Genetic selection strategies for generating and characterizing catalysts

Peter Kast and Donald Hilvert

Departments of Chemistry and Molecular Biology, The Scripps Research Institute,
10666 North Torrey Pines Road, La Jolla, California 92037, USA

Abstract: Molecular evolution is a powerful tool for exploring and creating catalytic function in biological macromolecules. This review focuses on examples and current strategies for the application of random mutagenesis of target genes coupled with direct selection *in vivo*. The technology is illustrated by the dissection of catalytic features in chorismate mutase. Future avenues for identifying and evolving catalytic antibodies are also discussed.

INTRODUCTION

Reengineering of biological macromolecules to elucidate or alter their function is currently one of the most significant research enterprises in modern biological science (1,2). Typically, specific positions in a target molecule, which are assumed to possess functional importance, are modified by site-directed mutagenesis and the effects of the resulting change(s) are assessed. Although sometimes successful (3,4), "rational" design frequently fails to yield the expected effect because of our limited understanding of the relationship between sequence, tertiary structure and function (5,6). Molecular evolution is a powerful alternative approach that involves random mutagenesis of coding nucleic acids and subsequent selection of individual clones which are capable of performing a desired task (7,8). The latter strategy is well-suited to scanning for functional alternatives and creating novel activities in macromolecules such as RNA, DNA and proteins. While both "rational" and "random" design can be applied in a complementary manner (9), the random approach described here may prove particularly useful in cases where structural data are not readily available or where mutation is likely to perturb structure.

IN VITRO EVOLUTION OF NUCLEIC ACIDS

While this review will focus on the application of *in vivo* selection strategies to protein catalysts, it must be pointed out that molecular evolution techniques have been applied extensively to nucleic acids. This class of molecules has the peculiarity that genotype (coding information) and phenotype (encoded function) can reside on the same molecule. This intrinsic simplicity and the availability of suitable RNA and DNA polymerases has led to the development of genuine Darwinian-type molecular evolution *in vitro* ["*in vitro* genetics" (10)], consisting in its complete form of iterative cycles of mutagenesis, selection and amplification of the selected variants (11). By implementing this technology, it has been possible to directly select nucleic acid molecules that recognize ligands with high specificity and/or catalyze diverse reactions. For more in depth discussion of strategies and developments in the field of *in vitro* RNA or DNA (12-14) evolution, the reader is referred to recent reviews (15-21). The reactions accelerated by evolved catalytic RNA molecules (ribozymes) usually involve phosphate ester chemistry (17), as do naturally occurring ribozymes (22,23). That this is not a fundamental limitation for nucleic acid catalysis was demonstrated recently by the creation of ribozymes capable of cleaving or forming carboxylate ester bonds (24,25), cleaving amides (26), transferring alkyl groups (27) and isomerizing substituted biphenyls (28). Furthermore, it is expected (29,30) that the scope of ribozyme-catalyzed reactions can be further expanded by the inclusion of cofactors

(in addition to metal ions) such as ATP (29), cyanocobalamin (31), biotin (27), nicotinamide cofactors (29,32,33), flavin cofactors (32,33) and analogs of natural cofactors (30) for which specific interactions with selected RNA molecules have already been demonstrated. This assumption is supported by the utilization of ATP by a ribozyme with polynucleotide kinase activity which was evolved from an ATP-binding RNA (34).

Nevertheless, ribozymes may be inferior to protein catalysts for most catalytic tasks. The more versatile chemistry and interactions possible with the diverse array of functional groups in proteins, as opposed to the limited functionalities of the four RNA building blocks, is thought to have led to the replacement of the ribozymes present in a primordial RNA world by proteins (35).

MUTAGENESIS AND SELECTION *IN VIVO*: THE TECHNOLOGY

It is desirable to subject proteins to molecular evolution using strategies similar to the RNA work described above. The application of *in vitro* selection schemes to proteins, however, appears impractical because the phenotype (selectable activity) is physically separated from the genotype, and the process of protein synthesis is far more complex than the transcription of RNA. Nevertheless, an interesting approach for *in vitro* selection of small peptides from large libraries has been put forward recently (36); it iteratively employs *in vitro* transcription/translation and subsequent affinity selection of the large complexes (polysomes) between mRNA, translating ribosomes, various translation factors and the nascent target peptides, but the system has not yet been tested with larger proteins or enzymes. Alternatively, the *in vitro* selection step may be replaced by selection *in vivo* (Fig. 1), if a protein catalyst can provide a selective advantage for growth or survival to a specific organism. *In vivo* selection systems are differentiated from their *in vitro* counterparts by the limitations on the library sizes that can be sampled, which are currently estimated to be at 10^9 to 10^{10} or at 10^{14} to 10^{15}, respectively (36). Furthermore, *in vivo* selection may not be as direct as selection *in vitro*, and therefore more prone to artifacts. On the other hand, its application may be advantageous if the catalyst to be selected is intended for subsequent use *in vivo*, i. e. as a therapeutic drug. Ribozymes created by *in vitro* molecular evolution often fail to function *in vivo* due to specific constraints imposed by living organisms (37,38).

