}$ and thus a 10^6-fold decrease of the observed deactivation rate constant $k_{d,obs}$ at corresponding temperatures (50). The key residues for stabilization (global stabilizers), Glu170 (Lys or Arg) and Gln252Leu, had already been found via random mutagenesis (51).

The B-FIT method targets the most mobile residues from B-factors in a protein crystal structure, with the premise that freezing those residues will stabilize the whole protein. With the help of the B-FIT approach, Reetz et al. (52) found that the melting temperature T_m of a pentavariant of lipase A (LipA) from *Bacillus subtilis* was increased by 45°C, accompanied by improved reversibility. This stabilization was found to be correlated with, and likely caused by, the formation of a H-bonding network extending over almost the whole size of the protein, as well as additional charged residues (52). The structure-guided consensus approach and the B-FIT method were combined in the stabilization of amino ester hydrolase (AEH) from *Xanthomonas campestris* by 8°C for T_m (53) and 13-fold for $k_{d,obs}$.

Crystal structures in combination with a host of computational techniques, including molecular dynamics simulations, have been utilized successfully for biocatalyst stabilization. FRESCO (Framework for Rapid Enzyme Stabilization by Computational libraries; **Figure 5**) allowed 1,634 initially suggested point mutations to be reduced, after manual inspection of all last-round computationally suggested mutations, to 64 variants that were experimentally screened. Three variants of limonene epoxide hydrolase (LEH) from *Rhodococcus erythropolis* stood out: the best stabilized LEH by 35°C for a T_m of 85°C while increasing kinetic stability (i.e., decreasing $k_{d,obs}$) by a factor of 250 (54).

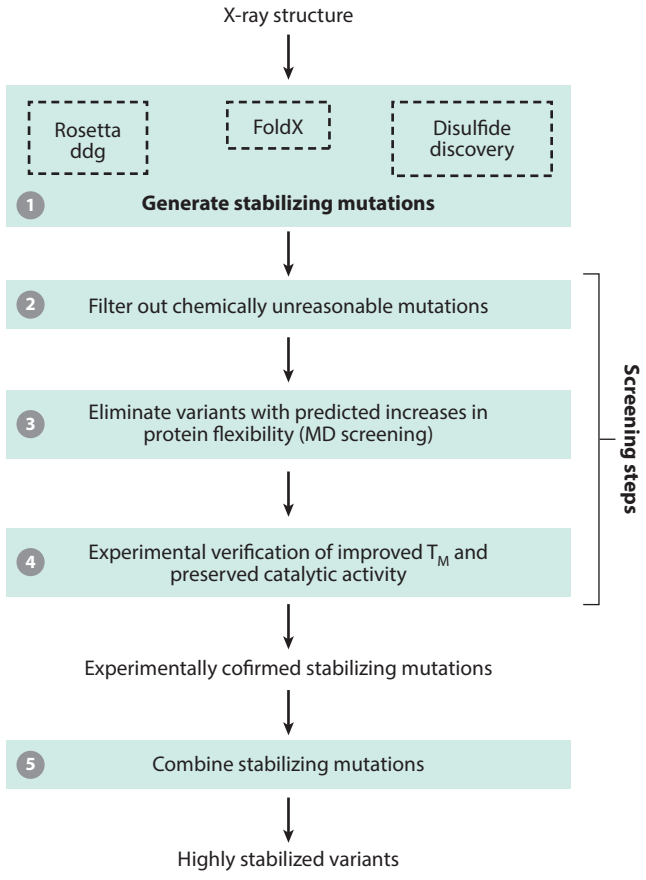


Figure 5

FRESCO (Framework for Rapid Enzyme Stabilization by Computational libraries). Reproduced from Wijma et al. (54).

Stabilization of Enzymes Against Several Stresses

Biocatalysts are sensitive against various stresses, such as high and low temperature, ionic strength, chaotropic salts, organic solvents, denaturants, and gas-liquid and solid-liquid interfaces. The following examples have been discussed by Bommarius & Paye (55) and are briefly summarized here.

Thermal and organic solvent stability. Naturally occurring homologs of several different thermophilic proteins demonstrated correlated thermal stability and tolerance to organic solvents (56, 57). A thermostable fructose bisphosphate aldolase found via directed evolution also featured increased stability in organic solvents (58).

Variants of *B. subtilis* lipase A, which after directed evolution based on the B-FIT method were found to display enhanced thermostability (see Stabilization of Enzymes Through Protein Engineering, above), also show significantly increased tolerance to hostile organic solvents (59).

Thermophilic Old Yellow Enzyme (TOYE) from *Thermoanaerobacter pseudethanolicus* E39 adapts to octamers and dodecamers in solution under stress, in contrast to a tetrameric state at ambient conditions. TOYE is stable at high temperatures ($T_m > 70°$) and simultaneously features

increased resistance to denaturation in water-miscible organic solvents compared to a closely related mesophilic Old Yellow Enzyme family member, pentaerythritol tetranitrate reductase (60).

Thermal, acid, and solvent stress. Encapsulation in a carboxymethylcellulose (CMC)-Ca alginate liquid membrane kept cysteine protease cathepsin B purified from goat brain >60% active during five batches. The maximum temperature T_{max} of the free and immobilized enzyme was 50°C and 55°C, respectively. Microencapsulation significantly stabilized the enzyme against pH, thermal, and solvent instability, opening up the potential for use in catalyzing transesterifications or transamidations (61).

Thermal, solvent, and interfacial stress. Formate dehydrogenase (FDH) from *Candida boidinii* has shown limiting stability in cofactor regeneration schemes with an amino acid dehydrogenase. Variants with the free Cys residues mutated to Ser or Ala, C23S and C23S/C262A, were found to stabilize FDH even in the absence of oxygen (62). The two changes also stabilized FDH against interfaces with nitrogen bubbles, shear stress in Couette viscometers, and shear stress and cavitation in gear pumps (63).

Thermostable glucose dehydrogenase (GDH) variants developed via an amino acid sequence-based consensus method also showed improved stability in solutions with high concentrations of either kosmotropic or chaotropic salts, as well as water-miscible organic solvents. Only the most stable variants showed little deactivation dependence on salt types along the Hofmeister series and on salt concentration. Kinetic stability, expressed by the deactivation rate constant $k_{d,obs}$, did not always correlate with thermodynamic stability of variants, as measured by melting temperature, T_m. However, a strong correlation ($R^2 > 0.95$) between temperature stability and organic solvent stability was found when plotting T_{50} (63) versus C_{50} (63) values (50).

INSUFFICIENT NUMBER OF BIOCATALYSTS: DEVELOPMENT OF NOVEL ENZYMES

Novel Biocatalysts Through Finding New Reactions

Finding novel biocatalysts through new reactions requires good knowledge of chemistry and usually relies on very good assays to analyze for an anticipated target product.

Sulfatases. Through follow-up of an observation that cyano- and thiobacteria from *Synechococcus* and *Paracoccus* species, respectively, furnished either (*R*)- or (*S*)-secondary alcohols from the corresponding racemic sulfate esters in an enantiocomplementary fashion, Kurt Faber's group at the University of Graz (Austria) found broad-spectrum inverting alkylsulfatases, for which no complement in chemistry exists (64). Medium engineering dramatically improved enantioselectivity: addition of lower alcohols or use of methyl tert-butyl ether in a biphasic system improved the E value from approximately 4 to >200. Inverting alkylsulfatases, which belong to the family of metallo-β-lactamases and feature a conserved binuclear Zn^{2+} cluster in their active site, catalyze the enantioselective hydrolysis of *sec*-alkylsulfates with strict inversion of configuration, yielding homochiral (*S*)-configurated alcohols and nonreacted (*S*)-configurated sulfate esters. The latter can be chemically hydrolyzed to (*S*)-secondary alcohols, so that a single (*S*)-configurated product with up to >99% e.e. results (**Figure 6**) (65).

