**Protein Engineering, Design & Selection** vol. 19 no. 10 pp. 439–442, 2006
Published online July 25, 2006 doi:10.1093/protein/gzl029

**REVIEW**

# Natural history as a predictor of protein evolvability

**Taryn L.O'Loughlin, Wayne M.Patrick and Ichiro Matsumura[1]**

Department of Biochemistry, Center for Fundamental and Applied Molecular Evolution, Emory University, 1510 Clifton Road, Atlanta, GA 30322, USA

[1]To whom correspondence should be addressed.
E-mail:imatsum@emory.edu

**Natural selection generally produces specific and efficient enzymes. In contrast, directed evolution experiments usually produce enzyme variants with broadened substrate specificity or enhanced catalytic promiscuity. Some proteins may be more evolvable than others, but few workers consider this problem when choosing starting points for laboratory evolution. Here, we review the variables associated with enzyme evolvability, namely promiscuity and mutational robustness. We present a qualitative model of adaptive evolution and recommend that protein engineers exploit their knowledge of natural history to identify evolvable wild-type proteins. Three examples of 'generalist' proteins that evolved in the laboratory into 'specialists' are described to illustrate the practical utility of this point.**
*Keywords*: adaptive evolution/catalytic promiscuity/directed evolution/evolvability/robustness

## Introduction

Adaptive molecular evolution is a fundamental biological process, yet it remains poorly understood. The slow pace of natural selection generally precludes direct observation. Furthermore, retrospective comparisons of natural homologues generally do not elucidate structural mechanisms of adaptation because (i) most fixed mutations are apparently neutral (Kimura, 1983); (ii) the fitness of any sequence is a function of multiple parameters (such as thermostability, $k_{cat}$, $K_M$ and substrate specificity) and is difficult to define; and (iii) the relationships between protein structure and function are often difficult to understand. Ancestral gene resurrection and biochemical characterization have revealed instances of adaptive molecular evolution (Thornton, 2004; Thomson *et al*., 2005; Tocchini-Valentini *et al*., 2005; Bridgham *et al*., 2006), although ancient adaptive events are usually less amenable to this approach (an exception is described by Gaucher *et al*., 2003). The complexity of protein evolution, however, does not diminish its fundamental importance. Molecular adaptation is one mechanism of speciation, so a better understanding of this process would further articulate the Darwinian paradigm. It would also increase the efficiency of directed evolution protocols for protein engineering.

In this review, we consider the enigma of protein evolvability, which is defined as the 'capacity of a lineage to evolve' (Kirschner and Gerhart, 1998). Previous studies that suggest correlations between evolvability and mutational robustness, modularity, promiscuity or substrate range are briefly reviewed. Natural selection generates particularly evolvable enzymes in response to rapidly fluctuating selection conditions; evolvability itself could thus be a selectable trait, despite its apparently anticipatory nature (Earl and Deem, 2004). We recommend that engineers consider a protein's 'natural history' before choosing it as a target for mutation and selection. Finally, we use this rationale to highlight proteins that most readily adapt to novel substrates, both in nature and in the laboratory.

## Specificity

Proteins that require the fewest mutations to adapt to novel reaction conditions, are the most likely to survive environmental changes (Voigt *et al*., 2004). Thus, the evolvability of an enzyme is partly a function of its capacity to catalyze secondary reactions that did not initially contribute to organismal fitness. Most enzymes possess some 'substrate range', which is to say that they react with structurally similar substrates and cause them to undergo mechanistically identical chemical transformations (Jensen, 1976). Some enzymes exhibit 'catalytic promiscuity', which means that they catalyze mechanistically different types of reactions (O'Brien and Herschlag, 1999). Several studies have demonstrated that random mutagenesis and selection or screening can broaden substrate specificity (Zhang *et al*., 1997; Yano *et al*., 1998; Olsen *et al*., 2000; Matsumura and Ellington, 2001) or enhance the promiscuous activities of enzymes (Schmidt *et al*., 2003; Aharoni *et al*., 2005a, b). In either case, selection should favor any amino acid replacement that improves recognition between the enzyme and the novel transition state (at least in cases in which transition state stabilization is rate-limiting).

## Mutational robustness and thermostability

Some have argued that evolvability is a function of 'robustness', i.e. the capacity of a protein to withstand variations in amino acid sequence and/or reaction conditions without disruption of its binding or catalytic properties (Wagner, 2005). Mutational robustness is determined by the number, order and distribution of a protein's intramolecular interactions, as well as the modularity of its domains (Li *et al*., 1996; England and Shakhnovich, 2003; Voigt *et al*., 2004). Loeb and co-workers have defined an '*x* factor' as the probability that a particular protein will be inactivated (within the context of an artificial high throughput screen or selection) by a single, random amino acid substitution. In their model system (3-methyladenine DNA glycosylase), the *x* factor was found to be 34 ± 6% (Guo *et al*., 2004). Earlier studies of T4 lysozyme (Rennell *et al*., 1991) and barnase (Axe *et al*., 1998) suggested *x* factors

of ~16 and ~5%, respectively. In the absence of a standardized system for generating diversity and for assaying 'retention of function', direct comparisons of these values are somewhat fraught. However, the development of such a system would prove insightful for experimentally quantifying the contribution of mutational robustness to evolvability.