A general scheme for mutagenesis coupled with selection *in vivo* which is applicable to enzymes is depicted in Fig. 1. A gene designed to serve as the starting point is cloned into a plasmid vector. This gene is then subjected to randomizing mutagenesis. Depending on the specific case, this can be accomplished by a variety of procedures that include chemical mutagenesis (39,40), error-prone PCR (polymerase chain reaction) (41), DNA shuffling (8,42,43) and degenerate oligonucleotide mutagenesis (44-46). The library of resulting plasmid mutants is then introduced into a suitable selection organism, preferably a mutant strain of the well-characterized
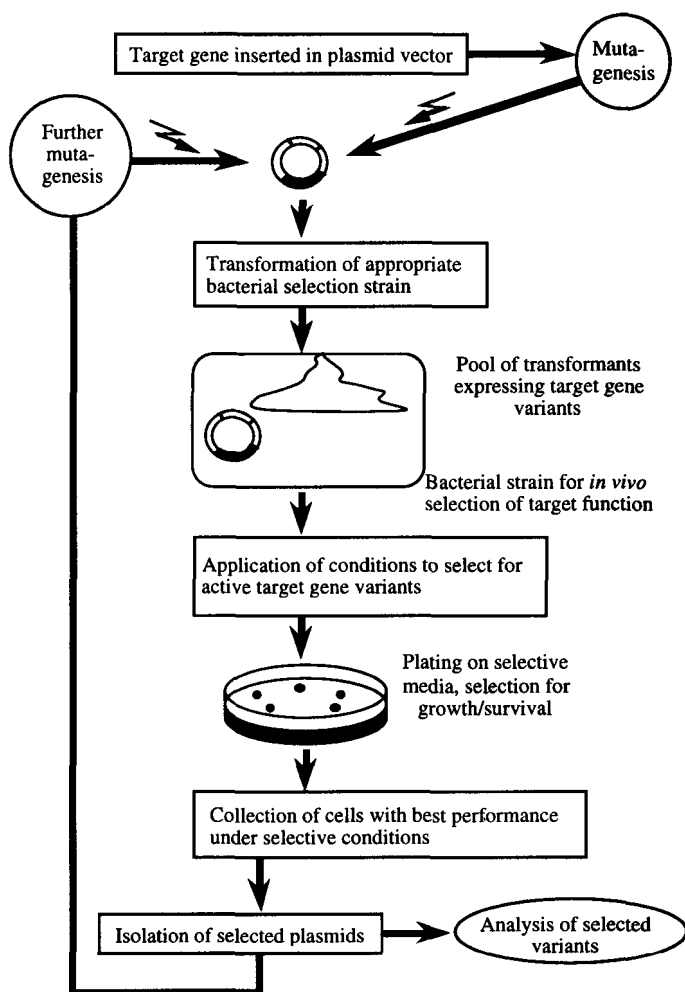

Fig. 1   Strategy for catalyst evolution and characterization by random mutagenesis and *in vivo* selection.

bacterium *Escherichia coli*. Selective conditions for the desired plasmid-encoded target function are imposed, and growing clones able to satisfy the selection criteria are collected. Finally, analysis of the selected variants yields information about active constellations of amino acid residues. Further rounds of mutagenesis and selection starting from partially active variants will imitate Darwinian evolution giving this strategy the potential to create catalysts *de novo*.

ENZYME CATALYSIS DISSECTED

Space limitations and the wealth of examples described in the literature make it impossible to cover all the experiments in which random mutagenesis and *in vivo* selection have been used to probe enzyme function. We will therefore highlight just a few of the diverse applications of this strategy.

One of the first enzymes to be studied by this methodology was the penicillin and cephalosporin-cleaving TEM-1 β-lactamase (47,48). By partially randomizing its gene and selecting for resistance against a variety of β-lactam and cephalosporin antibiotic derivatives, enzyme variants with dramatically altered catalytic activities and novel substrate specificities were readily obtained (42,47-50). The results emerging from these investigations have contributed to our understanding of the evolution of clinical antibiotic resistance and have important implications for the design of improved drugs.

Elegant work by Knowles and coworkers (51,52) has shown that the catalytic activity of a sluggish mutant triosephosphate isomerase can be partially restored by second-site amino acid changes. Their *in vivo* selection system was based on metabolic complementation of an *E. coli* strain from which the endogenous triosephosphate isomerase gene was excised. The rescue replacements, all close to the active site of the enzyme, demonstrate that multiple solutions for a similar functional outcome exist. Energetic characterization of two pseudorevertants revealed increased transition state stabilization in both cases, when compared to the starting mutant (52).