Perfect enantioselectivities were obtained inverting alkylsulfatase from DSM 6611 for a variety of *sec*-alkyl sulfate ester substrates bearing olefinic, acetylenic, and aromatic moieties next to the sulfate ester group; in addition, chirally pure products with anti-Kazlauskas' configuration were found (66). Addition of DMSO as cosolvent increased insufficient enantioselectivity up to twofold. Recently, employing a mixture of inverting and retaining stereocomplementary sulfatases acting

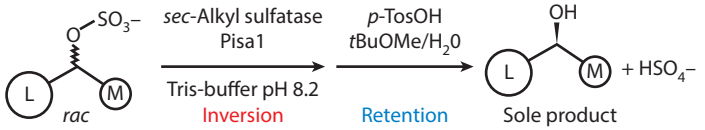


Figure 6

Enantiocomplementary synthesis of (*S*)-secondary alcohols. Reproduced from Schober et al. (65), copyright 2011 ACS.

on opposite substrate enantiomers enabled syntheses with a high degree of conversion (80 to >99%) and a high e.e. value (91 to >99% e.e.) (67).

Alkene oxidases (ozonidases). No enzyme was known to catalyze the equivalent of alkene ozonolysis to yield two smaller moieties of aldehydes or ketones until Wolfgang Kroutil's group in the University of Graz (Austria) obtained carbonyl compounds from phenyl-conjugated alkenes and molecular oxygen at 2 bar pressure, catalyzed by a cell-free extract of the wood-degrading fungus *Trametes hirsuta* FCC 047 (68). Ninety-six-well plates specially adapted for oxygen pressure revealed successful conversion of a broad range of aromatic alkenes to the respective aldehydes or ketones, with excellent chemoselectivity, i.e., no overreaction to carboxylic acids (**Figure 7**). Not only were various substituted styrenes cleaved but also cyclic alkenes, such as indene or 1,2-dihydronaphthalene (69). Mechanic investigations reveal that the new biocatalyst follows neither the standard dioxygenase nor the monooxygenase mechanism. It instead seems to prefer a mechanism akin to a 2 + 2 cycloaddition, as two oxygen atoms derived from two different molecules of dioxygen are incorporated into the carbonyl products (**Figure 8**) (70).

Very recently, this new biocatalyst was identified as a novel Mn^{3+}-containing proteinase A analog, in which Mn^{3+} is uniquely bound to oxygen ligands only (71).

Pictet-Spengler cyclisation: berberine bridge enzyme (BBE). Berberine-bridge enzyme (BBE) belongs to the class of bicovalent flavoproteins, discovered in 2005, that often feature unusually high redox potentials and thus can catalyze unusual reactions (72). Correct positioning of the flavin seems to depend on this bicovalent linkage. Recently, BBE was reported to catalyze the intramolecular oxidative deracemization of 1-benzyl-1,2,3,4-tetrahydroisoquinoline and molecular oxygen to (*S*)-berbine and (*R*)-1-benzyl-1,2,3,4-tetrahydroisoquinoline (73). Both substrates and products are nearly insoluble in aqueous solutions, but BBE was found to be compatible with a broad range of organic solvents. For the conversion of (*S*)-reticuline to (*S*)-scoulerine, a nearly 1:1 toluene/water biphasic system allowed, at 20 g/L substrate concentration and at conditions

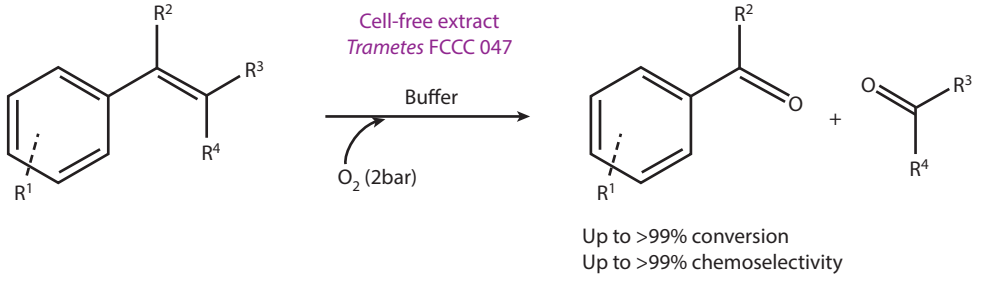


Figure 7

Reaction products of alkene double bond oxidation akin to ozonolysis. Reproduced from Reproduced from Lara et al. (68), copyright 2008 Wiley-VCH.

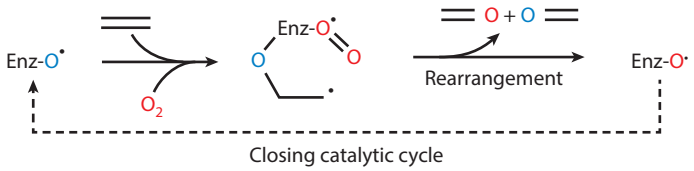


Figure 8

Presumed mechanism of alkene oxidase. Reproduced from Reference 69. Abbreviation: Enz, enzyme.

between pH 8 and 11 and 30–50°C, to achieve >97% e.e. for the product (74). For the deracemization of racemic benzylisoquinolines to optically pure (*S*)-berbines with up to 100% yield, a novel three-reaction cascade combining stereoinversion and simultaneous kinetic resolution was developed. The racemic substrates were transformed to optically pure (>97% e.e.) (*S*)-berbine products with nearly complete conversion (98%) and up to 88% isolated yield (75).

Novel Biocatalysts Derived Through Protein Engineering

Finding novel biocatalysts function through protein engineering requires a very stringent assay for the target molecule(s), often in high-throughput format, and a protocol for either gene mutation or gene recombination.

Monoamine oxidase N (MAO-N) for deracemization of racemic amines. About one third of all active pharmaceutical ingredients (APIs) contain chiral amine groups, so routes to enantiomerically pure amines are very important. Wild-type monoamine oxidase (MAO) catalyzes the oxidation of primary amines in the presence of molecular oxygen; however, Nicholas Turner's group from the University of Manchester (United Kingdom) has engineered MAO-N from *Aspergillus niger* to also catalyze the enantiospecific oxidation of secondary and even tertiary amines. Altogether, variants labeled D1 through D11 are available from 11 generations of protein engineering, including both rational design and high-throughput screening of random mutations. Generally, the subsequent generations resulted in increasing size of the substrate pocket, allowing conversion of increasingly more bulky amine substrates (**Figure 9**) (76).

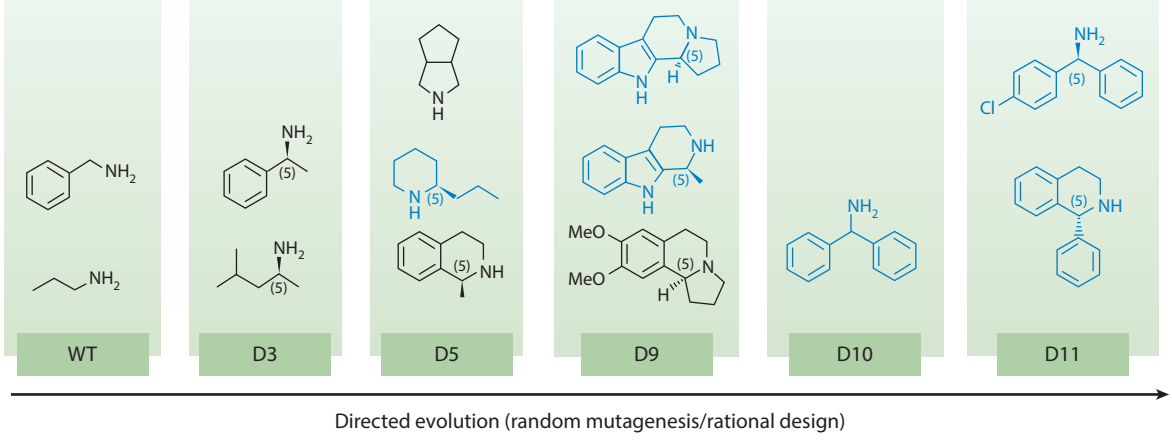


Figure 9

Panel of MAO-N variants from different generations. Reproduced from Ghislieri et al. (76), copyright 2013 ACS.

The overall deracemization process consists of both the enantiospecific oxidation of the amine to the imine state, as well as a typically nonspecific reduction of the imine or iminium ion back to the amine, leading to an enrichment of the remaining enantiomer. Crystal structures have been obtained for the variants MAO-N-D3 (Asn336Ser/Met348Lys/Ile246Met), which is able to oxidize secondary amines, and MAO-N-D5 (Asn336Ser/Met348Lys/Ile246Met/Thr384Asn/Asp385Ser), which is able to oxidize tertiary amines; both were obtained through point mutations (77). The function of the mutated amino acids can thus be explained. A noncovalently bound FAD molecule forms the base of an aromatic cage flanked by Trp430 and Phe466. Asn336Ser alleviates steric bulk behind Trp430; Ile246Met increases flexibility of the substrate binding site. Thr384Asn and Asp385Ser as part of the secondary shell perturb the side chain of Phe382 and thus alter sterics and electronics near the FAD.

