Directed evolution of enzyme catalysts

Olga Kuchner and Frances H. Arnold

Directed enzyme evolution has emerged in the past few years as a powerful alternative to rational approaches for engineering biocatalysts. Prerequisites for successful directed evolution are functional expression in a suitable microbial host, a rapid screen for the desired feature(s) and a well-thought-out working strategy for navigating protein landscapes. The rapidly growing body of literature on enzyme evolution *in vitro* includes techniques for creating and searching combinatorial enzyme libraries, as well as several successful examples of different evolutionary strategies being used.

Enzymes' efficiency and precision are unmatched by most conventional industrial catalysts. However, naturally occurring enzymes often lack features necessary for commercial applications. Although protein chemists continue to elucidate the relationships between the sequence, structure and function of proteins, the extensive knowledge that is necessary for the application of rational engineering approaches is available for only a tiny fraction of known enzymes. Directed evolution has proved to be useful for modifying enzymes in the absence of such knowledge.

Directed enzyme evolution generally begins with the creation of a library of mutated genes. Gene products that show improvement with respect to the desired property or set of properties are identified by selection or screening, and the gene(s) encoding those enzymes are subjected to further cycles of mutation and screening in order to accumulate beneficial mutations. This evolution can involve few or many generations, depending on how far one wishes to progress and the effects of mutations observed in each generation.

The main requirements for successful directed evolution are (1) the functional expression of the enzyme in a suitable microbial host, (2) the availability of a screen (or selection) sensitive to the desired properties and (3) identifying a workable evolution strategy. The vast majority of possible evolutionary paths lead to poorer enzymes; the strategic challenge is to identify a path that will result in the improvement of the desired feature(s). The number of possible variants increases rapidly with the size of the enzyme and with the number of amino acids that are allowed to vary simultaneously (Box 1). Even for a small protein, an impossibly large number of variants can be generated when multiple mutations are introduced. Because most mutations are deleterious[1,2], the chances of identifying improved enzymes in libraries containing large numbers of mutations are very small indeed. Therefore, the mutation rate must be tuned to the power of the screen or selection[3]. We will comment on some strategies for creating and searching libraries of enzyme variants. In addition, selected theoretical and experimental investigations of protein 'fitness landscapes' that may be applied to optimize these strategies will be introduced.

Interest in engineering enzymes by directed evolution has grown significantly in the past few years. Table 1 summarizes some key examples of directed enzyme evolution in which multiple generations of mutagenesis and screening or selection have been performed[4–13].

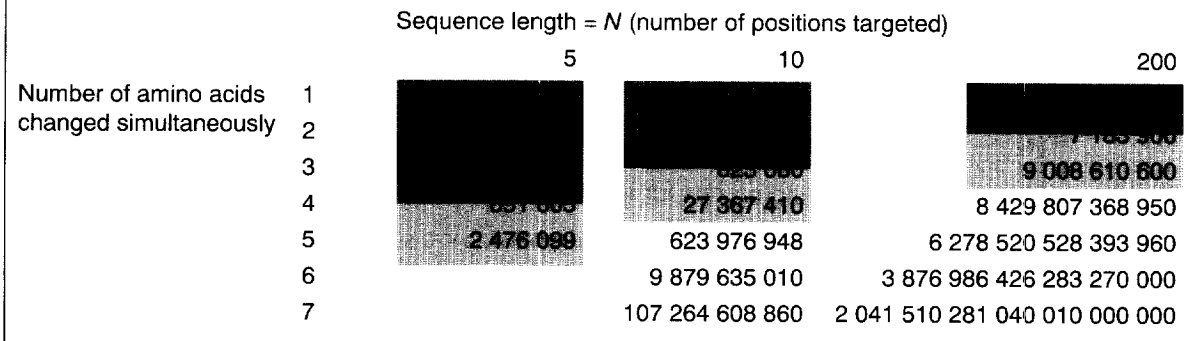
The evolution of thermostability

Improvements in stability, particularly thermostability, are commonly sought by protein engineers. In industrial processes, high temperatures impart such benefits as increased substrate solubility, decreased viscosity of the medium, lower risk of microbial contamination or higher rates of concurrent non-enzymatic reactions. In the large number of random-mutagenesis experiments that have targeted thermostability, single beneficial mutations typically increase the melting temperature (T_m) of the enzyme by 1–2°C[14–18]. Much larger changes are possible, but rare. Pjura *et al.*[15] identified eleven single mutations that increased the thermostability of T4 lysozyme by 0.8 to 1.4°C. Haruki *et al.*[14] found eight mutations that increased the stability of *Escherichia coli* ribonuclease HI; seven impart a ΔT_m between 0.3°C and 4.2°C, while

O. Kuchner and F. H. Arnold (frances@cheme.caltech.edu) are in the Division of Chemistry and Chemical Engineering 210-41, California Institute of Technology, Pasadena, CA 91125, USA.

 0167 – 7799/97/$17.00. PII: S0167-7799(97)01138-4

Box 1. The number of possible variants of a protein that can be created by introducing *M* substitutions simultaneously over *N* amino acids is 19^M [*N*! / (*N* – *M*)! *M*!]



This table lists the number of variants that could be made if every possible amino acid substitution were accessible (for example, by codon-synthesis methods). If point mutations are introduced randomly at the DNA level, the accessible sequence diversity is limited. On average, only 5.7 different amino acids are accessible by single-base-pair changes in a codon. Furthermore, the number of clones that must be screened to see a given number of unique amino-acid-level variants must increase to account for wild-type variants arising from the degeneracy of the genetic code. To have 95% confidence that a given number of amino-acid-level variants has been sampled in a screen or selection, the number of unique clones must be multiplied by a factor of approximately ten.

