nature chemical biology

Directed evolution drives the next generation of biocatalysts

Nicholas J Turner

Enzymes are increasingly being used as biocatalysts in the generation of products that have until now been derived using traditional chemical processes. Such products range from pharmaceutical and agrochemical building blocks to fine and bulk chemicals and, more recently, components of biofuels. For a biocatalyst to be effective in an industrial process, it must be subjected to improvement and optimization, and in this respect the directed evolution of enzymes has emerged as a powerful enabling technology. Directed evolution involves repeated rounds of (i) random gene library generation, (ii) expression of genes in a suitable host and (iii) screening of libraries of variant enzymes for the property of interest. Both *in vitro* screening–based methods and *in vivo* selection–based methods have been applied to the evolution of enzyme function and properties. Significant developments have occurred recently, particularly with respect to library design, screening methodology, applications in synthetic transformations and strategies for the generation of new enzyme function.

Biocatalysis involves the use of enzymes, either in semipurified or immobilized form, or as whole-cell systems, for the preparation of molecules that are traditionally made using chemical synthesis. The field continues to expand its horizons, particularly as applied to the synthesis of fine chemicals and pharmaceuticals[1–4]. A useful biocatalyst (in particular, one that is suitable for application in an industrial context) is characterized by a number of features, including high catalytic turnover (>500 min^{-1}), high selectivity for the particular transformation (for example, enantio-, regio- and chemoselectivity) and high process stability under the conditions that are required for the chemical transformation[5]. Biocatalysis often complements (rather than directly competes with) chemocatalysis in that it is probably best used for preparing relatively small (molecular weight < 500), enantiomerically pure chiral building blocks that are then combined with other chemically derived molecules to produce the final product. Biocatalysts have historically been widely applied in the brewing and food industries but are now finding increased application in the synthesis of fine chemicals, pharmaceuticals, agrochemicals and even bulk chemicals. A recent and important trend is the development of new biocatalyst-based processes for the production of biofuels using renewable starting materials as an alternative to fossil fuels. In this sense biocatalysis should be viewed as an essential ingredient of the industrial biotechnology revolution that is happening all around us and that will alter the landscape of where the products and processes of the future are derived from. This review covers recent developments in the field of directed evolution of enzymes, highlighting advances in screening and library design and also giving examples of where enzyme evolution has had a major impact on the development of new biocatalysts.

School of Chemistry, University of Manchester, Manchester Interdisciplinary Biocentre, Manchester, UK. Correspondence should be addressed to N.J.T. (nicholas.turner@manchester.ac.uk).

Published online 20 July 2009; doi:10.1038/nchembio.203

Irrespective of the target application that is envisaged, there are a number of key challenges that must be addressed in the development of a robust biocatalyst for practical applications (**Fig. 1**). The initial phase of enzyme discovery involves (i) screening available microbial culture collections or libraries of enzymes from a variety of sources in order to identify an enzyme or group of enzymes with some activity toward the substrate of interest, (ii) construction of synthetic genes based on literature precedents or (iii) use of bioinformatics to identify homologous genes. In some cases the wild-type enzyme may have the required activity and selectivity for further development work. However, a more likely scenario is that improvements in catalytic activity, selectivity and (importantly) stability under process conditions are required for practical applications. Even with the benefit of a high-resolution crystal structural of an enzyme, rational redesign to produce optimized biocatalysts is complex and has proven successful in only a small number of cases.

The early 1990s witnessed an alternative approach to the problem of enzyme optimization with the publication of some of the key early papers in directed evolution. These papers described methods for library generation via DNA shuffling[6,7] and error-prone PCR, and also gave examples of changing the properties of an enzyme, such as organic solvent tolerance[8] and enantioselectivity[9]. However, if the directed evolution of enzymes (**Fig. 1**) were to become a generally applicable technique, then significant developments would be needed in both methods for library generation, and we would need to develop the ability to screen for a wide range of different chemical transformations[10–21]. In the early days, emphasis was placed on the development of screening methods that could assess very large libraries of variants (for example, 10^8–10^{10}) in the expectation that larger libraries meant a greater chance of finding the desired improvements in activity or selectivity. As more examples were published of successful improvements with much smaller libraries (for example, 10^3–10^6), the emphasis shifted to the generation of smaller, higher quality libraries, as outlined below.

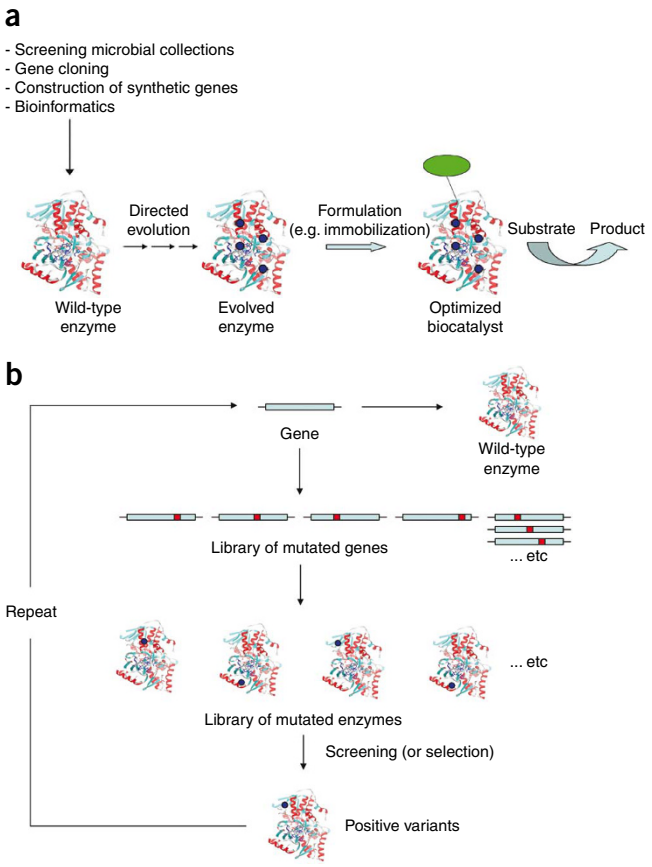

Figure 1 Biocatalysis in practice. (**a**) Stages in the discovery and development of a robust biocatalyst for practical applications. The advantages of biocatalysis include the high catalytic rate, the potential for high selectivity of the reaction (including enantio-, regio- and chemoselectivity) and the potential for low specificity, thus allowing broad substrate acceptance. Biocatalysis is also important in the development of sustainable processes. (**b**) Directed evolution of enzymes by iterative cycles of random mutagenesis, protein expression and screening or selection of the desired activity.