MAO-N-D5 was employed to oxidize O-Methyl-N-hydroxyamines to the oxime with 99% e.e., exclusively in the (E)-configuration (78). Again, MAO-N-D5 was successful in oxidizing meso-pyrrolidines to Δ^1-pyrrolines, which are valuable synthons for highly enantiopure (*S*)-proline analogs and (*S*)-α-amino nitriles (79). MAO-N-D5 and amine oxidation again were crucial in a shortened synthesis of the antihepatitis C drug telaprevir, in which a cyclic meso-amine, 3,4-cyclopentyl-pyrroline, was oxidized to leave the building block 3,4,-*cis*-substituted 1-pyrroline (80).

MAO-N-D-10 exhibits high activity and enantioselectivity toward aminodiphenylmethane (benzhydrylamine)-containing substrates and thus has been applied in deracemization reactions for the efficient asymmetric synthesis of the generic active pharmaceutical ingredients Solifenacin and Levocetirizine (76). To access chiral 2,5-disubstituted pyrrolidines, a one-pot ω-transaminase (TA)/monoamine oxidase (MAO-N) cascade process on preparative scale was developed (81). As the regio- and stereoselectivity of both enzymes is complementary, the enzyme acting second, the monoamine oxidase, does not affect the chiral center established by the transaminase. The cascading reactions resulted in excellent enantio- and diastereoselectivity, >94% e.e. and >98% d.e., respectively.

Amine dehydrogenase (AmDH) from amino acid dehydrogenase (AADH). Two key mutations in positions 67 and 260 converted the amino acid dehydrogenase leucine dehydrogenase (LeuDH) from *Bacillus stearothermophilus* into an amine dehydrogenase (AmDH), termed L-AmDH (14). Both utilize the biological cofactor NADH and ammonia as a nitrogen source to reduce a prochiral α-keto acid or ketone to an α-amino acid or amine, respectively (**Figure 10**).

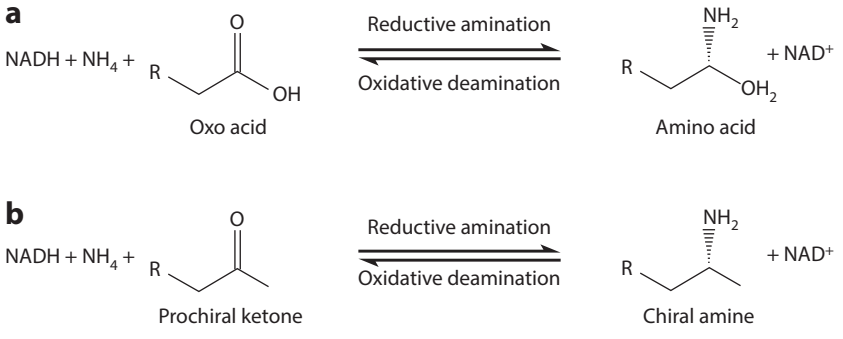


Figure 10

Reaction scheme of wild-type amino acid dehydrogenase (*a*) and amine dehydrogenase (*b*).

The direct amination of a ketone with ammonia was second on the priority list of aspirational reactions of the ACS Green Chemistry Institute Pharmaceutical Roundtable (82). Building blocks containing enantiomerically pure amines are required for a third of the 200 most prescribed brand name drugs in the United States; examples include sitagliptin (Januvia® and Janumed®) and rasagiline (Azilect®).

Mutation of the equivalent positions in phenylalanine DH (PheDH) from *Bacillus badius* (77 and 275) also yielded an active amine DH, F-AmDH (27). Whereas the new function was found through rational protein design based on the structure and mechanism of PheDH from *Rhodococcus rhodocrous* (83), optimization of rate and specificity was achieved by two-site CASTing (see Insufficient Designability of Proteins: Progress in Protein Engineering, above) to yield K67S/N260L for L-AmDH or K77S/N270L for F-AmDH and mutation of further sites, such as Glu114Val and Val291Cys in L-AmDH (14).

Cyclopropanation and amination biocatalysts: enhancement of the cytochrome P450 repertoire. Commonly, native cofactors or biobiological catalytic centers, such as metals, have been incorporated into nonfunctional protein scaffold to impart nonnative activity on the protein-catalytic center construct (84–86). The cytochrome P450 enzyme family already catalyzes several challenging chemical reactions, including some that were only known in synthetic chemistry but not biocatalysis [**Figure 11**; for a first review, see McIntosh (87)]. Frances Arnold's group at Caltech (Pasadena, CA) recently demonstrated that altered cofactor sites in enzymes can catalyze new reactions, notably the cyclopropanation of alkene via carbenes on the heme-containing cytochrome P450-BM3 scaffold from *B. megaterium*, never before observed in biocatalysis (15). Mutation of the conserved axial cysteine to serine enabled further novel catalysis, C-H amination (88) and N-H insertion (89). Lastly, a cytochrome P450 oxidase was discovered that catalyzed nitration of tryptophan derivatives (90).

Protein engineering in the vicinity of the heme group of the wild-type cytochrome P450-BM3 significantly altered and often improved the new functionality. The C400S variant, with the serine still bound to the heme, which is termed cyt P411 as the mutation to serine shifts the Soret band maximum to 411 nm, raises the reduction potential from Fe^{3+} to Fe^{2+} and as a consequence greatly improves NAD(P)H-driven activity of the variant (91). The rate of these novel P450 variants is not necessarily slow. A highly active His variant, cytochrome P450-BM3-Hstar, with turnover $>15\ s^{-1}$, accepted N,N-diethyl-2-phenylacrylamide for cyclopropanation under aerobic conditions en route to the enantioselective synthesis of a building block of the antidepressant levomilnacipran (Fetzima®) (92). The proximal ligand of the heme group seems to control cyclopropanation activity.

With the C400S variant (cyt P411) the activation and incorporation of a nitrogen moiety, more specifically azide, in the form of N-H amination could be demonstrated for the first time, leading to a benzosultam (88). Compared to wild-type cytochrome P450-BM3, where N-H amination had been observed as a contamination with very low turnover (TTN ≈ 2!), P450-CIS with 13 mutations gave a TTN of 9, cyt P411 a TTN of 32 (and 20% e.e.), and the best variant P411-CIS-T438S yielded a TTN of 383 (73% e.e.); consequently, improving turnover also increased selectivity.

Cytochrome P450-BM3 variants were used to create secondary amines from a primary amine and a carbenoid, more specifically ethyldiazoacetate (89). The reaction is chemoselective: no double insertion to a tertiary amine was observed (**Figure 11*c***), even though the diazo compound was not added continuously to keep down its concentration. Values of up to 83% e.e. were reached.