The feasible upper limit for screening with a high-throughput assay is indicated by the darker gray area. The upper limit of feasibility for selection in microbial hosts is indicated by lighter gray. Neither screening nor selection allow exhaustive searches of libraries containing multiple mutations distributed over long sequences.

one amino acid substitution (which is thought to fill a cavity in the protein's interior) yielded a full 7.8°C increase. T_m values reported for double and triple mutants[19,20] appear consistent with the generalization that the free-energy effects of individual stabilizing mutations are roughly additive[21].

Significantly larger increases in thermostability resulting from single amino acid substitutions have also been reported. Joyet *et al.*[22] combined a stabilizing mutation they had identified using sequence comparisons between mesophilic and thermophilic versions of α-amylases with a new mutation generated by random mutagenesis to raise the T_m of barley α-amylase by 11°C and increase its half-life at 90°C nine- to tenfold. In an early directed-evolution experiment, Liao *et al.*[4] found two amino acid substitutions in the antibiotic-resistance enzyme kanamycin nucleotidyltransferase that resulted in a 15°C increase in T_m. They performed two sequential rounds of mutagenesis and selection by growing *Bacillus stearothermophilus* expressing the kanamycin-nucleotidyltransferase gene (*kan*) in a chemostat and identified one mutation in each round. Remarkably, the same two amino acid substitutions had also been found by Matsumura and Aiba[23], who found the two mutations in a single clone after a single round of random mutagenesis and screening. Stable variants that imparted kanamycin resistance were identified by infecting *E. coli* cells with *kan*-transducing phage and incubating at increasingly higher temperatures.

Table 1. Summary of key directed-enzyme-evolution experiments utilizing sequential generations of random mutagenesis and recombination with selection or screening.

Enzyme	Altered property	Mutagenesis method	Screened or selected	Refs
Kanamycin nucleotidyltransferase	Thermostability	Mutator strain	Selected	4
Subtilisin E	Activity in organic solvents	Error-prone PCR	Screened	5,6
β-Lactamase	Total activity and substrate specificity	DNA shuffling	Selected	7
Subtilisin BPN'	Stability	Cassette	Screened	8
Paranitrobenzyl esterase	Substrate specificity and activity in organic solvents	Error-prone PCR and recombination	Screened	9
Thymidine kinase	Substrate specificity	Cassette	Both	10
β-Galactosidase	Substrate specificity	DNA shuffling	Screened	11
Arsenate detoxification pathway	Arsenic resistance	DNA shuffling	Screened	12
Paranitrobenzyl esterase	Substrate specificity and activity in organic solvents	Error-prone PCR and DNA shuffling	Screened	13

Various strategies have been devised for combating the difficulties encountered in screening highly stable enzymes, which include temperature limitations of the host organisms and the awkwardness of assaying for activity at very high temperatures. Several groups have devised schemes to select for thermostability in thermophilic organisms[4,14,16,24]; some contend that their systems can be readily adapted to the evolution of other proteins, but each of these experiments was performed using an enzyme model system (kanamycin nucleotidyltransferase[4], 3-isopropylmalate dehydrogenase[16,24] or ribonuclease HI[14]) that is highly amenable to selection – that is, the functional enzyme is critical to the survival and growth of the host. Four groups[14,16,22,24] created temperature-sensitive variants of the enzymes they wished to stabilize and subsequently screened for intragenic suppressor mutations that compensated for the loss of stability. This approach offers the advantage that screening can be performed at a lower temperature than that needed to deactivate the wild-type enzyme. However, it poses the risk of identifying mutations that compensate for the initial destabilization but do not confer an advantage to the wild-type sequence. In the experiments reported by Joyet *et al.*[22] and Kotsuka *et al.*[16], the single second-site suppressor mutations also increased the thermostability of the wild-type enzyme. Haruki *et al.*[14] identified eleven second-site suppressor mutations, eight of which were also beneficial in the wild-type enzyme.

Screening may offer a more efficient and generally applicable approach to evolution of thermostability than selection[25]. Using replica plates during screening avoids subjecting host organisms to the temperatures needed to deactivate thermostable enzymes. A typical result of screening paranitrobenzyl-esterase variants for enhanced thermostability is shown in Fig. 1. In this example, a significant fraction of the variants in each 96-well plate appear more thermostable than the parent. Eleven amino acid substitutions, accumulated during sequential mutagenesis and screening, have resulted in a 14°C increase in the esterase T_m (L. Giver, P. O. Freskgard and F. H. Arnold, unpublished).