For a protein that is composed of say 400 amino acids, there are approximately 20^{400} different possible sequences that could exist. This level of diversity should be compared with the fact that the total number of atoms in the universe is estimated to be between 10^{70} and 10^{80} (see http://en.wikipedia.org/wiki/observable_universe)! Approaching the problem in a different way, if for example only one amino acid in a protein containing 400 amino acids is exchanged for any of the other 19 amino acids, there are already a total of 7,600 different variants. For two amino acids, the number rises to 144,000, for three the total is about 2.7×10^6 and for four replacements the total is 5.2×10^7! The challenge of accessing and screening such large sequence space brings with it an additional problem: the issue of how to navigate between different sequences. Proteins are generally regarded as having 'rugged' landscapes in terms of catalytic activity, which makes the correlation of sequence with activity difficult. Typically researchers will try to identify the 'hot spots' in an enzyme via high-throughput screening and then subsequently optimize individual amino acids via saturation mutagenesis (**Fig. 2**). In order to address the problem of navigating sequence space, researchers have begun to apply multivariate statistical techniques to model protein sequence-function relationships and hence guide the evolutionary process by rapidly identifying beneficial diversity for recombination. For example, scientists at Codexis have developed an algorithm termed Pro-SAR (protein sequence-activity relationships) in which all mutations are assigned as either beneficial, neutral or deleterious based on the impact they have on the enzyme function that is being screened for[22,23]. Thus all mutations are assessed and may be incorporated into the final sequence of the fully optimized protein. The Pro-SAR approach was used to optimize an enantioselective haloalkane dehalogenase used in the commercial manufacture of the side chain of atorvastatin[24]. In this process, ethyl 4-chloro-3-ketobutyrate is first treated with a ketoreductase to catalyze asymmetric reduction of the keto group to the alcohol, followed by the introduction of the cyano group catalyzed by a dehalogenase (**Fig. 3**). The dehalogenase was subjected to multiple rounds of directed evolution in order to improve the volumetric productivity ~4,000-fold. Altogether, a total of 38 mutations were introduced compared to the wild-type enzyme.

Reducing the screening effort in directed evolution

As a consequence of the problems associated with accessing even a small fraction of the vast sequence space of proteins as outlined above, researchers have sought to develop strategies that significantly reduce the size of the libraries that need to be screened without compromising the likelihood of finding significant improvements in performance. For example, for many enzymes it may be sufficient just to target the active site residues[25]. Typically an active site of a protein will contain about 10–15 amino acid residues that make direct contact with the substrate (first coordination sphere) and a further 20–30 in the second coordination sphere. Morley *et al.*[26] analyzed results from their own work and from the work of other groups and concluded that in general, closer mutations are more beneficial, particularly in terms of the effect that they have on substrate specificity and enantioselectivity. Reetz *et al.*[27–29] have developed a very useful method termed iterative CASTing (combinatorial active site testing) in which active site residues are selected and organized into groups of two or three residues from which libraries can be subsequently generated. Screening of these initial libraries is undertaken, the hits from these libraries are combined, and new libraries are generated and are further screened in an iterative fashion. Clearly either a crystal structure or homology model is required for CASTing to be applied. Another approach to reduce the screening effort is to use reduced amino acid alphabets in the construction of proteins[30]. For example, by changing

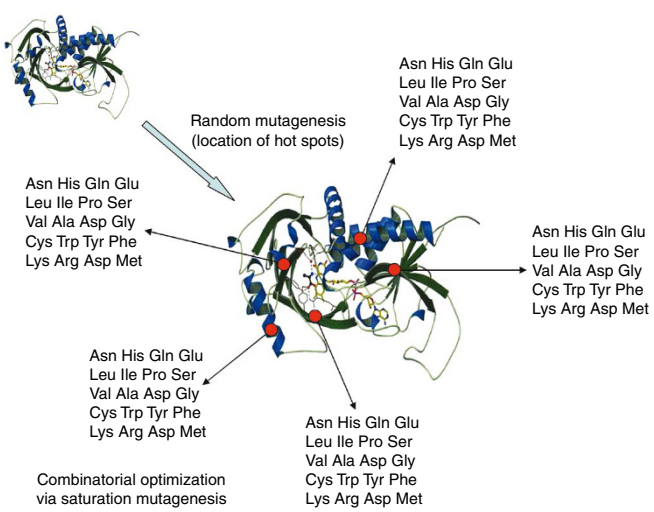

Figure 2 Location of 'hot spots' in an enzyme by random mutagenesis coupled with high-throughput screening. Subsequent saturation mutagenesis can be used to optimize at specific amino acid sites.

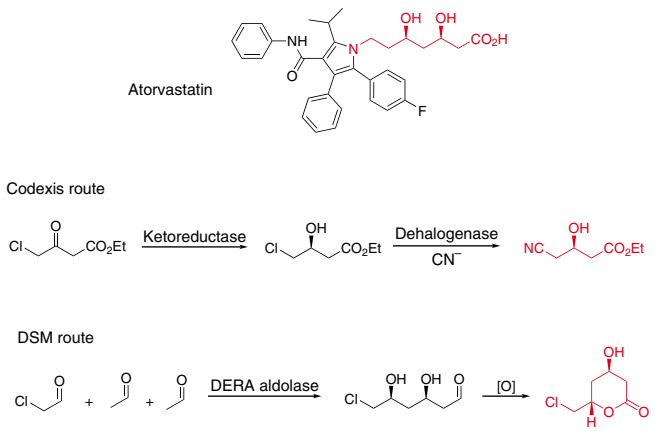

Figure 3 Two complementary biocatalytic approaches to the synthesis of the side chain of atorvastatin. The Codexis route uses two successive enzymes—a ketoreductase followed by a dehalogenase. The DSM route uses an aldolase to construct two new carbon-carbon bonds.

the nucleotide components during the PCR reaction it is possible to generate proteins that are composed of only 12 rather than 20 different amino acids. Reetz and Wu[31] have recently shown that the screening of such libraries can lead to significant improvements in activity by using saturation mutagenesis at homologous enzyme positions, as shown by the directed evolution of enantioselective mutants of the thermally robust phenylacetone monooxygenase (PAMO).