Genetic selection in a thymidine kinase-deficient *E. coli* strain was used to identify active clones from random gene libraries encoding variants of herpes simplex virus thymidine kinase (7,53-56). These studies examined determinants of the presumed nucleoside binding site, a target of antiviral chemotherapy. Among the selected variants were enzymes which phosphorylated thymidine at a higher rate than the wild type (54) and kinases with altered thermostability and, most interestingly, with increased activity towards nucleoside analogs such as the antiviral drugs AZT, gancyclovir and acyclovir (7,55,56). Selective phosphorylation of these chemotherapeutics by a viral kinase might be exploited for gene therapy in cancer treatment by making tumor cells into which the mutant viral gene is specifically introduced susceptible to nucleoside analogs. DNA replication is not affected in other cells whose endogenous thymidine kinase does not phosphorylate the drug.

In vivo selection systems for proteolytic (57,58) and nucleolytic enzymes (59) have made use of substrates that, upon hydrolysis by functional variants of the protein, liberate a nutrient needed to restore growth of a correspondingly auxotrophic *E. coli* strain. Studies with rat anionic trypsin, for example, established the importance of a negative charge at the base of the primary substrate binding pocket for cleavage of arginyl or lysyl derivatives (57,58). Gerlt and coworkers have used this approach in an attempt to improve the potency of a low-activity staphylococcal nuclease variant in which the general base catalyst Glu43 was replaced by Asp (59). However, selection of clones from a library in which the codons for the two adjacent amino acids were randomized led to the fortuitous isolation of a variant with an unexpected deletion. The active enzyme had Glu43 restored but was lacking six residues from a prominent peptide loop in the active site. This seemingly drastic perturbation was obviously less detrimental to catalytic activity than the removal of a single methylene in the Glu-to-Asp variant. The result also demonstrates the potential of direct selection to obtain unplanned, but highly interesting results.

A different type of *in vivo* selection can be applied to potentially toxic proteins to identify essential residues, as exemplified by experiments with HIV protease (60). Transformed *E. coli* cells were shown to survive only if the gene for the viral protein suffered an inactivating mutation. Knowledge of essential – and therefore potentially invariable – positions may help in the development of new target-bound drugs and vaccines.

Finally, the application of evolutionary engineering techniques may prove particularly valuable for improving the performance of catalysts under extreme conditions. By using selection for Leu prototrophy (61) or kanamycin resistance (62,63) in a thermophilic organism at high temperature, the (thermo-) stability of mesophilic enzymes has been significantly augmented.

The principles of such *in vivo* selection experiments are illustrated in more detail in the following section with an example from our own laboratory.

CATALYSIS BY CHORISMATE MUTASE

Chorismate mutase (CM) catalyzes the Claisen rearrangement of chorismate **1** to prephenate **3** (Scheme 1), a key reaction in the biosynthesis of the aromatic amino acids Phe and Tyr in bacteria, fungi and higher plants (64). The enzymatic reaction (65), like the spontaneous thermal rearrangement (66), proceeds through the chair-like transition state **2**, but the origins of the 10^6 to 10^7-fold rate acceleration by CM are poorly understood, despite extensive study. In the uncatalyzed reaction, ether bond cleavage substantially precedes C-C bond formation, resulting in a dipolar transition state (67-69). To examine whether electrostatic stabilization of a charge-separated transition state is important for catalysis by CM, we investigated the roles of a set of active site residues in *Bacillus subtilis* CM by random mutagenesis and selection *in vivo* (70,71).

Scheme 1
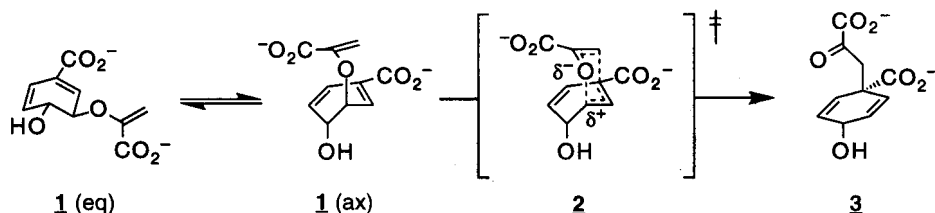
A selection system in *E. coli* for CM activity was devised by reengineering the metabolic pathways from chorismate to phenylpyruvate and 4-hydroxyphenylpyruvate. As shown in Fig. 2, the genes for the two bifunctional *E. coli* enzymes CM-prephenate dehydratase and CM-prephenate dehydrogenase were replaced by the genes for monofunctional versions of prephenate dehydratase (72) and prephenate dehydrogenase (73) from two other bacterial species. The resulting host strain made it possible to conduct *in vivo* selection experiments with the monofunctional *B. subtilis* CM, encoded by *aroH* (74).