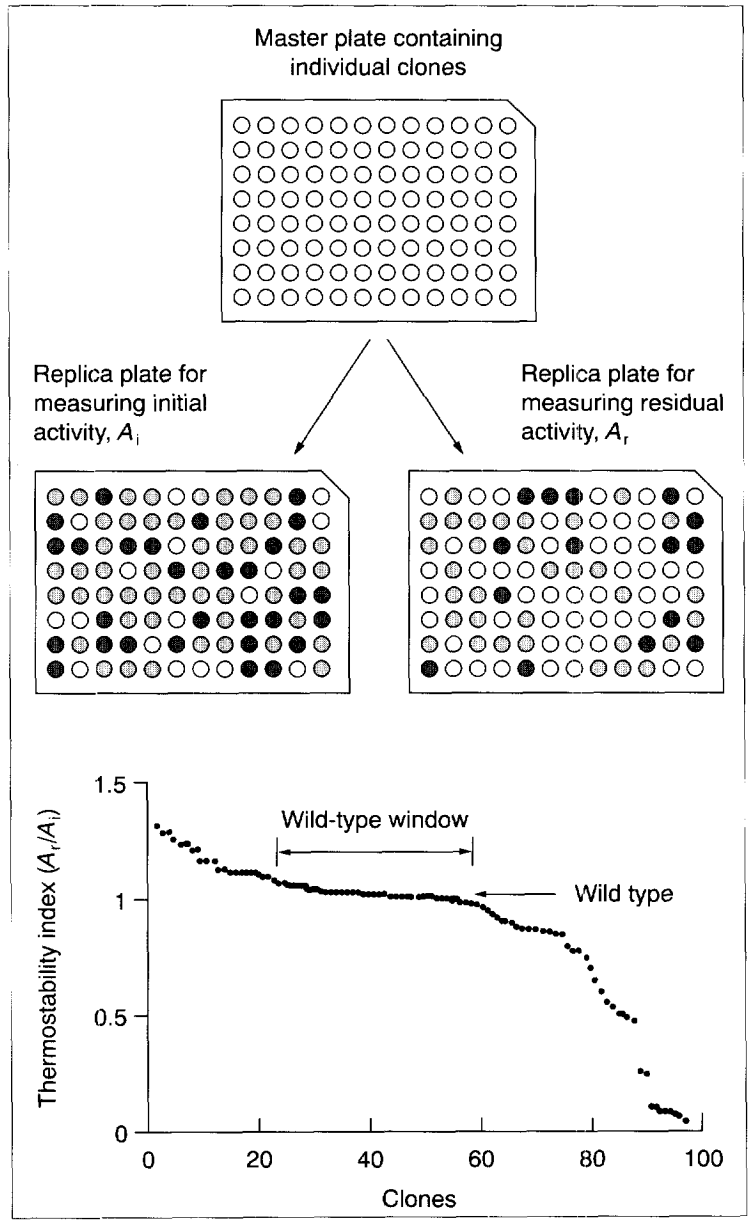


Figure 1

A simple screen for enzyme thermostability based on catalytic activity (*A*) before and after incubation at high temperature. Transformants of a randomly recombined paranitrobenzyl-esterase library were picked and grown in 96-well plates. Initial activity (A_i) and residual activity (A_r) after incubation at the T_m of the parent were measured in two separate replica plates using a microtitre plate reader. A_r/A_i values are sorted and plotted in descending order to show the local fitness profile. Useful parameters derived from this profile include the frequency and fitness of improved clones, the frequency of inactive clones (related to the point-mutagenesis rate) and the size of the wild-type window. Accumulation of mutations over five generations of random mutagenesis and screening has increased the T_m from 52 to 66°C (L. Giver, P. O. Freskgard and F. H. Arnold, unpublished).

The evolution of stability or activity in artificial environments

The use of organic solvents can facilitate solubilization of a substrate or enhance the rate of one reaction over another. Naturally occurring enzymes, however, generally lose activity in organic solvents or aqueous–organic mixtures, even when the native structure remains stable. In an early application of directed evolution, the protease subtilisin E was subjected to sequential generations of random mutagenesis and screening[5] to yield a variant that is active in 60% dimethylformamide (DMF) almost as efficiently as the wild type is in water, an increase of nearly 500 times in the k_{cat}:K_M ratio (Ref. 6). More recently, both sequential random mutagenesis and gene recombination were used to evolve paranitrobenzyl esterase for deprotection of an antibiotic synthetic intermediate in aqueous DMF. The total activity in 30% DMF was increased more than 100 times[9,13].

The first use of random mutagenesis to improve the activity of a mesophilic enzyme at low temperatures was recently reported by Kano *et al.*[26], who screened random variants of subtilisin BPN′ for activity at 10°C. The k_{cat} value of one improved variant was close to that of the wild-type enzyme, and decreased to the same extent as the k_{cat} of the wild-type enzyme in response to a shift from 25°C to 5°C. However, its K_M

value decreased much more sharply than that of the wild type in response to the same temperature shift, resulting in a 10% increase in hydrolytic activity at 10°C and a 30% increase in activity at 1°C. The stability of the evolved enzyme to higher temperature was unchanged. Many enzymes from psychrophilic organisms are more thermolabile than their mesophilic counterparts, and it has been claimed that the flexibility required for activity at cold temperatures inherently limits stability at moderate temperatures. Directed evolution should be able to answer the question of whether psychrophilic enzymes indeed have such intrinsic limitations to thermostability. It may be that these properties have simply not been subjected to evolutionary pressures.

Changing or eliminating an enzyme's requirements for metal ions can be useful in laundry detergents, which commonly contain chelating agents that maintain low free metal-ion concentrations and often use proteases that depend on metal ions for their activity or stability. Bryan and co-workers[8] evolved a subtilisin that is stable in the absence of calcium. After first deleting the enzyme's calcium-binding loop (creating a highly destabilized subtilisin), they randomly mutated ten amino acids that had interacted with that loop and were thus assumed no longer to be optimized for stability in the modified enzyme. The half-life of the evolved subtilisin in low free-calcium concentrations is 12.5 times higher than that of wild-type subtilisin BPN'. Additional rounds of mutagenesis and screening have resulted in an even more stable calcium-independent subtilisin (P. Bryan, pers. commun.).