Enzyme promiscuity and evolution of new function

A powerful theme that has recently emerged and that is highly relevant to directed evolution of enzymes is that of promiscuous activities[32–34]. Enzyme promiscuity refers to situations where an enzyme is able to catalyze a reaction that is quite distinct from its natural function but for which there are links between the two chemical transformations at the mechanistic rather than structural level. In order to try to understand the relationship between promiscuity and evolution of function, both Tawfik and Arnold have recently introduced the concept of neutral drift[35–39]. Neutral drift is the gradual accumulation of mutations under selection to maintain a protein's original function and structure. Such drifts have been demonstrated experimentally in the laboratories of Arnold and Tawfik by generating libraries and then screening for the original function rather than a new function. For example, Gupta *et al.*[38] studied the neutral drift of serum paraoxonase (PON1) and found that almost half of the 311 neutral variants that were characterized exhibited substantial changes in promiscuous activity. Similar changes in promiscuous activities were found by Bloom *et al.*[39] with P450 BM3 neutral variants.

It is generally thought that the earliest forms of enzymes had broad substrate specificity or reaction ambiguity and subsequently diverged and became more reaction-specific to give today's modern enzymes. Such enzyme superfamilies are exemplified by the enolase[40,41] and crotonase superfamilies[42]. For example, the enolase superfamily comprises a group of related enzymes that have a common ability to stabilize transition states involving the generation of enolates. This common mechanism is then manifested in a range of different enzymes that catalyze different reactions, including N-acylamino acid racemases, mandelate racemase and O-succinylhydrolase. Currently there is great interest in exploiting reaction promiscuity as a means of developing enzymes with new functionalities. The halohydrin dehalogenase from *Agrobacterium radiobacter* catalyzes the nucleophilic ring opening of epoxides by various nucleophiles to afford the corresponding 1-substituted butan-2-ols with moderate conversion and good enantioselectivity (*E* value = 24–31). Ring opening with cyanate afforded oxazolidinone in high enantiomeric excess (e.e.) as a result of subsequent cyclization[43]. The arylmalonate decarboxylase from *Alcaligenesis bronchisepticus* was mutated to a racemase by introduction of an additional active site base, allowing the racemization of enantiomerically pure amino acids. The mutant was tolerant of different aromatic substituents and to a lesser extent replacement of the methyl group[44]. A single active site mutation in *Bacillus stearothermophilus* alanine racemase resulted in the introduction of new aldolase activity. The activity toward phenylserine was increased compared to the wild type by five orders of magnitude[45]. A particularly powerful approach is the combination of natural promiscuity with directed evolution. Recently Jochens *et al.*[46] showed that an esterase can be converted to an epoxide hydrolase by initial redesign of the active site followed by random mutagenesis and screening for epoxide hydrolase activity.

The creation of enzymes with truly new function represents one of the greatest current challenges in the field of biocatalysis[47–54]. Recently Röthlisberger *et al.*[49] reported the computational design of retro-aldolases that use four different catalytic motifs to catalyze cleavage of a carbon-carbon bond in a non-natural substrate. Interestingly, the accuracy of the design at the molecular level was confirmed by an X-ray crystal structure of the active proteins, both of which were essentially superimposable on the design model. Seelig *et al.*[54] reported the evolution of a new enzyme activity by selection with mRNA display from a randomized zinc finger scaffold that had no enzymatic activity. Because single-turnover product formation could be observed, variants that exhibited weak activity could be identified.

New screening and selection strategies

Whatever strategy is used for library generation, it is vital to have in place a robust and high-throughput method for either screening or selection of the desired function[55]. In the past ten years, much effort and ingenuity have gone into the development of a broad range of screening and selection methods, including those based on ribosome display and mRNA display, to allow the identification of variants that have the desired characteristics. The essential functions of an enzyme that are desirable to be screened for include (i) catalytic activity, (ii) selectivity, including enantioselectivity, diastereoselectivity and regioselectivity,

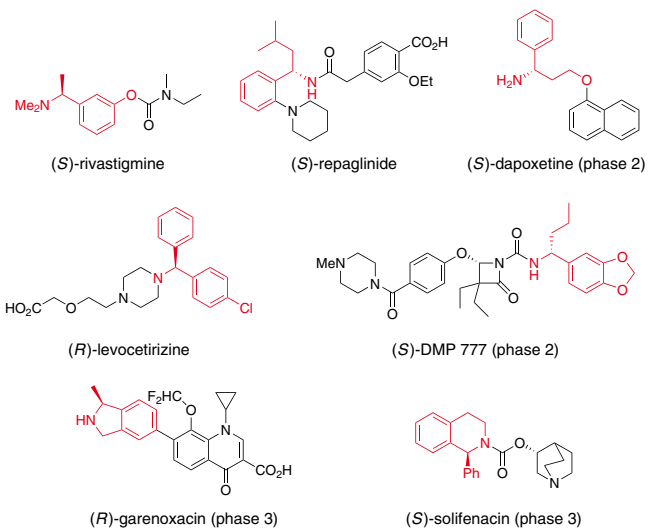

Figure 4 Pharmaceutical drugs in development or in the clinic that contain chiral amine building blocks highlighted in red.

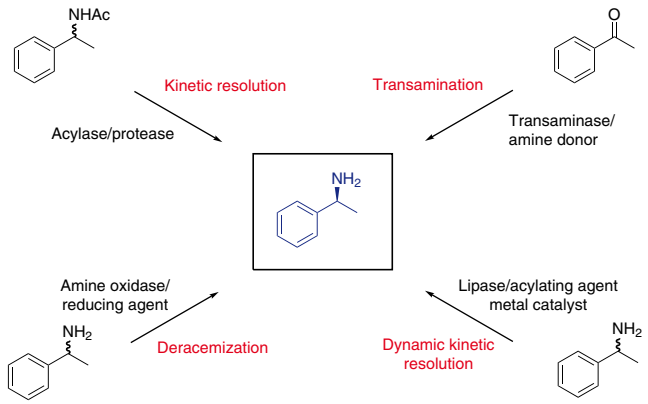

Figure 5 Deracemization of racemic amines and amino acids by combination of an oxidase enzyme with a chemical reducing agent.

and (iii) thermal stability, including stability to organic solvents. In order to be able to screen for any of these desired functions, it is essential to establish a rapid method for monitoring the reaction that is catalyzed by the enzyme. For example, lipase-catalyzed hydrolysis of an ester under aqueous conditions yields a carboxylic acid product, which results in a pH change that can be monitored by the introduction of pH indicator. Screening for epoxide hydrolase activity on a chip-based system has recently been reported[56]. Sorting cells by fluorescence-activated cell sorting (FACS) allows very large library sizes to be addressed and has been coupled with *in vitro* compartmentalization as a very effective high-throughput screening method[57,58]. Likewise, phage display[59] and yeast display[60] techniques have been developed and are effective for evolving enzymes. Circular permutation is an interesting approach in which the N and C termini of a protein can be systematically moved, one amino acid at a time, around the protein of interest. This method has been used to screen for variants of *Candida antarctica* lipase B with improved activity and stability[61].