As apparent from the crystal structure of a complex between *B. subtilis* CM and a transition state analog (75,76), the putative dipolar transition state could be stabilized by electrostatic interactions with Arg90 and Glu78 (Fig. 3). The codons for these residues were therefore randomized, the two gene libraries were introduced into the host strain, and clones were selected for their ability to provide their own Tyr and Phe.
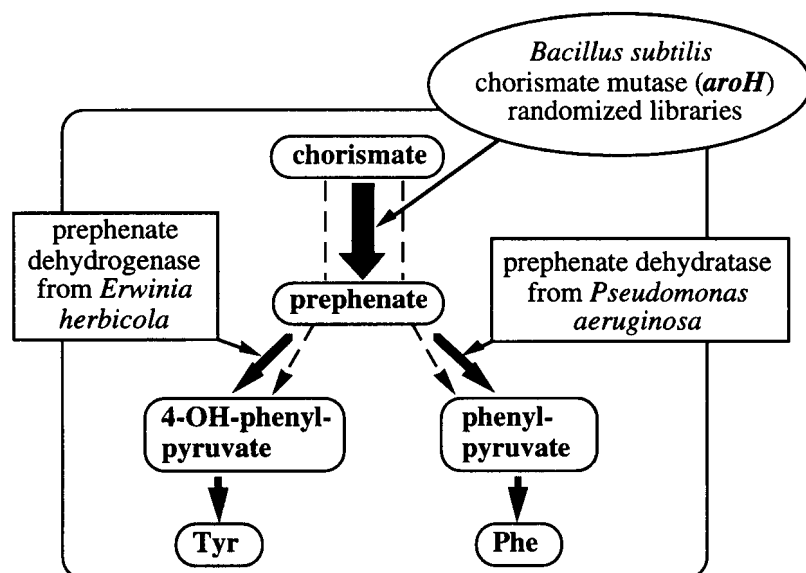

Fig. 2 Metabolic reengineering to create an *in vivo* selection system for chorismate mutase activity. Dashed lines indicate the transformations catalyzed by the two replaced bifunctional enzymes from *E. coli*.

While Arg90 was found to be essential, Glu78 can be replaced by other residues capable of hydrogen bonding interactions, albeit with substantial losses in CM activity of at least a thousand-fold. Two additional libraries which allow for more flexibility in the disposition of active site residues were constructed by simultaneously randomizing the codons for Cys88 and Arg90, or Cys75 and Glu78, respectively (Fig. 3). It emerged that a positively charged residue is required, either at position 88 or 90; and good complementation was observed for variants with an acid at either position 75 or 78. Thus, the pattern of selected active variants obtained from these randomization and selection experiments (70,71) supports the hypothesis that a dipolar active site is important for catalysis. In addition, some of the alternative active site

configurations that were selected have features which resemble those of a non-related CM from *E. coli* (77), suggesting that an electrostatic gradient across the active site may be a general requirement for an efficient CM. Interestingly, the catalytic antibody 1F7 with very modest CM activity possesses an Arg residue in the vicinity of the ether oxygen of the bound transition state analog, but lacks the negative pole in the active site (78).

The results from our evolutionary approach are supported by very recent experiments using the more traditional site-directed mutagenesis of active site residues in *B. subtilis* CM (79).

Fig. 3 Schematic of the active site from *B. subtilis* chorismate mutase with the presumed transition state of the reaction (thin lines).

CATALYTIC ANTIBODIES

Another possibility for the creation of catalysts *de novo* is the generation of catalytic antibodies. By exploiting the enormous capacity of the mammalian immune system to provide tight binding antibodies to virtually any compound, catalytic immunoglobulins can be elicited with a hapten that mimics the transition state of a desired chemical transformation (80-85). An alternative approach for obtaining antibody catalysts makes use of phage display technology (86-89) for the sampling of large combinatorial libraries composed of immuno-selected (90,91) or unselected (92) immunoglobulin genes. The products obtained from both strategies, the monoclonal antibodies secreted from hybridomas and the affinity-selected antibody fragments displayed on phage, are actually derived from evolutionary processes consisting of gene rearrangements coupled with selection for hapten binders. Somatic hypermutation or *in vitro* randomization of complementary determining regions (CDRs) (92) add additional diversity to the pool of available molecules. However, selection for binding to a transition state analog is still only an indirect method of finding catalysts. A binding site for a transition state analog of a reaction does not guarantee catalysis, and even the most sophisticated hapten designs (93) do not ensure precise positioning of catalytic groups or optimized active site dynamics. Thus, the screening of the immune response and affinity-selected phage libraries directly for catalysts rather than for transition state analog binders is an important, but often tedious step in the quest for catalytic antibodies. Although efficient general methods have been developed for the detection of catalysis by antibodies (94-96), they require the handling of large numbers of clones. In contrast, selection directly for active catalysts would be far more effective.