The evolution of substrate specificity and increased activity on novel substrates

Modifying enzyme substrate specificity represents a critical frontier for adapting enzymes to industrially desirable reactions. A number of random-mutagenesis and directed-evolution experiments have demonstrated that substrate preferences can be evolved. In many of these experiments[10,11,27–30], the increased catalytic efficiency of improved variants resulted from a decrease in K_M. For example, Sidhu and Borgford[29] found variants of *Streptomyces griseus* protease B with dramatically increased activity towards peptides with methionine at the P1 binding site; these variants displayed practically no change in k_{cat}. In other cases, such as the modification of a T7 RNA polymerase to function as a DNA polymerase[31], changes in both k_{cat} and K_M are observed. Beuve and Danchin[27] demonstrated a change in primary substrate specificity of adenylate cyclase by isolating a variant that exhibited a four-and-a-half-fold increase in guanylate-cyclase activity, while retaining only 4% of the wild-type adenylate-cyclase activity. A second mutation restored the wild-type adenylate-cyclase activity and increased guanylate-cyclase activity tenfold relative to the wild type, in effect broadening the enzyme's substrate specificity. An *E. coli* β-galactosidase evolved *in vitro* by Zhang *et al.*[11] to have greater activity towards fucose substrates showed a twofold decrease in k_{cat}, but a 20-fold decrease in K_M.

The directed-evolution experiments reported by Zhang *et al.*[11] and Moore and Arnold[9] are among the very few in which random mutagenesis was not restricted to the amino acids believed to constitute the substrate-binding site. Zhang *et al.* performed seven rounds of DNA shuffling and screening using chromogenic substrates. Six amino acid substitutions were accumulated during the evolution, three of which are located in domain 3 (residues 334–627), which contains most of the substrate-binding pocket. The other three amino acid changes are not located near the binding pocket or active site of the enzyme. The paranitrobenzyl esterase that was evolved for activity towards a novel antibiotic substrate in organic solvents[9] initially displayed only low activity towards this substrate. The increased specific activity achieved by directed evolution reflected adaptation towards the new substrate as well as the solvent environment, and yet none of the mutations accumulated in the evolved enzyme were in direct contact with the substrate: some were as far away as 20 Å. It is clear that limiting the search for possible solutions to residues that line the substrate-binding site will overlook important beneficial mutations.

Two groups have evolved glutathione transferases to create more efficient drug-detoxification enzymes. Gulick and Fahl[32] isolated a variant of glutathione-*S*-transferase that confers ninefold-greater resistance to an alkylating agent used in cancer chemotherapy by selecting for catalytic efficiency. Using phage display to identify glutathione transferases with improved binding to several electrophilic substrates, Widersten and Mannervik[33] failed to isolate any variants more active than the wild-type enzyme. One disadvantage of screening for binding rather than catalysis is that transition-state stabilization, the phenomenological definition of catalysis, involves mechanisms besides transition-state binding. Reinforcing transition-state binding beyond an optimum level may restrict the motion of the enzyme necessary for conformational changes during catalysis, as shown by the small increase in binding affinity that is used for transition-state stabilization compared with the total increase in binding affinity acquired by catalytic antibodies during somatic hypermutation[34].

The evolution of enantioselectivity

Matcham and Bowen[35] reported the use of directed-evolution methods to enhance the enantioselectivity of an *R*-transaminase. Transaminases have been used to produce chiral amines either directly from the precursor ketone or by resolution of the racemic amine through conversion of the amine with incorrect stereochemistry back to the ketone. Most transaminases convert ketones to the corresponding amine with a greater-than-99% enantiomeric excess, but one *S*-selective transaminase converted a β-tetralone to the desired amine with only a 65% enantiomeric excess. 10 000 randomly mutagenized colonies were screened to yield ten variants that increased the enantiomeric excess of this conversion to 80–94%. Some of the