Directed evolution of monoamine oxidase

To illustrate a number of the points discussed above, I will now present a more in-depth case study. The aim of this case study is to highlight some of the key strategies involved in an enzyme evolution program and also to present an analysis of the learning process. We set out to develop a platform biocatalytic technology for the preparation of enantiomerically pure chiral amines, which are an important class of chiral building block, particularly for the synthesis of biologically active pharmaceutical drugs (**Fig. 4**). Despite their importance, relatively few biocatalytic methods have been developed for their preparation, and existing methods rely principally on either the kinetic resolution of racemates using hydrolytic enzymes such as lipases and acylases or the more recent development of the use of transaminases for the conversion of ketones to chiral amines (**Fig. 5**). A conceptually different approach to preparation of enantiomerically pure chiral amines relies on deracemization of a racemic starting amine to provide either enantiomer in high optical purity. The deracemization process requires the combination of an enantioselective amine oxidase enzyme (conversion of amine to imine) together with a nonselective chemical reducing agent (reduction of imine to amine). Repeated cycles of enantioselective oxidation coupled with nonselective reduction result in accumulation of either the (*R*)- or (*S*)-amine enantiomer, depending on the stereopreference of the oxidase enzyme.

Prior studies in our group had established that deracemization could be successfully applied to α-amino acids[62,63]. These studies led to the identification of suitable L- and D-amino acid oxidases, together with a range of different chemical reducing agents that were compatible with the enzymes (for example, catalytic transfer hydrogenation using metal/formate or amine boranes). However, in order to extend the deracemization methodology to chiral amines, it was necessary to develop enantioselective amine oxidases with the desired substrate specificity by directed evolution of existing amine oxidases. Monoamine oxidase (MAO-N) from *Aspergillus niger* was identified as an attractive starting point for directed evolution on account of the high catalytic activity (k_{cat} ~ 500 min^{-1}) of the wild-type enzyme. However, the native enzyme was found to have narrow substrate specificity, typically limited to simple nonchiral amines such as butylamine and benzylamine.

In order to screen libraries of variants of MAO-N, we developed a high-throughput screening method based on capture of the byproduct hydrogen peroxide using horseradish peroxidase and a suitable substrate[64]. *Escherichia coli* colonies expressing the amine oxidase were grown on nitrocellulose membranes and then assayed for activity, resulting in the production of dark pink colonies indicative of high enzyme activity. It was envisaged that this solid-phase screen would be able to assess about 100,000 variants per round of screening. Libraries of MAO-N were prepared initially by using a mutator strain, and subsequently by error-prone PCR, and screened against α-methylbenzylamine as a model substrate. Several rounds of directed evolution led to the identification of a mutant (D5) containing 5 mutations that was found to catalyze the enantioselective oxidation of a wide range of chiral 1°, 2° and 3° amines[65–67]. The MAO-N D5 mutant was used for the preparative deracemization of a range of racemic amines; interestingly, it was also found to oxidize several racemic O-methyl-N-hydroxylamines, yielding unreacted enantiomerically pure (*R*) isomer as well as the corresponding (*E*)-oximes[68].

Screening the D5 variant against a wide range of amine substrates revealed that it has broad substrate specificity and high (*S*) enantioselectivity. To encourage greater usage of the enzyme by the community of synthetic chemists, a simple desktop model was devised to assess the likelihood of a specific substrate being oxidized by MAO-N[69]. Initially, a hypothetical molecule was generated by superimposing all known substrates on top of each other. Next, the model was tested by examining the deracemization of racemic crispine A, an alkaloid found in the welted thistle that has pronounced biological activity. Crispine A was found to map perfectly onto the hypothetical molecule and indeed underwent deracemization using MAO-N and ammonia borane; the racemic starting material was converted to the biologically active (*S*) enantiomer in >95% e.e.

Structure of MAO-N provides insights into directed evolution

The directed evolution experiments described above were carried out in the absence of the structure of MAO-N. Variants were simply selected on the basis of their ability to catalyze oxidation of a particular amine substrate. However, recently the three-dimensional structure of the D5 mutant of MAO-N from *A. niger* was solved, at a resolution of 1.8 Å, offering the prospect of shedding light on the nature of the mutations and how they might affect substrate specificity[70]. The protein crystallized with a molecule of the inhibitor L-proline at the active site, thus allowing the position of the amine substrates to be estimated. **Figure 6** shows the structure with four of the five key mutations found in the D5 variant mapped onto the sequence, and it is clear that these mutations are clustered close to the active site of the protein. Examination of the structure in more detail reveals the presence of two channels, labeled A and B, that pass from the surface of the protein and that intersect at the active site. The mutations at positions 336, 384 and 385 are located within channel A on loops,

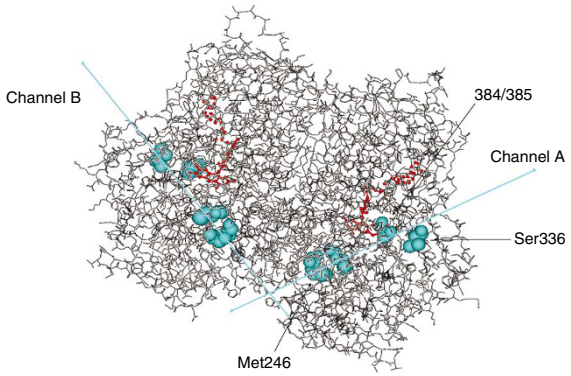

Figure 6 Structure of MAO-N showing the locations of channels A and B and four residues that are important for determining the substrate specificity.

whereas the mutation at position 246 is located within channel B. None of the 4 residues makes direct contact with the substrate; rather, they are located within the second coordination sphere. These results are interpreted as implying that the mutations at positions 336, 384, 385 and 242 in some way affect the dynamics of the two channels and hence have an effect on access of the substrate or release of the product from the active site of MAO-N.