Scheme 2
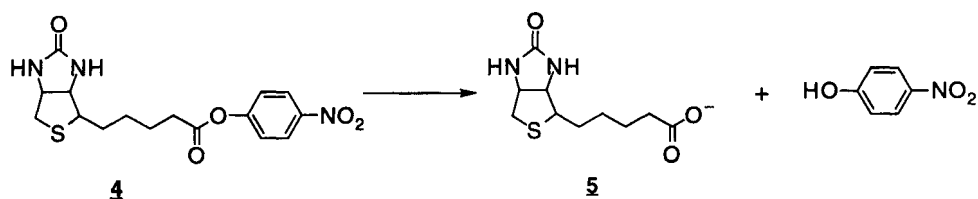
In vivo selection for catalysis represents a direct and elegant solution to the problem of identifying highly active catalytic antibodies. A general approach similar to the one used for proteases and nucleases (see above) has been proposed for hydrolytic antibodies (97). The system relies on *E. coli* mutant strains deficient in the biosynthesis of an essential cofactor, such as biotin (5) or thiamine. These bacteria can only grow if the missing nutrient is made available to the cells, for instance, by hydrolysis of a precursor compound to liberate the cofactor. Antibody 48G7 which was raised against a nitrophenyl phosphonate transition state analog and exhibits broad substrate specificity was shown to hydrolyze the biotin-*p*-nitrophenyl ester 4 (Scheme 2) and a *p*-nitrophenylthiazolyl carbonate *in vitro*. Although the genes for a Fab fragment of 48G7 could be expressed in the selection strain and hapten binding activity was demonstrated in the rele-
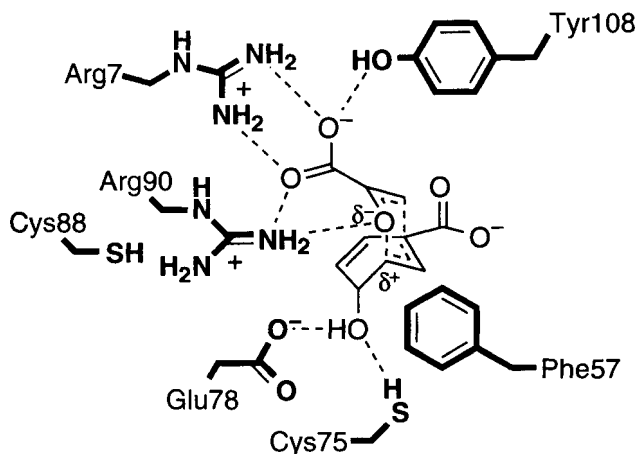
vant cellular fraction, no complementation of the biotin auxotrophy was observed. It is conceivable that selection experiments after randomization of the antibody genes will lead to the identification of catalysts active *in vivo*. No results were reported for the *in vivo* complementation of a thiamine-requiring strain (97).

Smiley and Benkovic described the *in vivo* selection of Fab-encoding antibody genes from a combinatorial immunoglobulin cDNA library prepared from mice which were immunized with an *N*-substituted 2,4-dihydroxyquinoline (98). The hapten was designed to elicit antibodies for the decarboxylation of orotate to yield uracil. Although 10^7-fold less active than the relevant natural enzyme, a single-chain version of one of the six selected Fab fragments was reported to complement the uracil-deficiency of an auxotrophic *E. coli* mutant strain (98).

The Fab fragment of 1F7, the CM catalyst mentioned above (99), was the first catalytic antibody shown to function *in vivo*. A permissive CM-deficient yeast strain could be complemented by introducing and inducing the genes for the light chain and the truncated heavy chain of the immunoglobulin (100). Expression of the 1F7 genes in yeast yielded correctly assembled, functional Fab protein in a ≈ 10-fold higher intracellular concentration than the natural CM (100,101). Since the Fab fragment possesses a 10^4-fold lower catalytic activity, 1F7 Fab-producing yeast cells which lack the natural enzyme may only exhibit 10^{-3} of the wild-type activity. This level of enzyme activity is insufficient to relieve the Tyr and Phe requirement of the original CM-deficient yeast strain YT-4Ca (101), but provides an obvious growth advantage on selection media to a permissive host strain which was derived from YT-4Ca by random chemical mutagenesis (100). Experiments are in progress to evolve the 1F7 genes by using *in vivo* selection in yeast and in the *E. coli* complementation system described above (Y. Tang, P. Kast and D. Hilvert, unpublished).

The broad application of *in vivo* selection to identify and evolve first-generation catalytic antibodies has to overcome three hurdles. In order to apply genetic selection pressure within a microorganism, the genes for the catalytic antibodies must first be cloned. Second, a suitable *in vivo* selection system needs to be available. Third, the antibody genes which are foreign to microbial selection hosts must be produced as soluble proteins at high levels, which is often problematic (97,98,101) (Y. Tang and D. Hilvert, unpublished). However, the advent of recombinant combinatorial antibody libraries (where the genes are already cloned), new general approaches to design and engineer *in vivo* selection systems [see above and ref. (97)] and improved technologies for high-level antibody gene expression in *E. coli* (102) are expected to greatly expand the utility and applicability of *in vivo* selection.