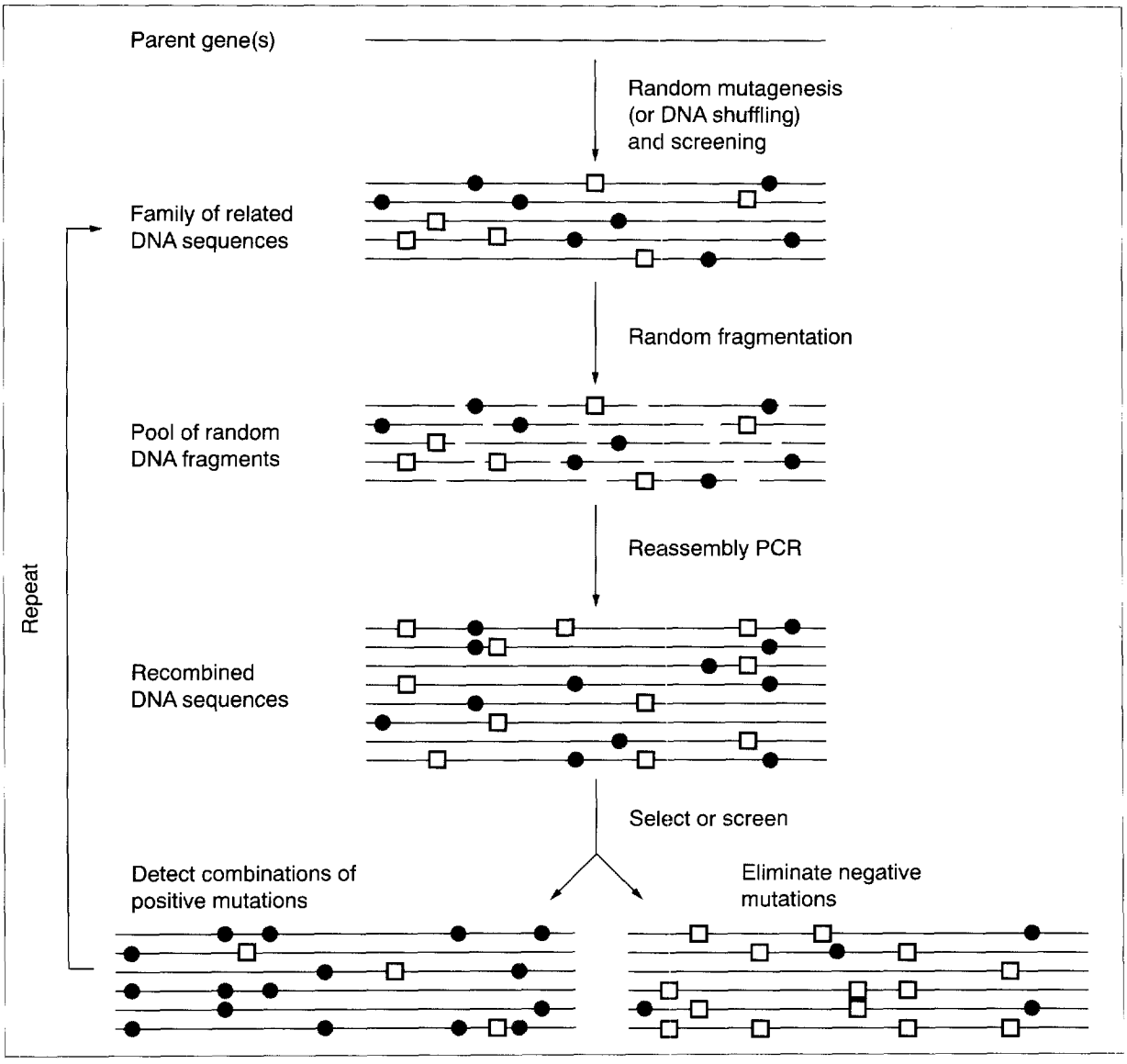


Figure 2

In vitro recombination by DNA shuffling. Random fragments of a single parent gene or pool of genes containing different mutations are reassembled in a PCR reaction, during which they are allowed to act as primers for one another. The fragments elongate, producing a family of DNA sequences containing different combinations of the mutations contained in the parent pool. Beneficial mutations are accumulated and deleterious mutations eliminated during recursive cycles of shuffling and screening or selection (adapted from Ref. 11).

individual mutations resulted in conversions with 85–91% enantiomeric excess.

Techniques and strategies for creating and searching genetic diversity

Efforts to develop and optimize techniques for introducing point mutations randomly into DNA segments facilitated many of the first experiments in directed enzyme evolution. Random mutations have been introduced continuously along a stretch of DNA using error-prone PCR, chemical mutagenesis, UV irradiation, mutator strains, poisoned nucleotides or concomitantly with DNA recombination. Among the methods that introduce mutations randomly along the entire length of a gene, error-prone PCR[36,37] has been characterized the most extensively. The protocol designed by Leung *et al.*[37] provides control of mutagenic rates by varying the concentrations of Mn^{2+} and dNTPs. However, this method exhibits a significant bias for transitions over transversions. The protocol of Joyce and co-workers[36] diminishes this bias but does not eliminate it[2]. Only a fraction of all the possible amino acid substitutions are accessible by point-mutation methods; sequence bias further limits the sequence space.

When a particular set of amino acids has been targeted for mutagenesis based on prior knowledge, region-specific random mutagenesis can be carried out using oligonucleotide-cassette mutagenesis[38,39]. This method is convenient when the targeted amino acids are all in the same region of primary sequence. The groups of Sauer[40], Youvan[41] and Loeb[1] have used this approach extensively to investigate structure–function relationships in enzymes and to probe the information content of amino acid sequences.

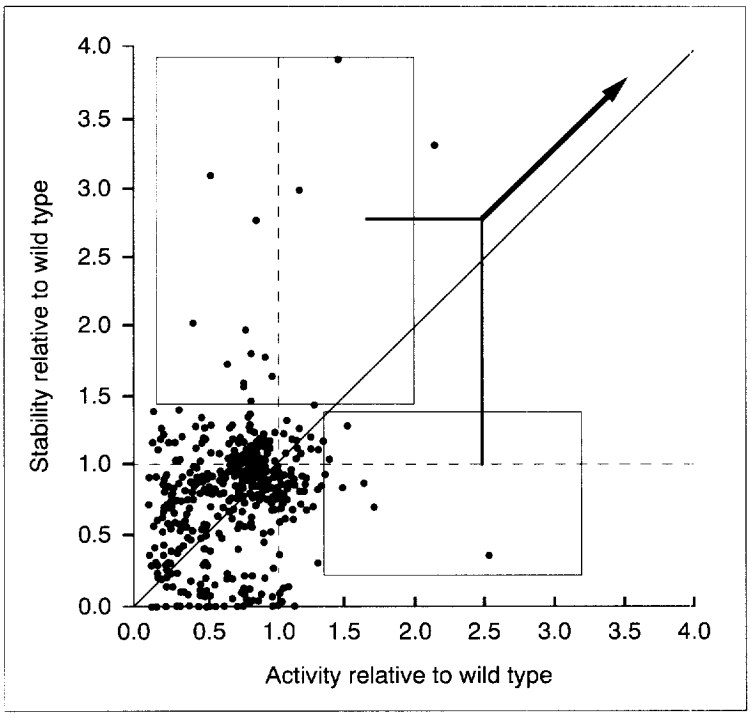


Figure 3

Activity and stability to hydrogen peroxide of a randomly mutagenized enzyme library. Variants exhibiting improvements in single properties are much more frequent than variants with improvements in both properties. The individual populations shown in the boxes can be evolved in parallel and recombined to obtain enzymes that are both more active and more stable.