It is interesting to reflect on the outcome of the MAO-N directed evolution work, particularly in terms of what it might teach us for future approaches. A key decision was to screen initially for catalytic activity followed by a secondary check for enantioselectivity. This strategy turned out to be successful in that it led ultimately to the identification of enzymes with high turnover and high (*S*) enantioselectivity. However, in retrospect we may not have been so lucky, and indeed with some enzymes, particularly those with low starting enantioselectivity, it is necessary to devise truly enantioselective screening methods as the primary screen. We were also fortunate to identify enzyme variants with broad substrate specificity, despite using single substrates in the screen—thereby challenging the paradigm that you 'get what you screen for'! Finally, analysis of the mutations in the light of a structure of MAO-N reminded us that protein structures are complex and subtle, particularly with respect to how enzymes interact with their substrates. Even if we had started the evolution program with the structure available to us, it is highly unlikely that we would have rationally identified the mutations that ultimately turned out to be important. Thus directed evolution not only provides us with superior biocatalysts but also teaches us about the way in which enzymes function and control substrate specificity.

Directed evolution provides a range of superior biocatalysts

Using the principles of directed evolution outlined above by the combination of random mutagenesis and high-throughput screening, many other enzymes have been evolved in an analogous fashion. In many cases, particularly for industrial applications, the overall aim of all of these projects is to develop robust and highly selective biocatalysts for applications in synthetic organic chemistry. The chiral 1,3-diol–containing side chain of the cholesterol-lowering statin drugs (for example, atorvastatin and rosuvastatin) has been a major target for biocatalysis, thus leading to the development of a number of different industrial processes that use one or more biocatalysts in the key step. One of the most elegant approaches developed involves the use of deoxyribose aldolase (DERA) to catalyze the combination of two molecules of acetaldehyde and one molecule of chloroacetaldehyde to give a product with two asymmetric centers[71]. Both carbon-carbon bonds are formed with very high levels of enantio- and diastereoselectivity (**Fig. 3**). The wild-type enzyme was found to be inhibited by the chloroacetaldehyde and hence was subjected to rounds of directed evolution to improve its process stability. Other groups have evolved aldolases[72,73] with altered reaction stereochemistry, including a particularly impressive example by Williams *et al.*[74] with N-acetylneuraminic acid aldolase. The aim of their work was to generate aldolase variants with broader substrate specificity with respect to the electrophilic aldehyde substrate. Interestingly, they chose to screen for variants that were able to catalyze the reverse reaction—retro-aldol cleavage of the substrate. Variants were identified that not only had broader specificity but that were also able to generate aldol products with both (*R*) and (*S*) configuration. Transketolase also catalyzes carbon-carbon bond formation, and the wild-type enzyme from *E. coli* exhibits high enantioselectivity (e.e. = 95%) toward glycolaldehyde but low stereoselectivity toward propanal (e.e. = 58%). Libraries obtained by saturation mutagenesis of His26, Asp469 and His100 were generated to enhance enantioselectivity toward propanal. Several mutants were identified, with mutants from Asp469 being (*S*)-selective and mutants from His26 being (*R*)-selective[75].

A number of groups have used directed evolution to improve the properties of the P450 BM3. A variant of P450 BM3 was used to optimize a new screen for epoxygenase activity using styrene as a substrate. The epoxide product reacts with *para*-nitrothiophenolate, causing a measurable decrease in absorption at 405 nm. The sensitivity limit of the assay was shown to be 400 M for styrene oxide. This assay was used to detect new, improved variants of BM3[76]. Epoxide hydrolases catalyze the enantioselective hydrolysis of racemic chiral epoxides, which are valuable building blocks in organic synthesis. Directed evolution has been used to improve the enantioselectivity of these enzymes, including a detailed analysis of how the selectivity changes as a function of the individual mutations[77]. Optically pure cyanohydrins can be prepared in either (*S*) or (*R*) form using evolved hydroxynitrile lyases[78]. Among the most powerful of nature's biocatalysts are the alkane hydroxylases[79] and P450 monooxygenases[80], which are able to catalyze reactions that are difficult to emulate using conventional catalysis. Directed evolution has been used to change the substrate specificity of galactose oxidase to generate mutants that are able to enantioselectively oxidize racemic 1-phenylethanol[81]. The wild-type galactose oxidase enzyme has very high activity but a rather restricted substrate specificity, as evidenced by the fact that it has $>10^6$-fold higher activity toward galactose compared with glucose, which only differs in structure by the stereochemistry of the C4 hydroxyl group. These evolved galactose oxidase variants have also been used for the rapid screening of activity and enantioselectivity of ketoreductase enzymes. In this work, the activity of the ketoreductases is determined spectrophotometrically by monitoring NADH consumption, and simultaneously the e.e. of the product alcohol is determined by using an enantioselective alcohol oxidase[82].

Hybrid enzymes with new catalytic function

A promising alternative to generating enzymes with new function is to generate artificial or hybrid enzymes by inserting transition metal catalysts into the active sites of proteins[83–89]. The transitional metal catalyst provides the essential chemistry and first coordination sphere, whereas the protein provides additional interactions via a second coordination sphere. Initial success in this field came in the generation of hybrid catalysts for asymmetric hydrogenation of olefins[83–86]. However, the scope has been broadened recently to include asymmetric sulfoxidation[87,88] and allylic alkylation[89]. Klein

et al.[83] have applied the principles of directed evolution to optimize the hybrid catalysts by means of chemogenetic protein engineering. Carbonic anhydrase has also been used as a host protein. Site-directed mutagenesis and combined site-directed mutagenesis with chemical modification were used to bind rhodium to the active site of a carbonic anhydrase with minimal binding to the surface. Hydrogenation and isomerization of *cis*-stilbene using rhodium(I)-bound carbonic anhydrase was found to have high selectivity (20:1) for *cis*-stilbenes over *trans*-stilbenes[86].