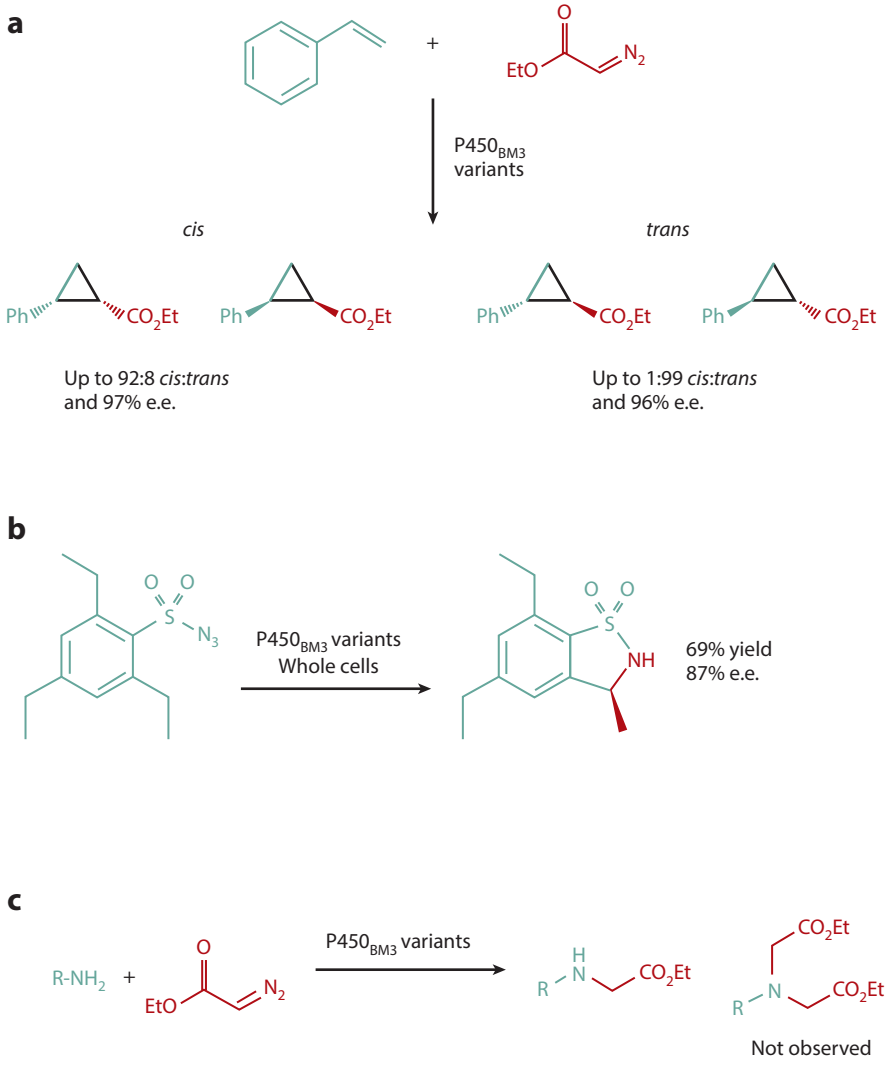


Figure 11

Novel reaction with cytochrome P450 enzymes. Reproduced from McIntosh et al. (87), copyright 2014 Current Opinions.

Novel Biocatalysts Through Computational Design

Recently, biocatalysts have been designed from scratch. Quantum mechanical and molecular mechanical models of the transition state and the first shell of residues around the active site were required, respectively.

Kemp eliminase. The one-step proton abstraction reaction from 5-nitroisooxazole to p-cyanonitrobenzene (Kemp elimination), in which proton abstraction results in irreversible N-O bond cleavage, does not feature any highly developed natural catalyst, so its uncatalyzed rate constant k_{uncat} is extremely low. This catalytically straightforward reaction thus was a good test case

for a novel method to develop biocatalysts, the design of an active enzyme by computation from scratch. With the aid of quantum mechanical models, transition states of the purported enzyme are created and then docked into a vast number of protein structures.

The original computationally designed Kemp eliminase KE70 employed a familiar His-Asp dyad for proton abstraction but produced an enzyme with low specificity, with a k_{cat}/K_M value of around 100 $(M{\cdot}s)^{-1}$, which is not synthetically useful. Protein engineering in the form of incorporating rationally designed mutations, as well as nine rounds of directed evolution featuring both random mutagenesis and DNA shuffling, resulted in a protein with a specificity of $5{\cdot}10^4$ $(M{\cdot}s)^{-1}$, which corresponds to a rate acceleration k_{cat}/k_{uncat} of approximately 10^7 over background (13).

KE07, another computationally designed Kemp eliminase, employed a glutamate, the catalytic base aiding proton abstraction, and an apolar binding site, which resulted in similarly low specificity as KE70. Seven rounds of evolution yielded 200-fold improvement to a k_{cat}/K_M value of $2.6{\cdot}10^3$ $(M{\cdot}s)^{-1}$, i.e., a rate acceleration k_{cat}/k_{uncat} of approximately 10^6, requiring eight mutations and resulting in decreased thermodynamic stability. The improvement of activity was attributed to modulation of the interaction of the catalytic glutamate with the substrate and the optimization of its pKa (93).

A third eliminase, KE59, featured the highest specificity of all computational designs and catalyzed elimination of unactivated benzisoxazoles but could not be subjected to the standard protocol of directed evolution owing to its lack of stability. However, upon the inclusion of stabilizing consensus mutations, the acceleration improved to 10^7 and catalytic specificity improved more than 2,000-fold to $6{\cdot}10^5$ $(M{\cdot}s)^{-1}$, mainly due to an improved k_{cat} value (94).

Another Kemp eliminase computationally designed on the basis of quantum mechanics and subject to extensive protein engineering achieved an acceleration ratio k_{cat}/k_{uncat} of $6{\cdot}10^8$ over background (95). On the basis of a very high-resolution crystal structure (1.09 Å), the success of the best variants was attributed to the high degree of precision of the positioning of catalytically relevant groups, as well as to shape complementarity between substrate and binding site, the same elements that render naturally evolved biocatalysts successful.

A recent review of de novo computational design of enzymes, while optimistic for the technique, stresses that protein engineering and structural studies are important for success; success is defined here as reaching catalytic activity, specificity, and acceleration on a par with naturally occurring enzymes (96).

SUCCESSFUL APPLICATIONS

Here, we limit our discussion to examples of biocatalyst development in pharma. Although there are good examples from industrial chemistry, the developments in pharma tend to have a more immediate impact.

Side Chain of Lipitor

Lipitor® (atorvastatin calcium) was the first drug to reach the annual sales of 10 billion dollars in the United States and until 2012 was the top-selling pharmaceutical product in the world (97). The statin drug is a 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibitor used to decrease levels of low density lipoprotein (LDL) cholesterol. The synthesis of the (3*R*,5*S*)-dihydroxyhexanoate side chain requires highly specific synthesis as high e.e. and d.e. values are required [99.5% and 99%, respectively (98)], encouraging the development of various biocatalytic routes to achieve these challenging goals (97).

In one example, a hydroxy ketoester intermediate to atorvastatin was obtained through the highly regio- and stereoselective single-site NADPH-dependent reduction of a β,δ-diketo

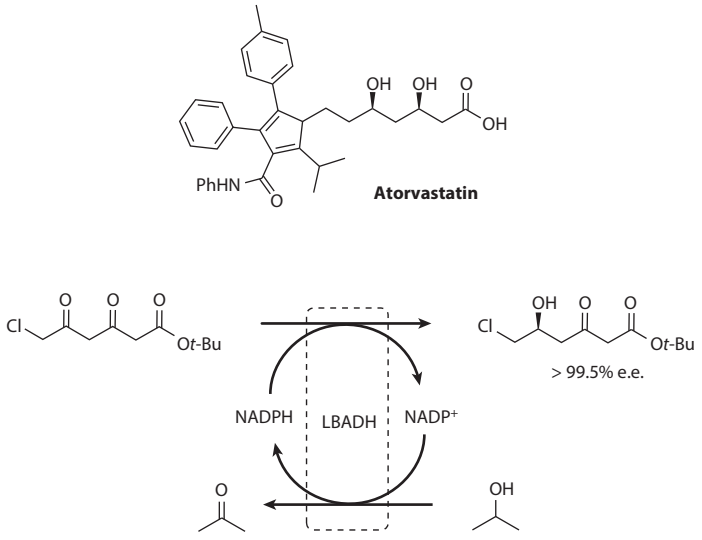


Figure 12

Enzymatic reduction using ADH from *Lactobacillus brevis* for the synthesis of a Lipitor® building block. Reproduced from Reference 109, copyright 2011 ACS.

ester using an ADH from *Lactobacillus brevis* (LADH). *tert*-Butyl 6-chloro-3,5-dioxohexanoate was converted in 72% yield to enantiopure *tert*-butyl (*S*)-6-chloro-5-hydroxy-3-oxohexanoate (**Figure 12**). 2-Propanol was used for cofactor regeneration as LADH catalyzes the oxidation of the alcohol to acetone at the expense of $NADP^+$, while simultaneously improving the solubility of the lipophilic substrate in the aqueous phase. A detailed kinetic study allowed improvement of the cofactor consumption by decreasing the amount of $NADP^+$ 60-fold while the cofactor total turnover number was improved 50-fold. The reaction was successfully scaled-up to 100 g (99–101).

Another route developed by Codexis involves both a ketoreductase (KRED) and a halohydrin dehalogenase (HHDH), starting from 4-chloroacetoacetate and giving access to (*R*)-4-cyano-3-hydroxybutyrate as key intermediate (**Figure 13**). Codexis won the 2006 Presidential Green Chemistry Challenge Award from the United States Environmental Protection Agency under the focus category "Greener Reaction Conditions" for the development of this process (102, 103).