The DNA-shuffling method developed by Stemmer[42] for random *in vitro* DNA recombination represents a significant advance in our ability to mimic the natural mechanisms of evolution. DNA shuffling (Fig. 2) creates libraries of genes containing combinations of mutations derived from a set of homologous natural sequences or arising as a result of point mutation; minor variations on this method have allowed greater control over the point mutagenesis associated with DNA shuffling[43,44]. A new method for *in vitro* mutagenesis and recombination uses extension of random-sequence oligonucleotides to create the gene fragments for reassembly (Z. Shao, *et al.*, unpublished). The utility of recombination in directed enzyme evolution has been discussed[13,45] and strategies for combining sequential point mutagenesis with recombination explored[13]. Recombination allows the rapid accumulation of beneficial mutations identified in separate genes. Screening protocols, for example, often produce multiple improved sequences. The beneficial mutations normally discarded when a single clone is chosen from these to parent the subsequent generation can be efficiently recombined by DNA shuffling. In addition, delerious mutations are removed.

Recombination is also a promising approach to combining two or more improved properties evolved separately. This concept is illustrated in Fig. 3, which shows the results of screening a randomly mutagenized enzyme library (e.g. of subtilisin) for two separate properties, activity towards a particular substrate and stability towards hydrogen peroxide. Although few variants exhibit improvements in both properties, several are either more stable or more active (sometimes at the cost of the second property). These populations can be recombined in order to create enzymes that are both more stable and more active. With efficient *in vitro* recombination methods, such as DNA shuffling, it should be possible to combine several desired features, starting from libraries of enzymes evolved separately for individual properties.

Delagrave *et al.*[41] applied combinatorial cassette mutagenesis in an approach they call recursive combinatorial mutagenesis. This algorithm uses information about functionally conserved residues learned during each round of combinatorial cassette mutagenesis to guide the next round of oligonucleotide-cassette synthesis. The objective of sampling combinations of several amino acid substitutions at once is similar to the goal of DNA shuffling[42]. In combinatorial cassette mutagenesis, random cassettes are created to replace a limited number of contiguous amino acids, whereas DNA shuffling is appropriate for mutagenizing and recombining much longer sequences.

The decision whether to restrict mutagenesis to a target set of residues can be a difficult one. Limiting the number of amino acids to be mutated greatly reduces the sequence space that must be searched (Box 1), allowing the sampling of multiple simultaneous substitutions in key regions of the enzyme. However, the fact that mutations far from the active site or binding pocket modulate activity, substrate recognition or binding[9,11] underscores the risks inherent in restricting the regions undergoing evolution.

The array of available screening and selection methods has also expanded with the proliferation of directed-evolution experiments. Screening[25] and selection[46] approaches useful for directed enzyme evolution have recently been reviewed. Selections are attractive for searching larger libraries of variants, but are difficult to devise for enzymes that are not critical to the survival of the host organism. Furthermore, organisms are notoriously adept at evading imposed selective pressures by unexpected mechanisms. Less-stringent functional complementation can be very useful for identifying variants that retain biological activity in libraries generated using relatively high mutagenic rates[1,2,25]. In choosing an assay or selection, it is useful to remember the first law of random mutagenesis: you get what you screen for[6].

Local fitness landscapes

A local fitness landscape is a representation of the fitness – in this context, the extent to which an enzyme exhibits a desired property – of not-too-distant sequence variants of an enzyme. Such landscapes are often greatly simplified for depiction in two or three dimensions (Box 2). (The term 'fitness landscape' often refers to the fitness values of all possible sequences[47,48]. Such a landscape is clearly not accessible by any experimental measures.) Although the

shape of local protein-fitness landscapes is subject to debate[49], the high dimensionality of sequence space, the near-additivity of the free energies of individual mutations for many properties[21] and statistical models[50] all suggest that local landscapes are more 'Fujiyama-like' (i.e. increase more-or-less monotonically towards a fitness optimum) than highly rugged.