CONCLUSIONS

Mutagenesis coupled to genetic selection is a very powerful strategy for analyzing catalytic mechanisms and for improving or creating biological catalysts. Direct selection of the few interesting molecules from large random libraries of variants alleviates tedious and impractical screening procedures. Whereas RNA evolution experiments can employ selection steps *in vitro*, the corresponding approaches for catalytic proteins generally require *in vivo* strategies. It has been demonstrated that tailored and versatile *in vivo* systems can be engineered by altering whole metabolic pathways within the cell, by coupling the catalytic reaction of interest to a metabolically relevant compound, or by arrangements leading to the destruction of a compound toxic to the cell. Future applications of protein evolution in the laboratory may not only complement but also outperform current "rational" protein design strategies.

ACKNOWLEDGMENTS

REFERENCES

1.   G. Petsko and S. Sligar. *Curr. Opin. Struct. Biol.* **4**, 591-592 (1994).
2.   J.M. Thornton and J.M. Berg. *Curr. Opin. Biotechnol.* **6**, 367-369 (1995).
3.   D.A. Estell, T.P. Graycar, J.V. Miller, D.B. Powers, J.P. Burnier, P.G. Ng and J.A. Wells. *Science* **233**, 659-663 (1986).
4.   J.A. Wells, D.B. Powers, R.R. Bott, T.P. Graycar and D.A. Estell. *Proc. Natl. Acad. Sci. USA* **84**, 1219-1223 (1987).
5.   L. Hedstrom. *Curr. Opin. Struct. Biol.* **4**, 608-611 (1994).

6. S.F. Betz, D.P. Raleigh and W.F. DeGrado. *Curr. Opin. Struct. Biol.* **3**, 601-610 (1993).
7. D.K. Dube, M.E. Black, K.M. Munir and L.A. Loeb. *Gene* **137**, 41-47 (1993).
8. W.P.C. Stemmer. *Bio/Technology* **13**, 549-553 (1995).
9. R.R. Breaker and G.F. Joyce. *Trends Biotech.* **12**, 268-275 (1994).
10. J.W. Szostak. *Trends Biochem. Sci.* **17**, 89-93 (1992).
11. A.A. Beaudry and G.F. Joyce. *Science* **257**, 635-641 (1992).
12. B. Cuenoud and J.W. Szostak. *Nature* **375**, 611-614 (1995).
13. D.E. Huizenga and J.W. Szostak. *Biochemistry* **34**, 656-665 (1995).
14. R.R. Breaker and G.F. Joyce. *Chem. Biol.* **1**, 223-229 (1994).
15. S.J. Klug and M. Famulok. *Mol. Biol. Reports* **20**, 97-107 (1994).
16. K.B. Chapman and J.W. Szostak. *Curr. Opin. Struct. Biol.* **4**, 618-622 (1994).
17. G.F. Joyce. *Curr. Opin. Struct. Biol.* **4**, 331-336 (1994).
18. B.L. Iverson. *Chem. Biol.* **2**, 67-70 (1995).
19. M.M. Ouellette and W.E. Wright. *Curr. Opin. Biotechnol.* **6**, 65-72 (1995).
20. I. Hirao and A.D. Ellington. *Curr. Biol.* **5**, 1017-1022 (1995).
21. P. Burgstaller and M. Famulok. *Angew. Chem. Int. Ed. Engl.* **34**, 1189-1192 (1995).
22. T. Tuschl, J.B. Thomson and F. Eckstein. *Curr. Opin. Struct. Biol.* **5**, 296-302 (1995).
23. R.H. Symons. *Curr. Opin. Struct. Biol.* **4**, 322-330 (1994).