Pregabalin via Early Resolution of a Diester

Pregabalin (Lyrica®), (*S*)-3-(aminomethyl)-5-methylhexanoic acid, used to treat fibromyalgia and control epilepsy, has been accessible in the discovery and first full-scale synthesis via chemical

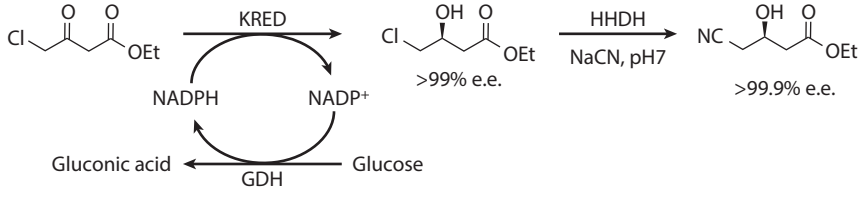


Figure 13

Codexis process for the production of a key intermediate in atorvastatin synthesis. Reproduced from Fox et al. (103), copyright 2007 Nature.

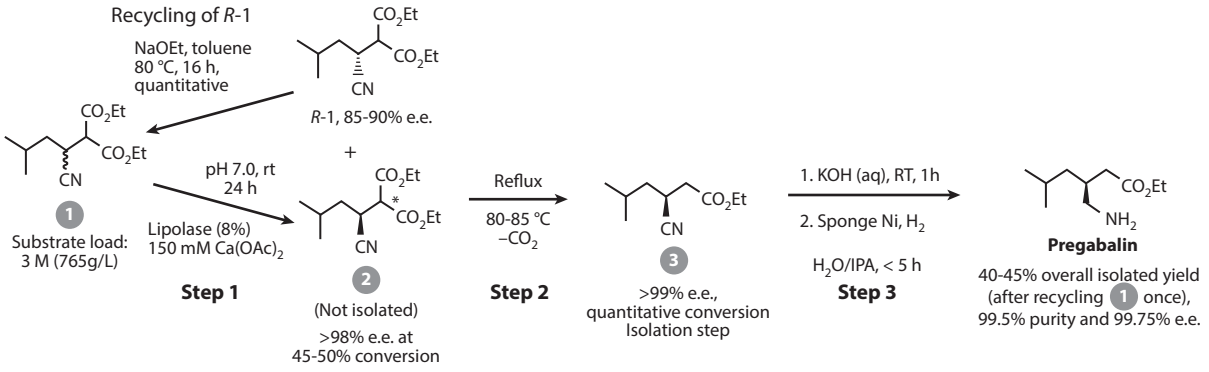


Figure 14

Second-generation manufacturing process for pregabalin (Lyrica®). Reproduced from Martinez et al. (104); copyright 2008 ACS.

technology employing resolution of the racemate, which resulted in low overall yield. The second-generation process in the key step resolves racemic 2-carboxyethyl-3-cyano-5-methylhexanoic acid ethyl ester to yield (*S*)-2-carboxyethyl-3-cyano-5-methylhexanoic acid (**Figure 14**) (104). The remaining (*R*)-enantiomer is racemized and reused. Thermally induced decarboxylation leaves (*S*)-3-cyano-5-methylhexanoic acid ethyl ester. Catalytic hydrogenation with Raney Ni/H_2 forms the API. This second-generation process, still practiced today, led to increased yield of 45% based on racemate (without recycle) and decreased solvent use such that the E factor [unit mass of waste (not water)/unit mass of API] was reduced from 86 to 17. Very recently, the ene reductase-catalyzed asymmetric hydrogenation of 3-cyano-5-methylhex-2-enoic acid (ethyl ester) to (*S*)-3-cyano-5-methylhexanoic acid (ethyl ester) was tested (105, 106).

Saxagliptin via Reductive Amination of Keto Acid Precursor

Another interesting application of amino acid dehydrogenases concerns the preparation of (*S*)-3-hydroxyadamantylglycine from 2-(3-hydroxy-1-adamantyl)-2-oxoethanoic acid using a modified form of a recombinant PheDH cloned from *Thermoactinomyces intermedius* and expressed in *Pichia pastoris* or *Escherichia coli*. The nonproteinogenic amino acid is a key intermediate required for the synthesis of Saxagliptin, a dipeptidyl peptidase IV inhibitor under development for treatment of type 2 diabetes mellitus (**Figure 15**). *E. coli* coexpressing PheDH and FDH from *P. pastoris* for cofactor regeneration allowed the successful reductive amination of the ketoacid on several kilogram scale (40 kg in 800-liter vessel and 99% conversion) (107).

Sitagliptin via Transamination of Ketositagliptin

A truly spectacular biocatalyst redesign was the substrate specificity enhancement of (*R*)-amino acid amino transferase ATA-117 to accommodate ketositagliptin, the key intermediate en route to sitagliptin phosphate, the API in Januvia™ and Janumet™ (108) (**Figure 16**). As the wild-type enzyme showed no activity toward ketositagliptin, a truncated substrate, a 3-keto-butyrpiperazine, was employed to overcome the lack of activity on their target substrate. Wild-type transaminase had low but detectable activity of 4% conversion in 24 h toward 3-keto-butyrpiperazine. Transition

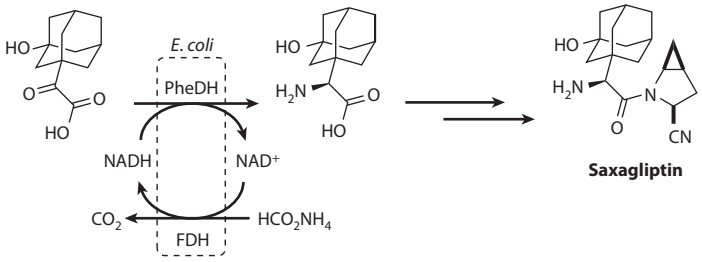


Figure 15

Reductive amination to (*S*)-3-hydroxyadamantylglycine, a precursor to Saxagliptin, using a PheDH/FDH system. Reproduced from Reference 109, copyright 2011 ACS.

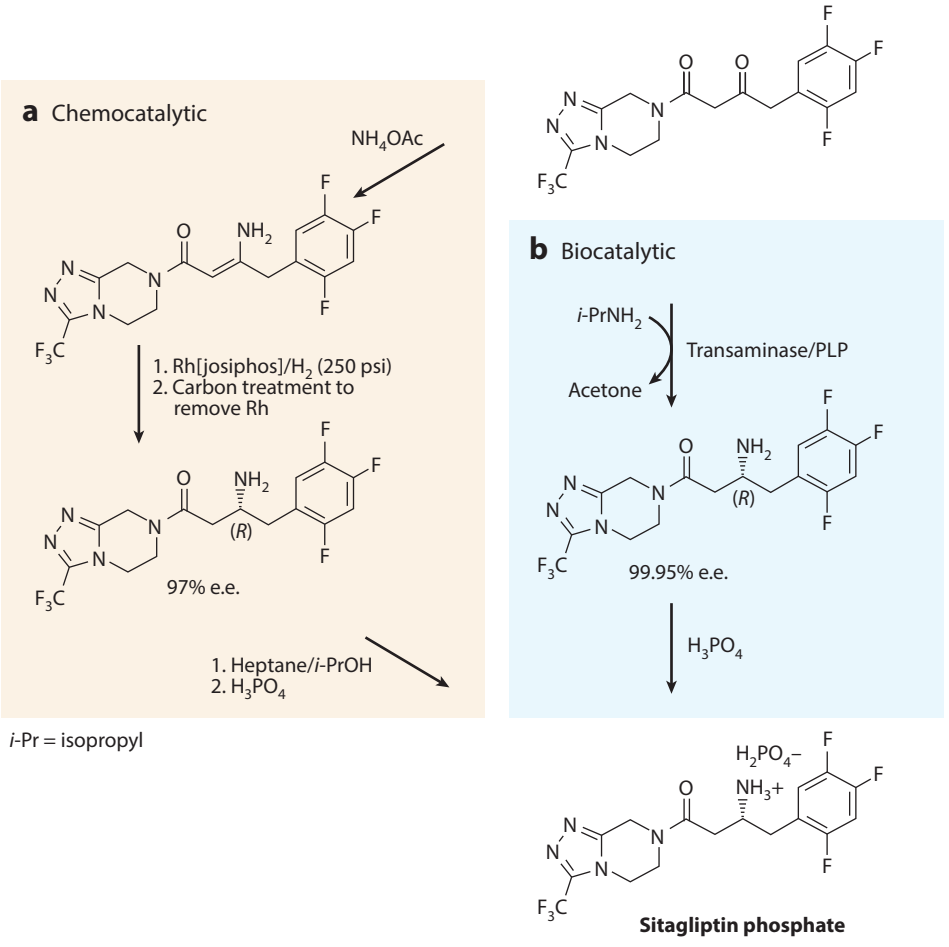
state docking of variants of the binding pocket of the truncated substrate yielded an enzyme with very low activity of <1% over 24 h, but again above background. After several rounds of random and targeted mutagenesis, employing restricted codon usage and ProSAR, an enzyme with much higher conversion and sufficient stability in the reaction medium 60:40 water:DMSO was developed. Ultimately, the final variant contains 27 mutations around the active-site and at the dimer interface. The enzymatic process, certified by the Federal Drug Administration in 2012, resulted in 3–5% points higher yield over the first-generation chemical process and one-third less solvent waste.