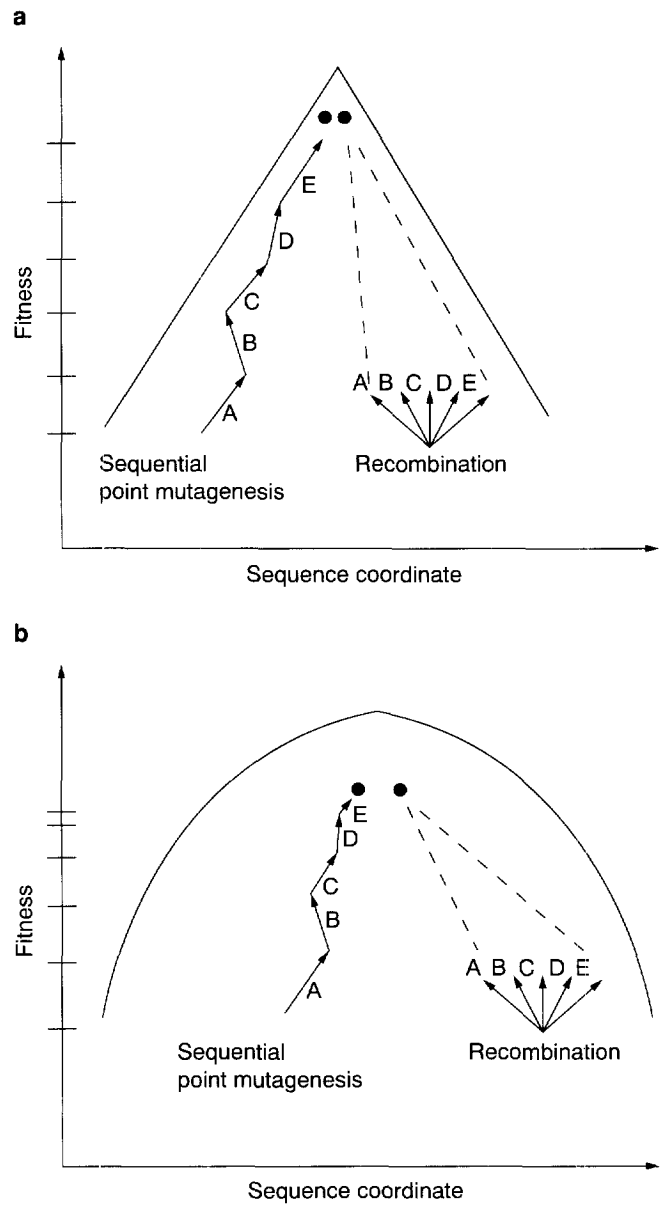
Several groups have attempted to bring the concept of protein landscapes to bear on practical issues of devising strategies for directed evolution. Urabe and coworkers[51] created a library of versions of *B. stearothermophilus* catalase I using an average mutagenic rate of one mutation per gene and characterized 82 of these variants for fitness with respect to their activity as a catalase, activity as a peroxidase and thermostability. The profiles generated by plotting each of these characteristics from best to worst were termed 'single-mutant-neighbour' fitness landscapes. Such simplified fitness profiles are very useful for evaluating the feasibility of a particular evolutionary experiment. For example, the position of the fitness of the wild-type or parent enzyme relative to the remainder of the profile provides an indication of the 'evolvability' of a particular trait. The fact that the esterase profile in Fig. 1 shows a number of mutants more stable than the wild-type enzyme suggests that thermostability in this system is highly amenable to directed evolution. Urabe and coworkers reached a similar conclusion for *B. stearothermophilus* catalase I. Parameters such as the proportion of clones that are inactive, easily obtained from fitness profiles, are also very useful for estimating mutagenic rates[2] and optimizing mutagenesis conditions.

Aita and Husimi[50] built a model protein landscape based on the additivity of the free-energy effects of amino acid substitutions and the introduction of functions to describe the tolerance to amino acid substitution at each site. Such tolerance functions have been measured experimentally by Delagrave *et al.*[41] and Suzuki *et al.*[1] Although an *a priori* assumption of additivity automatically prejudices the landscape to be Fujiyama-like, the methods developed by Aita and Husimi[50] allow shape parameters such as the slope of the landscape or the fitness of the global optimum to be estimated from local fitness profiles. Different evolutionary strategies, such as sequential point mutagenesis and recombination, can be depicted on the simplified landscape (Box 2).

Future directions

The first *in vitro* evolution of a metabolic pathway was recently reported by Crameri *et al.*[12], who evolved a three-gene operon coding for arsenic resistance. Several rounds of DNA shuffling were used to achieve a 40-fold increase in arsenate detoxification in the absence of any structural information or knowledge of which gene products are rate-limiting. Extending the power of directed evolution to groups of genes offers a glimpse of exciting future avenues for enzyme and metabolic engineering. A trend toward further miniaturization of assay formats and higher-throughput automation of screening promises to increase the libary

Box 2. A depiction of two 'Fujiyama-type' local protein-fitness landscapes with different slope characteristics



Fitness, for our purposes, is the extent to which the enzyme exhibits a desired characteristic. Beneficial mutations (A–E) increase enzyme fitness by an amount equal to the *y* component of the arrow. The landscape can be traversed by sequential point mutagenesis and selection or screening over multiple generations, or by recombination of beneficial mutations identified in earlier generations.

(a) A Fujiyama-type landscape with constant slope. The constant slope indicates that the mutations are completely additive in their effects on fitness.

(b) A Fujiyama-type landscape with decreasing slope approaching the fitness optimum. Here, the benefit of any given mutation depends on the rest of the sequence. Mutations in advanced generations are more difficult to identify owing to diminishing returns obtained for each mutation. Recombination of point mutations from earlier generations also yields a smaller increment in fitness than strict additivity would predict.

The complete landscape is multidimensional and includes fitnesses for all possible sequences and pathways. The 'sequence coordinate' is a one-dimensional representation of all possible amino acid combinations.

sizes accessible to searching. Increasing control over mutagenic methods and insight into optimal strategies for traversing protein landscapes will further expand the potential of directed evolution to create novel biocatalysts.

Acknowledgments

Support from the Army Research Office, the Office of Naval Research and the Department of Energy's program in Biological and Chemical Technologies Research is acknowledged. O. Kuchner is the recipient of an NIH predoctoral training fellowship from the National Institute of General Medical Sciences, NRSA Award 1 T32 GM 08346-01. The authors thank Daniel Song for helpful discussions.