24. J.A. Piccirilli, T.S. McConnell, A.J. Zaug, H.F. Noller and T.R. Cech. *Science* **256**, 1420-1424 (1992).
25. M. Illangasekare, G. Sanchez, T. Nickles and M. Yarus. *Science* **267**, 643-647 (1995).
26. X. Dai, A. De Mesmaeker and G.F. Joyce. *Science* **267**, 237-240 (1995).
27. C. Wilson and J.W. Szostak. *Nature* **374**, 777-782 (1995).
28. J.R. Prudent, T. Uno and P.G. Schultz. *Science* **264**, 1924-1927 (1994).
29. M. Sassanfar and J.W. Szostak. *Nature* **364**, 550-553 (1993).
30. R.R. Breaker and G.F. Joyce. *J. Mol. Evol.* **40**, 551-558 (1995).
31. J.R. Lorsch and J.W. Szostak. *Biochemistry* **33**, 973-982 (1994).
32. C.T. Lauhon and J.W. Szostak. *J. Am. Chem. Soc.* **117**, 1246-1257 (1995).
33. P. Burgstaller and M. Famulok. *Angew. Chem. Int. Ed. Engl.* **33**, 1084-1087 (1994).
34. J.R. Lorsch and J.W. Szostak. *Nature* **371**, 31-36 (1994).
35. L.E. Orgel. *J. Mol. Biol.* **38**, 381-393 (1968).
36. L.C. Mattheakis, R.R. Bhatt and W.J. Dower. *Proc. Natl. Acad. Sci. USA* **91**, 9022-9026 (1994).
37. R.B. Denman, M. Smedman, W. Ju, R. Rubenstein, A. Potempska and D.L. Miller. *Nucl. Acids Res.* **22**, 2375-2382 (1994).
38. J.J. Rossi. *Curr. Biol.* **4**, 469-471 (1994).
39. C. Lopez-Camacho and J. Polaina. *Mutation Res.* **301**, 73-77 (1993).
40. R.M. Myers, L.S. Lerman and T. Maniatis. *Science* **229**, 242-247 (1985).
41. R.C. Cadwell and G.F. Joyce. *PCR Methods Appl.* **3**, S136-S140 (1994).
42. W.P.C. Stemmer. *Nature* **370**, 389-391 (1994).
43. W.P.C. Stemmer. *Proc. Natl. Acad. Sci. USA* **91**, 10747-10751 (1994).
44. J.F. Reidhaar-Olson and R.T. Sauer. *Science* **241**, 53-57 (1988).
45. S. Delagrave, E.R. Goldman and D.C. Youvan. *Protein Eng.* **6**, 327-331 (1993).
46. B. Virnekäs, L. Ge, A. Plückthun, K.C. Schneider, G. Wellnhofer and S.E. Moroney. *Nucl. Acids Res.* **22**, 5600-5607 (1994).
47. A.R. Oliphant and K. Struhl. *Proc. Natl. Acad. Sci. USA* **86**, 9094-9098 (1989).
48. D.K. Dube and L.A. Loeb. *Biochemistry* **28**, 5703-5707 (1989).
49. T. Palzkill, Q.-Q. Le, K.V. Venkatachalam, M. LaRocco and H. Ocera. *Mol. Microbiol.* **12**, 217-229 (1994).
50. H. Viadiu, J. Osuna, A.L. Fink and X. Soberón. *J. Biol. Chem.* **270**, 781-787 (1995).
51. J.D. Hermes, S.C. Blacklow and J.R. Knowles. *Proc. Natl. Acad. Sci. USA* **87**, 696-700 (1990).
52. S.C. Blacklow and J.R. Knowles. *Biochemistry* **29**, 4099-4108 (1990).
53. K.M. Munir, D.C. French, D.K. Dube and L.A. Loeb. *J. Biol. Chem.* **267**, 6584-6589 (1992).
54. K.M. Munir, D.C. French, D.K. Dube and L.A. Loeb. *Protein Eng.* **7**, 83-89 (1994).
55. M.E. Black and L.A. Loeb. *Biochemistry* **32**, 11618-11626 (1993).
56. K.M. Munir, D.C. French and L.A. Loeb. *Proc. Natl. Acad. Sci. USA* **90**, 4012-4016 (1993).