CONCLUSIONS

There is consensus in the community that biocatalysts are competitive or even superb in terms of activity and selectivity at near-ambient conditions. In the past, three points have often been raised to question the overall role of biocatalysis in catalytic processing: the lack of designability of biocatalysts, their limits of stability, and the insufficient number of well-characterized, ready-to-use biocatalysts.

Progress on all three fronts, outlined in the preceding section of this review, is very encouraging.

1. Novel protein engineering tools have significantly enhanced the toolbox for biocatalyst development (see above for descriptions of experimental and computational tools).
2. The deactivation of biocatalysts under various stresses, far more so than other kinds of catalysts, can be described quantitatively via rational models. Although protein stabilization commonly occurs in an incremental fashion as a function of amino acid sequence, there are now several cases of spectacular leaps of stabilization after accumulating all stabilizing mutations found in earlier rounds. Last but not least, the concept that stabilization against one type of stress, often thermoinstability, commonly also stabilizes against other types of stress, such as organic solvents, air-water or solid-water interfaces, or cavitation as a consequence of pumping or ultrasonication, is now experimentally considerably better founded than a few years ago.
3. A host of developments of novel biocatalysts in the past few years, in part fueled by improved designability and improved methods of stabilization, has ended up considerably broadening the toolbox for synthetic chemistry. Considering that the focus of this review on novel biocatalysts had to leave out the vastly improved characterization of several more established biocatalysts, such as transaminases, aldolases, and even the most frequently used lab-scale biocatalyst, *Candida antarctica* lipase B (CALB), the rate of progress of the field is truly encouraging. However, to claim victory, two additional aspects will have to fit into



i-Pr = isopropyl

Figure 16

New biocatalytic step (*a*) compared to first-generation, organometallically catalyzed steps (*b*). Reproduced from Savile et al. (108), copyright 2010 Science.

the picture as well: suitable synthetic targets for and commercial availability of biocatalysts, not just on a lab scale but on a full production scale.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The author owes a considerable amount of progress of biocatalysis in his lab to the funders of the NSF I/UCRC Center for Pharmaceutical Development. The author specially acknowledges support from NSF through this program under grant 0969003. Dr. Bettina Bommarius and Samantha Au are thanked for editorial advice.

LITERATURE CITED

1. Turner NJ. 2009. Directed evolution drives the next generation of biocatalysts. *Nat. Chem. Biol.* 5:568–74
2. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. 2012. Engineering the third wave of biocatalysis. *Nature* 485:185–94
3. Reetz MT. 2013. Biocatalysis in organic chemistry and biotechnology: past, present, and future. *J. Am. Chem. Soc.* 135:12480–96
4. Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR. 1986. The structure of the mouse glutathione-peroxidase gene: the selenocysteine in the active-site is encoded by the 'termination' codon, TGA. *EMBO J.* 5:1221–27
5. Zinoni F, Birkmann A, Stadtman TC, Bock A. 1986. Nucleotide-sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 83:4650–54
6. Srinivasan G, James CM, Krzycki JA. 2002. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* 296:1459–62
7. Hao B, Gong WM, Ferguson TK, James CM, Krzycki JA, Chan MK. 2002. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* 296:1462–66
8. Ngo JT, Tirrell DA. 2011. Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc. Chem. Res.* 44:677–85
9. Xiao H, Chatterjee A, Choi SH, Bajjuri KM, Sinha SC, Schultz PG. 2013. Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew. Chem. Int. Ed.* 52:14080–83
10. Rosenzweig R, Kay LE. 2014. Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu. Rev. Biochem.* 83:291–15
11. Oikawa H, Kobayashi T, Katayama K, Suzuki Y, Ichihara A. 1998. Total synthesis of (−)-solanapyrone A via enzymatic Diels-Alder reaction of prosolanapyrone. *J. Organ. Chem.* 63:8748–56
12. Reetz MT, Mondiere R, Carballeira JD. 2007. Enzyme promiscuity: first protein-catalyzed Morita-Baylis-Hillman reaction. *Tetrahedron Lett.* 48:1679–81
13. Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, et al. 2008. Kemp elimination catalysts by computational enzyme design. *Nature* 453:190–95
14. Abrahamson MJ, Vazquez-Figueroa E, Woodall NB, Moore JC, Bommarius AS. 2012. Development of an amine dehydrogenase for synthesis of chiral amines. *Angew. Chem. Int. Ed.* 51:3969–72
15. Coelho PS, Brustad EM, Kannan A, Arnold FH. 2013. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* 339:307–10
16. Fersht A. 1999. *Structure and Mechanism in Protein Science*. New York: Freeman
17. Radzicka A, Wolfenden R. 1995. A proficient enzyme. *Science* 267:90–93
18. Taylor EA, Palmer DRJ, Gerlt JA. 2001. The lesser "burden borne" by *o*-succinylbenzoate synthase: an "easy" reaction involving a carboxylate carbon acid. *J. Am. Chem. Soc.* 123:5824–25
19. Wolfenden R, Snider MJ. 2001. The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.* 34:938–45
20. Notomista E, Cafaro V, Bozza G, Di Donato A. 2009. Molecular determinants of the regioselectivity of toluene/*o*-xylene monooxygenase from *Pseudomonas* sp. strain OX1. *Appl. Environ. Microbiol.* 75:823–36
21. He ZQ, Nadeau LJ, Spain JC. 2000. Characterization of hydroxylaminobenzene mutase from pNBZ139 cloned from *Pseudomonas pseudoalcaligenes* JS45. *Eur. J. Biochem.* 267:1110–16
22. Hu S, Huang J, Mei LH, Yu Q, Yao SJ, Jin ZH. 2010. Altering the regioselectivity of cytochrome P450 BM-3 by saturation mutagenesis for the biosynthesis of indirubin. *J. Mol. Catal. B: Enzymatic* 67:29–35
23. Chenault HK, Dahmer J, Whitesides GM. 1989. Kinetic resolution of unnatural and rarely occurring amino acids: enantioselective hydrolysis of N-acyl amino acids catalyzed by acylase I. *J. Am. Chem. Soc.* 111:6354–64
24. Wu Q, Soni P, Reetz MT. 2013. Laboratory evolution of enantiocomplementary *Candida antarctica* lipase B mutants with broad substrate scope. *J. Am. Chem. Soc.* 135:1872–81
25. Reetz MT, Wang LW, Bocola M. 2006. Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. *Angew. Chem. Int. Ed.* 45:1236–41

26. Bommarius AS, Drauz K, Hummel W, Kula MR, Wandrey C. 1994. Some new developments in reductive amination with cofactor regeneration. *Biocatalysis* 10:37–47
27. Abrahamson MJ, Wong JW, Bommarius AS. 2013. The evolution of an amine dehydrogenase biocatalyst for the asymmetric production of chiral amines. *Adv. Synth. Catal.* 355:1780–86
28. Reetz MT, Kahakeaw D, Lohmer R. 2008. Addressing the numbers problem in directed evolution. *ChemBioChem* 9:1797–804

29a. Reetz MT, Bocola M, Carballeira JD, Zha DX, Vogel A. 2005. Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew. Chem. Int. Ed.* 44:4192–96

29b. Reetz MT. 2011. Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. *Angew. Chem. Int. Ed.* 50:138–74