References

1 Suzuki, M., Christians, F. C., Kim, B., Skandalis, A., Black, M. and Loeb, L. A. (1996) *Mol. Diversity* 2, 111–118
2 Shafikhani, S., Siegel, R. A., Ferrari, E. and Schellenberger, V. (1997) *BioTechniques* 23, 304–310
3 Arnold, F. H. (1996) *Chem. Eng. Sci.* 51, 5091–5102
4 Liao, H., McKenzie, T. and Hageman, R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 576–580
5 Chen, K. and Arnold, F. H. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 5618–5622
6 You, L. and Arnold, F. H. (1996) *Protein Eng.* 9, 77–83
7 Stemmer, W. P. C. (1994) *Nature* 370, 389–391
8 Strausberg, S. L., Alexander, P. A., Gallagher, D. T., Gilliland, G. L., Barnett, B. L. and Bryan, P. N. (1995) *Biotechnology* 13, 669–673
9 Moore, J. C. and Arnold, F. H. (1996) *Nat. Biotechnol.* 14, 458–467
10 Black, M. E., Newcomb, T. G., Wilson, H. M. P. and Loeb, L. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 3525–3529
11 Zhang, J-H., Dawes, G. and Stemmer, W. P. C. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 4504–4509
12 Crameri, A., Dawes, G., Rodrigues, E., Jr, Silver, S. and Stemmer, W. P. C. (1997) *Nat. Biotechnol.* 15, 436–438
13 Moore, J. C., Jin, H. M., Kuchner, O. and Arnold, F. H. (1997) *J. Mol. Biol.* 272, 336–347
14 Haruki, M. *et al.* (1994) *J. Biol. Chem.* 269, 26 904–26 911
15 Pjura, P., Matsumura, M., Baase, W. A. and Matthews, B. W. (1993) *Protein Sci.* 2, 2217–2225
16 Kotsuka, T., Akanuma, S., Tomuro, M., Yamagishi, A. and Oshima, T. (1996) *J. Bacteriol.* 178, 723–727
17 Okada, Y. (1995) *Biosci. Biotechnol. Biochem.* 59, 1152–1153
18 Bryan, P. N. *et al.* (1986) *Proteins* 1, 326–334
19 Shinkai, A., Hirano, A. and Aisaka, K. (1996) *J. Biochem.* 120, 915–921
20 Rellos, P. and Scopes, R. K. (1994) *Protein Expres. Purif.* 5, 270–277
21 Wells, J. A. (1990) *Biochemistry* 29, 8509–8517
22 Joyet, P., Declerck, N. and Gaillardin, C. (1992) *Biotechnology* 10, 1579–1583
23 Matsumura, M. and Aiba, S. (1985) *J. Biol. Chem.* 260, 15 298–15 303
24 Tamakoshi, M., Yamagishi, A. and Oshima, T. (1995) *Mol. Microbiol.* 16, 1031–1036
25 Zhao, H. and Arnold, F. H. (1997) *Curr. Opin. Struct. Biol.* 7, 480–485
26 Kano, H., Taguchi, S. and Momose, H. (1997) *Appl. Microbiol. Biotechnol.* 47, 46–51
27 Beuve, A. and Danchin, A. (1992) *J. Mol. Biol.* 225, 933–938
28 Hawrani, A. S., Sessions, R. B., Moreton, K. M. and Holbrook, J. J. (1996) *J. Mol. Biol.* 264, 97–110
29 Sidhu, S. S. and Borgford, T. J. (1996) *J. Mol. Biol.* 257, 233–245
30 Ohnuma, S-I. *et al.* (1996) *J. Biol. Chem.* 271, 10 087–10 095
31 Sousa, R. and Padilla, R. (1995) *EMBO J.* 14, 4609–4621
32 Gulick, A. M. and Fahl, W. E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 8140–8144
33 Widersten, M. and Mannervik, B. (1995) *J. Mol. Biol.* 250, 115–122
34 MacBeath, G. and Hilvert, D. (1996) *Chem. Biol.* 3, 433–445
35 Matcham, G. W. and Bowen, A. R. S. (1996) *Chem. Today* 14, 20–24
36 Cadwell, R. C. and Joyce, G. F. (1994) *Mutagenic PCR*, Cold Spring Harbor Laboratory
37 Leung, D. W., Chen, E. and Goeddel, D. V. (1989) *Technique* 1, 11–15
38 Derbyshire, K. M., Salvo, J. J. and Grindley, N. D. F. (1986) *Gene* 46, 145–152
39 Oliphant, A. R., Nussbaum, A. L. and Struhl, K. (1986) *Gene* 44, 177–183
40 Reidhaar-Olson, J. F. and Sauer, R. T. (1988) *Science* 241, 53–57
41 Delagrave, S., Goldman, E. R. and Youvan, D. C. (1995) *Protein Eng.* 8, 237–242
42 Stemmer, W. P. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* 91, 10 747–10 751
43 Zhao, H. and Arnold, F. H. (1997) *Nucleic Acids Res.* 25, 1307–1308
44 Lorimer, I. A. and Pastan, I. (1995) *Nucleic Acids Res.* 23, 3067–3068
45 Stemmer, W. P. C. (1995) *Biotechnology* 13, 549–553
46 Hilvert, D. and Kast, P. (1997) *Curr. Opin. Struct. Biol.* 7, 470–479
47 Eigen, M. (1985) *Ber. Bunsenges. Phys. Chem.* 89, 658–667
48 Kauffman, S. A. (1993) *The Origins of Order*, Oxford University Press
49 Macken, C. A. and Perelson, A. S. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 6191–6195
50 Aita, T. and Husimi, Y. (1996) *J. Theor. Biol.* 182, 469–485
51 Trakulnaleamsai, S., Yomo, T., Yoshikawa, M., Aihara, S. and Urabe, I. (1995) *J. Ferment. Bioeng.* 79, 107–118