Summary and future perspectives

Directed evolution of enzymes has, in a relatively short period of time, made a rapid transition from being a new tool for studying the relationship between sequence and function to being an extremely useful and efficient method for optimizing biocatalysts for practical applications. Directed evolution is also being extended beyond single-enzyme systems to embrace metabolic pathways[90], including carotenoid[91] biosynthesis and enzymes involved in the production of complex natural products via assembly lines[92]. Inevitably, the fields of rational, semirational and random redesign of enzymes are moving ever closer together, such that most directed evolution programs make use, where possible, of structural data in order to inform the library generation processes[93,94].

The key to a directed evolution program is the development of a reliable high-throughput screen, coupled with effective methods for generating libraries that are rich in variants containing new function, and in this respect researchers will continue to develop new protocols for library generation and high-throughput screening. Some reaction types are notoriously difficult to screen for in a general sense—for example, the oxidation of C-H bonds to give C-OH—and yet this transformation is one of the most important reactions in synthesis catalyzed by an enzyme. The further development of algorithms that can understand the relationship between sequence space and function will be extremely valuable. For the present, the *de novo* design of enzymes from scaffolds appears highly challenging but in the medium term may offer the prospect of generating truly new functions that will then serve as highly valuable starting points for further optimization when coupled with directed evolution. Exploitation of enzyme promiscuity will also emerge as a major theme contributing to the development of new biocatalyst function. Many roads lead to Rome, but all have in common the objective of understanding the factors involved in enzyme evolution and exploiting this knowledge to develop biocatalysts with exquisite properties for future applications.

Published online at http://www.nature.com/naturechemicalbiology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Schoemaker, H. *et al.* Dispelling the myths – biocatalysis in industrial synthesis. *Science* **299**, 1694–1697 (2003).
2. Patel, R.N. Chemo-enzymatic synthesis of pharmaceutical intermediates. *Expert Opin. Drug Discov.* **3**, 187–245 (2008).
3. Pollard, D.J. & Woodley, J.M. Biocatalysis for pharmaceutical intermediates: the future is now. *Trends Biotechnol.* **25**, 66–73 (2007).
4. Panke, S., Held, M. & Wubbolts, M. Trends and innovations in industrial biocatalysis for the production of fine chemicals. *Curr. Opin. Biotechnol.* **15**, 272–279 (2004).
5. Fox, R.J. & Clay, M.D. Catalytic effectiveness, a measure of enzyme proficiency for industrial applications. *Trends Biotechnol.* **27**, 137–140 (2009).
6. Stemmer, W.P. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391 (1994).
7. Crameri, A. *et al.* DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291 (1998).
8. Moore, J.C. & Arnold, F.H. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467 (1996).
9. Reetz, M.T., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K.-E. Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew. Chem. Int. Ed.* **36**, 2830–2832 (1997).
10. Arnold, F.H. Directed evolution: creating biocatalysts for the future. *Chem. Eng. Sci.* **51**, 5091–5102 (1996).
11. Tracewell, C.A. & Arnold, F.H. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr. Opin. Chem. Biol.* **13**, 3–9 (2009).
12. Alexeeva, M., Carr, R. & Turner, N.J. Directed evolution of enzymes: new biocatalysts for asymmetric synthesis. *Org. Biomol. Chem.* **1**, 4133–4137 (2003).
13. Bershtein, S. & Tawfik, D.S. Advances in laboratory evolution of enzymes. *Curr. Opin. Chem. Biol.* **12**, 151–158 (2008).
14. Turner, N.J. Directed evolution of enzymes new biocatalysts for organic synthesis. *Chim. Oggi* **26**, 9–10 (2008).
15. Reetz, M.T. Directed evolution of enzymes for asymmetric syntheses. in *Asymmetric Synthesis* (eds. Christmann, M. & Bräse, S.) 207–211 (Wiley-VCH, Weinheim, Germany, 2007).
16. Johannes, T.W. & Zhao, H. Directed evolution of enzymes and biosynthetic pathways. *Curr. Opin. Microbiol.* **9**, 261–267 (2006).
17. Sylvestre, J., Chautard, H., Cedrone, F. & Delcourt, M. Directed evolution of biocatalysts. *Org. Process Res. Dev.* **10**, 562–571 (2006).
18. Hibbert, E.G. *et al.* Directed evolution of biocatalytic processes. *Biomol. Eng.* **22**, 11–19 (2005).
19. Valetti, F. & Gilardi, G. Directed evolution of enzymes for product chemistry. *Nat. Prod. Rep.* **21**, 490–511 (2004).
20. Lutz, S. & Patrick, W.M. Novel methods for directed evolution of enzymes: quality, not quantity. *Curr. Opin. Biotechnol.* **15**, 291–297 (2004).
21. Montiel, C. & Bustos-Jaimes, I. Trends and challenges in directed evolution. *Curr. Chem. Biol.* **2**, 50–59 (2008).
22. Fox, R.J. *et al.* Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344 (2007).
23. Fox, R.J. & Huisman, G.W. Enzyme optimization: moving from blind evolution to statistical exploration of sequence-function space. *Trends Biotechnol.* **26**, 132–138 (2008).
24. Grate, J. Directed evolution of three biocatalysts to produce the key chiral building block for atorvastatin, the active ingredient in Lipitor. *United States Environmental Protection Agency* <http://www.epa.gov/greenchemistry/pubs/pgcc/winners/grca06.html > (2006).
25. Park, S. *et al.* Focusing mutations into the *P. fluorescens* esterase binding site increases enantioselectivity more effectively than distant mutations. *Chem. Biol.* **12**, 45–54 (2005).
26. Morley, K.L. & Kazlauskas, R.J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **23**, 231–237 (2005).
27. Reetz, M.T., Kahakeaw, D. & Lohmer, R. Addressing the numbers problem in directed evolution. *ChemBioChem* **9**, 1797–1804 (2008).
28. Reetz, M.T., Wang, L.-W. & Bocola, M. Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. *Angew. Chem. Int. Ed.* **45**, 1236–1241 (2006).
29. Reetz, M.T. *et al.* Expanding the substrate scope of enzymes: combining mutations obtained by cASTing. *Chem. Eur. J.* **12**, 6031–6038 (2006).
30. Muñoz, E. & Deem, M.W. Amino acid alphabet size in protein evolution experiments: better to search a small library thoroughly or a large library sparsely? *Protein Eng. Des. Sel.* **21**, 311–317 (2008).
31. Reetz, M.T. & Wu, S. Greatly reduced amino acid alphabets in directed evolution: making the right choice for saturation mutagenesis at homologous enzyme positions. *Chem. Commun. (Camb.)* 5499–5501 (2008).
32. Hult, K. & Berglund, P. Enzyme promiscuity: mechanism and applications. *Trends Biotechnol.* **25**, 231–238 (2007).
33. Taglieber, A. *et al.* Alternate-site enzyme promiscuity. *Angew. Chem. Int. Ed.* **46**, 8597–8600 (2007).
34. Bornscheuer, U.T. & Kazlauskas, R.J. Reaction specificity of enzymes: catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed.* **43**, 6032–6040 (2004).
35. Peisajovich, S.G. & Tawfik, D.S. Protein engineers turned evolutionists. *Nat. Methods* **4**, 991–994 (2007).
36. Bershtein, S., Goldin, K. & Tawfik, D.S. Intense neutral drifts yield robust and evolvable consensus proteins. *J. Mol. Biol.* **379**, 1029–1044 (2008).
37. Aharoni, A. *et al.* The 'evolvability' of promiscuous protein functions. *Nat. Genet.* **37**, 73–76 (2004).
38. Gupta, R.D. & Tawfik, D.S. Directed enzyme evolution via small and effective neutral drift libraries. *Nat. Methods* **5**, 939–942 (2008).
39. Bloom, J.D., Romero, P.A., Lu, Z. & Arnold, F.H. Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. *Biol. Direct* **2**, 1–19 (2007).
40. Sakai, A. *et al.* Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of cis,cis-muconate lactonizing enzymes. *Biochemistry* **48**, 1445–1453 (2009).
41. Grogan, G. Emergent mechanistic diversity of enzyme-catalysed β-diketone cleavage. *Biochem. J.* **388**, 721–730 (2005).
42. Hamed, R.B., Batchelar, E.T., Clifton, I.J. & Schofield, C.J. Mechanisms and structures of crotonase superfamily enzymes-how nature controls enolate and oxyanion reactivity. *Cell. Mol. Life Sci.* **65**, 2507–2527 (2008).
43. Hasnaoui-Dijoux, G., Majerić Elenkov, M., Lutje Spelberg, J.H., Hauer, B. & Janssen, D.B. Catalytic promiscuity of halohydrin dehalogenase and its application in enantioselective epoxide ring opening. *ChemBioChem* **9**, 1048–1051 (2008).