30. Reetz MT, Carballeira JD. 2007. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* 2:891–903
31. Bommarius AS, Schwarm M, Drauz K. 2001. Comparison of different chemoenzymatic process routes to enantiomerically pure amino acids. *Chimia* 55:50–59
32. Rozzell JD. 1999. Biocatalysis at commercial scale: myths and realities. *Chim. Oggi-Chem. Today* 17:42–47
33. Wandrey A, Seelbach K, Wandrey C. 2006. *Industrial Biotransformations*. Weinheim, Ger.: Wiley-VCH. 570 pp.
34. Lumry R, Eyring H. 1954. Conformation changes of proteins. *J. Phys. Chem.* 58:110–20
35. Rogers TA, Bommarius AS. 2010. Utilizing simple biochemical measurements to predict lifetime output of biocatalysts in continuous isothermal processes. *Chem. Eng. Sci.* 65:2118–24
36. Eisenthal R, Danson MJ, Hough DW. 2007. Catalytic efficiency and k_{cat}/K_M: a useful comparator? *Trends Biotechnol.* 25:247–49
37. Ceccarelli EA, Carrillo N, Roveri OA. 2008. Efficiency function for comparing catalytic competence. *Trends Biotechnol.* 26:117–18
38. Fox RJ, Clay MD. 2009. Catalytic effectiveness, a measure of enzyme proficiency for industrial applications. *Trends Biotechnol.* 27:137–40
39. Palackal N, Brennan Y, Callen WN, Dupree P, Frey G, et al. 2004. An evolutionary route to xylanase process fitness. *Protein Sci.* 13:494–503
40. Voigt CA, Martinez C, Wang Z-G, Mayo SL, Arnold FH. 2002. Protein building blocks preserved by recombination, *Nature Struct. Biol.* 9:553–58
41. Meyer MM, Silberg JJ, Voigt CA, Endelman JB, Mayo SL, et al. 2003. Library analysis of SCHEMA-guided protein recombination, *Protein Sci.* 12:1686–93
42. Heinzelman P, Snow CD, Wu I, Nguyen C, Villalobos A, et al. 2009. A family of thermostable fungal cellulases created by structure-guided recombination. *Proc. Natl. Acad. Sci. USA* 106:5610–15
43. Heinzelman P, Snow CD, Smith MA, Yu XL, Kannan A, et al. 2009. SCHEMA recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability. *J. Biol. Chem.* 284:26229–33
44. Heinzelman P, Komor R, Kanaan A, Romero P, Yu XL, et al. 2010. Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination. *Protein Eng. Des. Select.* 23:871–80
45. Anbar M, Bayer EA. 2012. Approaches for improving thermostability characteristics in cellulases. *Cellulases* 510:261–71
46. Steipe B, Schiller B, Pluckthun A, Steinbacher S. 1994. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* 240:188–92
47. Lehmann M, Pasamontes L, Lassen SF, Wyss M. 2000. The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta-Protein Struct. Mol. Enzymol.* 1543:408–15
48. Lehmann M, Loch C, Middendorf A, Studer D, Lassen SF, et al. 2002. The consensus concept for thermostability engineering of proteins: further proof of concept. *Protein Eng.* 15:403–11
49. Amin N, Liu AD, Ramer S, Aehle W, Meijer D, et al. 2004. Construction of stabilized proteins by combinatorial consensus mutagenesis. *Protein Eng. Des. Sel.* 17:787–93
50. Vazquez-Figueroa E, Yeh V, Broering JM, Chaparro-Riggers JF, Bommarius AS. 2008. Thermostable variants constructed via the structure-guided consensus method also show increased stability in salts solutions and homogeneous aqueous-organic media. *Protein Eng. Des. Sel.* 21:673–80

51. Baik SH, Ide T, Yoshida H, Kagami O, Harayama S. 2003. Significantly enhanced stability of glucose dehydrogenase by directed evolution. *Appl. Microbiol. Biotechnol.* 61:329–35
52. Reetz MT, Soni P, Acevedo JP, Sanchis J. 2009. Creation of an amino acid network of structurally coupled residues in the directed evolution of a thermostable enzyme. *Angew. Chem. Int. Ed.* 48:8268–72
53. Blum JK, Ricketts MD, Bommarius AS. 2012. Improved thermostability of AEH by combining B-FIT analysis and structure-guided consensus method. *J. Biotechnol.* 160:214–21
54. Wijma HJ, Floor RJ, Jekel PA, Baker D, Marrink SJ, Janssen DB. 2014. Computationally designed libraries for rapid enzyme stabilization. *Protein Eng. Des. Sel.* 27:49–58
55. Bommarius AS, Paye MF. 2013. Stabilizing biocatalysts. *Chem. Soc. Rev.* 42:6534–65
56. Cowan DA. 1997. Thermophilic proteins: stability and function in aqueous and organic solvents. *Comp. Biochem. Physiol.-Mol. Integr. Physiol.* 118:429–38
57. Owusu RK, Cowan DA. 1989. Correlation between microbial protein thermostability and resistance to denaturation in aqueous-organic solvent 2-phase systems. *Enzyme Microb. Technol.* 11:568–74
58. Hao JJ, Berry A. 2004. A thermostable variant of fructose bisphosphate aldolase constructed by directed evolution also shows increased stability in organic solvents. *Protein Eng. Des. Sel.* 17:689–97
59. Reetz MT, Soni P, Fernandez L, Gumulya Y, Carballeira JD. 2010. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem. Commun.* 46:8657–58
60. Adalbjörnsson BV, Toogood HS, Fryszkowska A, Pudney CR, Jowitt TA, et al. 2010. Biocatalysis with thermostable enzymes: structure and properties of a thermophilic 'ene'-reductase related to old yellow enzyme. *ChemBioChem* 11:197–207
61. Sharma S, Mittal A, Gupta VK, Singh H. 2007. Improved stabilization of microencapsulated Cathepsin B in harsh conditions. *Enzyme Microb. Technol.* 40:337–42
62. Slusarczyk H, Felber S, Kula MR, Pohl M. 2000. Stabilization of NAD-dependent formate dehydrogenase from *Candida boidinii* by site-directed mutagenesis of cysteine residues. *Eur. J. Biochem.* 267:1280–89
63. Bommarius AS, Karau A. 2005. Deactivation of formate dehydrogenase (FDH) in solution and at gas-liquid interfaces. *Biotechnol. Prog.* 21:1663–72
64. Gadler P, Reiter TC, Hoelsch K, Weuster-Botz D, Faber K. 2009. Enantiocomplementary inverting *sec*-alkylsulfatase activity in cyano- and thio-bacteria *Synechococcus* and *Paracoccus* spp.: selectivity enhancement by medium engineering. *Tetrahedron-Asymmetry* 20:115–18
65. Schober M, Gadler P, Knaus T, Kayer H, Birner-Grunberger R, et al. 2011. A stereoselective inverting *sec*-alkylsulfatase for the deracemization of *sec*-alcohols. *Org. Lett.* 13:4296–99
66. Schober M, Knaus T, Toesch M, Macheroux P, Wagner U, Faber K. 2012. The substrate spectrum of the inverting *sec*-alkylsulfatase Pisa1. *Adv. Synth. Catal.* 354:1737–42
67. Schober M, Toesch M, Knaus T, Strohmeier GA, van Loo B, et al. 2013. One-pot deracemization of *sec*-alcohols: enantioconvergent enzymatic hydrolysis of alkyl sulfates using stereocomplementary sulfatases. *Angew. Chem. Int. Ed.* 52:3277–79
68. Lara M, Mutti FG, Glueck SM, Kroutil W. 2008. Biocatalytic cleavage of alkenes with O_2 and *Trametes hirsuta* G FCC 047. *Eur. J. Organ. Chem.* 2008(21):3668–72
69. Rajagopalan A, Mutti FG, Kroutil W. 2012. Biocatalytic cleavage of alkenes with oxygen and *Trametes hirsuta* G FCC047. In *Practical Methods for Biocatalysis and Biotransformations 2*, ed. J. Whittall, P. Sutton, pp. 199–202. West Sussex, UK: Wiley. 388 pp.
70. Lara M, Mutti FG, Glueck SM, Kroutil W. 2009. Oxidative enzymatic alkene cleavage: indications for a nonclassical enzyme mechanism. *J. Am. Chem. Soc.* 131:5368–69
71. Rajagopalan A, Schober M, Emmerstorfer A, Hammerer L, Migglautsch A, et al. 2013. Enzymatic aerobic alkene cleavage catalyzed by a Mn^{3+}-dependent proteinase A homologue. *ChemBioChem* 14:2427–30
72. Wallner S, Winkler A, Riedl S, Dully C, Horvath S, et al. 2012. Catalytic and structural role of a conserved active site histidine in berberine bridge enzyme. *Biochemistry* 51:6139–47
73. Schrittwieser JH, Resch V, Sattler JH, Lienhart WD, Durchschein K, et al. 2011. Biocatalytic enantioselective oxidative C-C coupling by aerobic C-H activation. *Angew. Chem. Int. Ed.* 50:1068–71
74. Resch V, Schrittwieser JH, Wallner S, Macheroux P, Kroutil W. 2011. Biocatalytic oxidative C-C bond formation catalysed by the berberine bridge enzyme: optimal reaction conditions. *Adv. Synth. Catal.* 353:2377–83