57. J.J. Perona, L.B. Evnin and C.S. Craik. *Gene* **137**, 121-126 (1993).
58. L.B. Evnin, J.R. Vásquez and C.S. Craik. *Proc. Natl. Acad. Sci. USA* **87**, 6659-6663 (1990).
59. L.B. Poole, D.A. Loveys, S.P. Hale and J.A. Gerlt. *Biochemistry* **30**, 3621-3627 (1991).
60. E.Z. Baum, G.A. Bebernitz and Y. Gluzman. *Proc. Natl. Acad. Sci. USA* **87**, 5573-5577 (1990).
61. T. Oshima. *Curr. Opin. Struct. Biol.* **4**, 623-628 (1994).
62. H. Liao, T. McKenzie and R. Hageman. *Proc. Natl. Acad. Sci. USA* **83**, 576-580 (1986).
63. M. Matsumura and S. Aiba. *J. Biol. Chem.* **260**, 15298-15303 (1985).
64. E. Haslam, *Shikimic Acid: Metabolism and Metabolites,* John Wiley & Sons, New York (1993).
65. S.G. Sogo, T.S. Widlanski, J.H. Hoare, C.E. Grimshaw, G.A. Berchtold and J.R. Knowles. *J. Am. Chem. Soc.* **106**, 2701-2703 (1984).
66. S.D. Copley and J.R. Knowles. *J. Am. Chem. Soc.* **107**, 5306-5308 (1985).
67. S.D. Copley and J.R. Knowles. *J. Am. Chem. Soc.* **109**, 5008-5013 (1987).
68. J.J. Gajewski, J. Jurayj, D.R. Kimbrough, M.E. Gande, B. Ganem and B.K. Carpenter. *J. Am. Chem. Soc.* **109**, 1170-1186 (1987).
69. L. Addadi, E.K. Jaffe and J.R. Knowles. *Biochemistry* **22**, 4494-4501 (1983).
70. P. Kast, M. Asif-Ullah, N. Jiang and D. Hilvert. *Proc. Natl. Acad. Sci. USA* **in press** (1996).
71. P. Kast, J.D. Hartgerink, M. Asif-Ullah and D. Hilvert. *J. Am. Chem. Soc.* **118**, 3069-3070 (1996).
72. G. Zhao, T. Xia, R.S. Fischer and R.A. Jensen. *J. Biol. Chem.* **267**, 2487-2493 (1992).
73. T. Xia, G. Zhao, R.S. Fischer and R.A. Jensen. *J. Gen. Microbiol.* **138**, 1309-1316 (1992).
74. J.V. Gray, B. Golinelli-Pimpaneau and J.R. Knowles. *Biochemistry* **29**, 376-383 (1990).
75. Y.M. Chook, H. Ke and W.N. Lipscomb. *Proc. Natl. Acad. Sci. USA* **90**, 8600-8603 (1993).
76. Y.M. Chook, J.V. Gray, H. Ke and W.N. Lipscomb. *J. Mol. Biol.* **240**, 476-500 (1994).
77. A.Y. Lee, P.A. Karplus, B. Ganem and J. Clardy. *J. Am. Chem. Soc.* **117**, 3627-3628 (1995).
78. M.R. Haynes, E.A. Stura, D. Hilvert and I.A. Wilson. *Science* **263**, 646-652 (1994).
79. S.T. Cload, D.R. Liu, R.M. Pastor and P.G. Schultz. *J. Am. Chem. Soc.* **118**, 1787-1788 (1996).
80. R.A. Lerner, S.J. Benkovic and P.G. Schultz. *Science* **252**, 659-667 (1991).
81. D. Hilvert. *Pure & Appl. Chem.* **64**, 1103-1108 (1992).
82. P.G. Schultz and R.A. Lerner. *Acc. Chem. Res.* **26**, 391-395 (1993).
83. J.D. Stewart, L.J. Liotta and S.J. Benkovic. *Acc. Chem. Res.* **26**, 396-404 (1993).
84. D. Hilvert. *Acc. Chem. Res.* **26**, 552-558 (1993).
85. D. Hilvert. *Curr. Opin. Struct. Biol.* **4**, 612-617 (1994).
86. L. Makowski. *Curr. Opin. Struct. Biol.* **4**, 225-230 (1994).
87. W.D. Huse, L. Sastry, S.A. Iverson, A.S. Kang, M. Alting-Mees, D.R. Burton, S.J. Benkovic and R.A. Lerner. *Science* **246**, 1275-1281 (1989).
88. J. McCafferty, A.D. Griffiths, G. Winter and D.J. Chiswell. *Nature* **348**, 552-554 (1990).
89. A.S. Kang, C.F. Barbas, K.D. Janda and S.J. Benkovic. *Proc. Natl. Acad. Sci. USA* **88**, 4363-4366 (1991).
90. C.F. Barbas III, A.S. Kang, R.A. Lerner and S.J. Benkovic. *Proc. Natl. Acad. Sci. USA* **88**, 7978-7982 (1991).
91. T. Clackson, H.R. Hoogenboom, A.D. Griffiths and G. Winter. *Nature* **352**, 624-628 (1991).
92. C.F. Barbas III, J.D. Bain, D.M. Hoekstra and R.A. Lerner. *Proc. Natl. Acad. Sci. USA* **89**, 4457-4461 (1992).
93. J. Wagner, R.A. Lerner and C.F. Barbas III. *Science* **270**, 1797-1800 (1995).
94. D.S. Tawfik, B.S. Green, R. Chap, M. Sela and Z. Eshhar. *Proc. Natl. Acad. Sci. USA* **90**, 373-377 (1993).
95. G. MacBeath and D. Hilvert. *J. Am. Chem. Soc.* **116**, 6101-6106 (1994).
96. H. Fenniri, K.D. Janda and R.A. Lerner. *Proc. Natl. Acad. Sci. USA* **92**, 2278-2282 (1995).
97. S.A. Lesley, P.A. Patten and P.G. Schultz. *Proc. Natl. Acad. Sci. USA* **90**, 1160-1165 (1993).
98. J.A. Smiley and S.J. Benkovic. *Proc. Natl. Acad. Sci. USA* **91**, 8319-8323 (1994).
99. D. Hilvert, S.H. Carpenter, K.D. Nared and M.-T.M. Auditor. *Proc. Natl. Acad. Sci. USA* **85**, 4953-4955 (1988).
100. Y. Tang, J.B. Hicks and D. Hilvert. *Proc. Natl. Acad. Sci. USA* **88**, 8784-8786 (1991).
101. K. Bowdish, Y. Tang, J.B. Hicks and D. Hilvert. *J. Biol. Chem.* **266**, 11901-11908 (1991).
102. H.D. Ulrich, P.A. Patten, P.L. Yang, F.E. Romesberg and P.G. Schultz. *Proc. Natl. Acad. Sci. USA* **92**, 11907-11911 (1995).