44. Terao, Y., Miyamoto, K. & Ohta, H. Introduction of single mutation changes arylmalonate decarboxylase to racemase. *Chem. Commun. (Camb.)* 3600–3602 (2006).
45. Seebeck, F.P., Guainazzi, A., Amoreira, C., Baldridge, K.K. & Hilvert, D. Stereoselectivity and expanded substrate scope of an engineered PLP-dependent aldolase. *Angew. Chem. Int. Ed.* **45**, 6824–6826 (2006).
46. Jochens, H. *et al.* Converting an esterase into an epoxide hydrolase. *Angew. Chem. Int. Ed.* **48**, 3532–3535 (2009).
47. Keefe, A.D. & Szostak, J. Functional proteins from a random sequence library. *Nature* **410**, 715–718 (2001).
48. Park, H.S. *et al.* Design and evolution of new catalytic activity with an existing protein scaffold. *Science* **311**, 535–538 (2006).
49. Röthlisberger, D. *et al*. Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
50. Jiang, L. *et al.* De novo computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008).
51. Sterner, R., Merkl, R. & Raushel, F.M. Computational design of enzymes. *Chem. Biol.* **15**, 421–423 (2008).
52. Damborsky, J. & Brezovsky, J. Computational tools for designing and engineering biocatalysts. *Curr. Opin. Chem. Biol.* **13**, 26–34 (2009).
53. Smith, A.J.T. *et al.* Structural reorganization and preorganization in enzyme active sites: comparisons of experimental and theoretically ideal active site geometries in the multistep serine esterase reaction cycle. *J. Am. Chem. Soc.* **130**, 15361–15373 (2008).
54. Seelig, B. & Szostak, J.W. Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **448**, 828–831 (2007).
55. Leemhuis, H., Kelly, R.M. & Dijkhuizen, L. Directed evolution of enzymes: library screening strategies. *IUBMB Life* **61**, 222–228 (2009).
56. Belder, D., Ludwig, M., Wang, L.-W. & Reetz, M.T. Enantioselective catalysis and analysis on a chip. *Angew. Chem. Int. Ed.* **45**, 2463–2466 (2006).
57. Mastrobattista, E. *et al.* High-throughput screening of enzyme libraries: in vitro evolution of a β-galactosidase by fluorescence-activated sorting of double emulsions. *Chem. Biol.* **12**, 1291–1300 (2005).
58. Griffiths, A.D. & Tawfik, D.S. Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J.* **22**, 24–35 (2003).
59. Fernandez-Gacio, A., Uguen, M. & Fastrez, J. Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol.* **21**, 408–414 (2003).
60. Lipovsek, D. *et al.* Selection of horseradish peroxidase variants with enhanced enantioselectivity by yeast surface display. *Chem. Biol.* **14**, 1176–1185 (2007).
61. Qian, Z., Fields, C.J. & Lutz, S. Investigating the structural and functional consequences of circular permutation on lipase B from *Candida antarctica*. *ChemBioChem* **8**, 1989–1996 (2007).
62. Enright, A. *et al.* Stereoinversion of β- and γ-substituted-α-amino acids using a chemoenzymatic oxidation-reduction procedure. *Chem. Commun. (Camb.)* 2636–2637 (2003).
63. Roff, G.J., Lloyd, R.C. & Turner, N.J. A versatile chemo-enzymatic route to enantiomerically pure β-branched-α-amino acids. *J. Am. Chem. Soc.* **126**, 4098–4099 (2004).
64. Alexeeva, M., Enright, A., Dawson, M.J., Mahmoudian, M. & Turner, N.J. Deracemisation of α-methylbenzylamine using an enzyme obtained by *in vitro* evolution. *Angew. Chem. Int. Ed.* **41**, 3177–3180 (2002).
65. Carr, R. *et al.* Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity. *Angew. Chem. Int. Ed.* **42**, 4807–4810 (2003).
66. Carr, R. *et al.* Directed evolution of an amine oxidase for the preparative deracemisation of cyclic secondary amines. *ChemBioChem* **6**, 637–639 (2005).
67. Dunsmore, C.J., Carr, R., Fleming, T. & Turner, N.J. A chemo-enzymatic route to enantiomerically pure cyclic tertiary amines. *J. Am. Chem. Soc.* **128**, 2224–2225 (2006).
68. Eve, T.S.C., Wells, A.S. & Turner, N.J. Enantioselective oxidation of *O*-methyl-*N*-hydroxylamines using MAO-N as catalyst. *Chem. Commun. (Camb.)* 1530–1531 (2007).
69. Bailey, K.R., Ellis, A.J., Reiss, R., Snape, T.J. & Turner, N.J. A template-based mnemonic for monoamine oxidase (MAO-N) catalyzed reactions and its application to the chemo-enzymatic deracemisation of the alkaloid (±)-crispine A. *Chem. Commun. (Camb.)* 3640–3642 (2007).
70. Atkin, K.E. *et al.* The structure of monoamine oxidase from *Aspergillus niger* provides a molecular context for improvements in activity obtained by directed evolution. *J. Mol. Biol.* **384**, 1218–1231 (2008).