75. Schrittwieser JH, Groenendaal B, Resch V, Ghislieri D, Wallner S, et al. 2014. Deracemization by simultaneous bio-oxidative kinetic resolution and stereoinversion. *Angew. Chem. Int. Ed.* 53:3731–34
76. Ghislieri D, Green AP, Pontini M, Willies SC, Rowles I, et al. 2013. Engineering an enantioselective amine oxidase for the synthesis of pharmaceutical building blocks and alkaloid natural products. *J. Am. Chem. Soc.* 135:10863–69
77. Atkin KE, Reiss R, Koehler V, Bailey KR, Hart S, et al. 2008. The structure of monoamine oxidase from *Aspergillus niger* provides a molecular context for improvements in activity obtained by directed evolution. *J. Mol. Biol.* 384:1218–31
78. Eve TSC, Wells A, Turner NJ. 2007. Enantioselective oxidation of O-methyl-N-hydroxylamines using monoamine oxidase N as catalyst. *Chem. Commun.* 2007(15):1530–31
79. Kohler V, Bailey KR, Znabet A, Raftery J, Helliwell M, Turner NJ. 2010. Enantioselective biocatalytic oxidative desymmetrization of substituted pyrrolidines. *Angew. Chem. Int. Ed.* 49:2182–84
80. Znabet A, Polak MM, Janssen E, de Kanter FJJ, Turner NJ, et al. 2010. A highly efficient synthesis of telaprevir by strategic use of biocatalysis and multicomponent reactions. *Chem. Commun.* 46:7918–20
81. O'Reilly E, Iglesias C, Ghislieri D, Hopwood J, Galman JL, et al. 2014. A regio-and stereoselective w-transaminase/monoamine oxidase cascade for the synthesis of chiral 2,5-disubstituted pyrrolidines. *Angew. Chem. Int. Ed.* 53:2447–50
82. Constable DJC, Dunn PJ, Hayler JD, Humphrey GR, Leazer JL, et al. 2007. Key green chemistry research areas—a perspective from pharmaceutical manufacturers. *Green Chem.* 9:411–20
83. Brunhuber NMW, Thoden JB, Blanchard JS, Vanhooke JL. 2000. Rhodococcus L-phenylalanine dehydrogenase: kinetics, mechanism, and structural basis for catalytic specificity. *Biochemistry* 39:9174–87
84. Ringenberg MR, Ward TR. 2011. Merging the best of two worlds: artificial metalloenzymes for enantioselective catalysis. *Chem. Commun.* 47:8470–76
85. Lo C, Ringenberg MR, Gnandt D, Wilson Y, Ward TR. 2011. Artificial metalloenzymes for olefin metathesis based on the biotin-(strept)avidin technology. *Chem. Commun.* 47:12065–67
86. Durrenberger M, Ward TR. 2014. Recent achievements in the design and engineering of artificial metalloenzymes. *Curr. Opin. Chem. Biol.* 19:99–106
87. Mclntosh JA, Farwell CC, Arnold FH. 2014. Expanding P450 catalytic reaction space through evolution and engineering. *Curr. Opin. Chem. Biol.* 19:126–34
88. McIntosh JA, Coelho PS, Farwell CC, Wang ZJ, Lewis JC, et al. 2013. Enantioselective intramolecular C-H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. *Angew. Chem. Int. Ed.* 52:9309–12
89. Wang ZJ, Peck NE, Renata H, Arnold FH. 2014. Cytochrome P450-catalyzed insertion of carbenoids into N-H bonds. *Chem. Sci.* 5:598–601
90. Dodani SC, Cahn JKB, Heinisch T, Brinkmann-Chen S, McIntosh JA, Arnold FH. 2014. Structural, functional, and spectroscopic characterization of the substrate scope of the novel nitrating cytochrome P450 TxtE. *Chembiochem* 15:2259–67
91. Coelho PS, Wang ZJ, Ener ME, Baril SA, Kannan A, et al. 2013. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins *in vivo*. *Nat. Chem. Biol.* 9:485–87
92. Wang ZJ, Renata H, Peck NE, Farwell CC, Coelho PS, Arnold FH. 2014. Improved cyclopropanation activity of histidine-ligated cytochrome P450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chem. Int. Ed.* 53:6810–13
93. Khersonsky O, Rothlisberger D, Dym O, Albeck S, Jackson CJ, et al. 2010. Evolutionary optimization of computationally designed enzymes: Kemp eliminases of the KE07 series. *J. Mol. Biol.* 396:1025–42
94. Khersonsky O, Kiss G, Rothlisberger D, Dym O, Albeck S, et al. 2012. Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59. *Proc. Natl. Acad. Sci. USA* 109:10358–63
95. Blomberg R, Kries H, Pinkas DM, Mittl PRE, Grutter MG, et al. 2013. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* 503:418–21
96. Hohne M, Bornscheuer UT. 2014. Protein engineering from "scratch" is maturing. *Angew. Chem. Int. Ed.* 53:1200–2
97. Patel JM. 2009. Biocatalytic synthesis of atorvastatin intermediates. *J. Mol. Catal. B: Enzym.* 61:123–28

98. Muller M. 2005. Chemoenzymatic synthesis of building blocks for statin side chains. *Angew. Chem. Int. Ed.* 44:362–65
99. Wolberg M, Hummel W, Wandrey C, Muller M. 2000. Highly regio- and enantioselective reduction of 3,5-dioxocarboxylates. *Angew. Chem. Int. Ed.* 39:4306–8
100. Wolberg M, Villela M, Bode S, Geilenkirchen P, Feldmann R, et al. 2008. Chemoenzymatic synthesis of the chiral side-chain of statins: application of an alcohol dehydrogenase catalysed ketone reduction on a large scale. *Bioprocess Biosyst. Eng.* 31:183–91
101. Wolberg M, Hummel W, Muller M. 2001. Biocatalytic reduction of β,δ-diketo esters: a highly stereoselective approach to all four stereoisomers of a chlorinated β,δ-dihydroxy hexanoate. *Chem. Eur. J.* 7:4562–71
102. Tao JH, Xu JH. 2009. Biocatalysis in development of green pharmaceutical processes. *Curr. Opin. Chem. Biol.* 13:43–50
103. Fox RJ, Davis SC, Mundorff EC, Newman LM, Gavrilovic V, et al. 2007. Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* 25:338–44
104. Martinez CA, Hu S, Dumond Y, Tao J, Kelleher P, Tully L. 2008. Development of a chemoenzymatic manufacturing process for pregabalin. *Organ. Process Res. Dev.* 12:392–98
105. Winkler CK, Clay D, Davies S, O'Neill P, McDaid P, et al. 2013. Chemoenzymatic asymmetric synthesis of pregabalin precursors via asymmetric bioreduction of β-cyanoacrylate esters using ene-reductases. *J. Organ. Chem.* 78:1525–33
106. Winkler CK, Clay D, Turrini NG, Lechner H, Kroutil W, et al. 2014. Nitrile as activating group in the asymmetric bioreduction of β-cyanoacrylic acids catalyzed by ene-reductases. *Adv. Synth. Catal.* 356:1878–82
107. Hanson RL, Goldberg SL, Brzozowski DB, Tully TP, Cazzulino D, et al. 2007. Preparation of an amino acid intermediate for the dipeptidyl peptidase IV inhibitor, saxagliptin, using a modified phenylalanine dehydrogenase. *Adv. Synth. Catal.* 349:1369–78
108. Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S, et al. 2010. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* 329:305–9
109. Hall M, Bommarius AS. 2011. Enantioenriched compounds via enzyme-catalyzed redox reactions. *Chem. Rev.* 111:4088–110