71. Jennewein, S. *et al*. Directed evolution of an industrial biocatalyst: 2-deoxy-D-ribose 5-phosphate aldolase. *Biotechnol. J.* **1**, 537–548 (2006).
72. Ran, N. & Frost, J.W. Directed evolution of 2-keto-3-deoxy-6-phosphogalactonate aldolase to replace 3-deoxy-D-arabino-heptulosonic acid 7-phosphate synthase. *J. Am. Chem. Soc.* **129**, 6130–6139 (2007).
73. Hsu, C.-C., Hong, Z., Wada, M., Franke, D. & Wong, C.-H. Directed evolution of D-sialic acid aldolase to L-3-deoxy-manno-2-octulosonic acid (L-KDO) aldolase. *Proc. Natl. Acad. Sci. USA* **102**, 9122–9126 (2005).
74. Williams, G.J., Woodhall, T., Farnsworth, L.M., Nelson, A. & Berry, A. Creation of a pair of stereochemically complementary biocatalysts. *J. Am. Chem. Soc.* **128**, 16238–16247 (2006).
75. Smith, M.E.B., Hibbert, E.G., Jones, A.B., Dalby, P.A. & Hailes, H.C. Enhancing and reversing the stereoselectivity of *Escherichia coli* transketolase via single-point mutations. *Adv. Synth. Catal.* **350**, 2631–2638 (2008).
76. Tee, K.L. & Schwaneberg, U. A screening system for the directed evolution of epoxygenases: importance of position 184 in P450 BM3 for stereoselective styrene epoxidation. *Angew. Chem. Int. Ed.* **45**, 5380–5383 (2006).
77. Reetz, M.T. *et al.* Directed evolution of an enantioselective epoxide hydrolase: uncovering the source of enantioselectivity at each evolutionary stage. *J. Am. Chem. Soc.* **131**, 7334–7343 (2009).
78. Liu, Z. *et al.* Laboratory evolved biocatalysts for stereoselective syntheses of substituted benzaldehyde cyanohydrins. *ChemBioChem* **9**, 58–61 (2008).
79. Koch, D.J., Chen, M.M., van Beilen, J.B. & Arnold, F.H. *In vivo* evolution of butane oxidation by terminal alkane hydroxylases AlkB and CYP153A6. *Appl. Environ. Microbiol.* **75**, 337–344 (2009).
80. Landwehr, M. *et al.* Enantioselective alpha-hydroxylation of 2-arylacetic acid derivatives and buspirone catalyzed by engineered cytochrome P450 BM-3. *J. Am. Chem. Soc.* **128**, 6058–6059 (2006).
81. Escalettes, F. & Turner, N.J. Directed evolution of galactose oxidase: generation of enantioselective secondary alcohol oxidases. *ChemBioChem* **9**, 857–860 (2008).
82. Truppo, M.D., Escalettes, F. & Turner, N.J. Rapid determination of both the activity and enantioselectivity of ketoreductases. *Angew. Chem. Int. Ed.* **49**, 2639–2641 (2008).
83. Klein, G. *et al.* Tailoring the active site of chemzymes by using a chemogenetic-optimization procedure: towards substrate-specific artificial hydrogenases based on the biotin-avidin technology. *Angew. Chem. Int. Ed.* **44**, 7764–7767 (2005).
84. Ward, T.R. Artificial enzymes made to order: combination of computational design and directed evolution. *Angew. Chem. Int. Ed.* **47**, 7802–7803 (2008).
85. Reetz, M.T., Peyralans, J.J.-P., Maichele, A., Fu, Y. & Maywald, M. Directed evolution of hybrid enzymes: evolving enantioselectivity of an achiral Rh-complex anchored to a protein. *Chem. Commun. (Camb.)* 4318–4320 (2006).
86. Jing, Q., Okrasa, K. & Kazlauskas, R.J. Stereoselective hydrogenation of olefins using rhodium-substituted carbonic anhydrase - a new reductase. *Chem. Eur. J.* **15**, 1370–1376 (2009).
87. Pordea, A. *et al.* Artificial metalloenzyme for enantioselective sulfoxidation based on vanadyl-loaded streptavidin. *J. Am. Chem. Soc.* **130**, 8085–8088 (2008).
88. Rousselot-Pailley, P. *et al.* The protein environment drives selectivity for sulfide oxidation by an artificial metalloenzyme. *ChemBioChem* **10**, 545 (2009).
89. Pierron, J. *et al.* Artificial metalloenzymes for asymmetric allylic alkylation on the basis of the biotin-avidin technology. *Angew. Chem. Int. Ed.* **47**, 701–705 (2008).
90. Umeno, D., Tobias, A.V. & Arnold, F.H. Diversifying carotenoid biosynthetic pathways by directed evolution. *Microbiol. Mol. Biol. Rev.* **69**, 51–78 (2005).
91. Chatterjee, R. & Yuan, L. Directed evolution of metabolic pathways. *Trends Biotechnol.* **24**, 28–38 (2006).
92. Fischbach, M.A., Lai, J.R., Roche, E.D., Walsh, C.T. & Liu, D.R. Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proc. Natl. Acad. Sci. USA* **104**, 11951–11956 (2007).
93. Chica, R.A., Doucet, N. & Pelletier, J.N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **16**, 378–384 (2005).
94. Kazlauskas, R.J. & Lutz, S. Engineering enzymes by intelligent design. *Curr. Opin. Chem. Biol.* **13**, 1–2 (2009).