Proper folding is a prerequisite for molecular recognition, so thermostability should correlate with mutational robustness and, therefore, with evolvability (Vendruscolo *et al*., 1997; Bornberg-Bauer and Chan, 1999; Tiana *et al*., 2001). However, some have argued that natural selection drives proteins towards marginal stability and into a less evolvable state. Most mutations are destabilizing, but are selectively neutral as long as the overall thermostability remains above the unfolding threshold (Taverna and Goldstein, 2002). The implication is that thermostable protein variants should be better starting points for directed evolution. This hypothesis has recently been demonstrated experimentally, by comparing the evolvability of marginally stable and thermostable variants of cytochrome P450 BM3. Random mutants derived from the latter were more likely to fold correctly than those derived from the marginally stable protein and were, consequently, more readily evolved to react with the novel substrate naproxen (Bloom *et al*., 2006). Taken together, these studies suggest that mutational robustness and thermostability are determinants of protein evolvability.

## Utility

Unfortunately, theoretical conjectures about the relationships between enzyme evolvability and substrate range, catalytic promiscuity and mutational robustness have had little influence upon protein engineers. These parameters are challenging to define in quantitative terms. Quantitative measurements such as the *x* factor (Guo *et al*., 2004), reactivity with novel substrates ($k_{cat}/K_M$) and thermostability ($\Delta G_{unfolding}$) are difficult to measure, and remain unproven predictors of evolvability. Protein engineers could begin by directing the evolution of extra-thermostable (Bloom *et al*., 2006) or extra-promiscuous (Schmidt *et al*., 2003; Chen and Zhao, 2005) variants, before using these variants as starting points to evolve the desired substrate specificity (Figure 1). However, few workers have made this extra effort, because these models have not yet been rigorously tested. In general, it remains impossible to predict the outcome of any directed evolution experiment *a priori*. Laboratory evolution remains more of an art than a science and will not mature from its current trial-and-error practices until the emergence of an accessible theoretical framework. Here we present a pre-paradigmatic hypothesis to focus attention upon an important and practical aspect of the problem.

## Hypothesis: nature selected certain enzymes for evolvability

Intuition suggests that natural selection optimizes most enzymes for specificity and catalytic efficiency ($k_{cat}/K_M$). This specialization is reflected in the ornate structures surrounding most enzyme active sites. For example, the tetrameric *Escherichia coli* β-galactosidase requires 4092 amino acids to catalyze the hydrolysis of a glycosidic bond (Jacobson *et al*., 1994). The active site structures are relatively rigid, so that the residues that interact with the preferred transition state are held

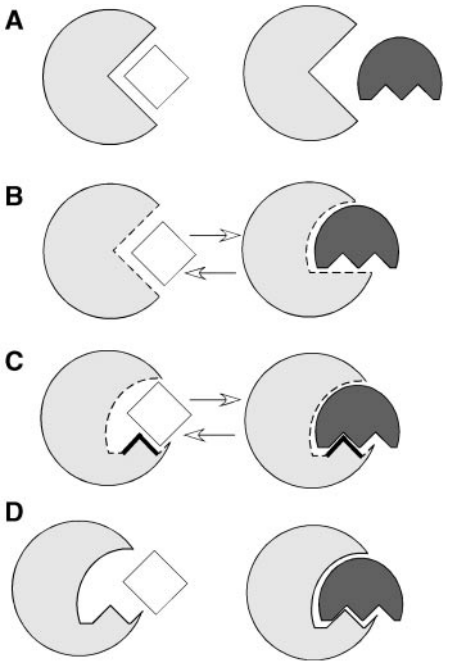


**Fig. 1.** Proposed model of adaptive enzyme evolution: (**A**) Wild-type enzyme is specific for the native diamond-shaped substrate. (**B**) Mutations that destabilize the conformation of the active site broaden substrate specificity so that the enzyme reacts with both substrates; most directed evolution experiments do not go beyond this point. (**C**) The evolving enzyme can accumulate mutations that introduce novel enzyme–substrate interactions, thereby, narrowing substrate specificity for the new ghost-shaped substrate. (**D**) Mutations throughout the protein can stabilize the new productive conformation, improving both specificity and catalytic efficiency.

in their proper conformations. Novel substrates are excluded, because they do not possess the right properties (size, shape, charge, etc.). Directed evolution tends to produce broad-specificity variants (Matsumura and Ellington, 2001; Aharoni *et al*., 2005b), perhaps by initially favoring mutations that increase the conformational flexibility of active site residues (Figure 1). The partial destabilization of specialized active site structures enables novel modes of transition state binding, often with only modest diminution of native activities (see also James and Tawfik, 2003).

In principle, these now evolvable enzymes could accumulate active site mutations that introduce novel enzyme–substrate interactions. Mutations throughout the protein that stabilize the new, productive active site conformation should also be favored. In practice, however, random mutagenesis of whole genes does not usually generate variants with narrowed substrate specificity (although exceptions include: Matsumura and Ellington, 2001; Varadarajan *et al*., 2005; O'Loughlin *et al*., 2006). Mutations that destabilize active site structures apparently occur more frequently than those that create new structures. Therefore, we postulate that most wild-type enzymes are sub-optimal starting points for directed evolution.

Nature may have selected particular enzymes for evolvability, rather than for catalytic efficiency or substrate specificity. We argue that it is presently easier to identify such enzymes by examining their natural history, rather than through difficult to assay properties such as catalytic promiscuity or mutational robustness. Evolvability is advantageous to organisms that must survive in rapidly changing environments (Elena and Sanjuan, 2003). Enzymes

that evolve under rapidly changing selection criteria are likely to become 'generalists' with regard to specificity. Such enzymes may be prompted to evolve into 'specialists' with relative ease. The identification of starting templates is a critical—but overlooked—aspect of directed evolution. Here we describe three examples of naturally evolvable proteins, which were identified by virtue of their functions in the context of natural, changing environments. This list is not comprehensive, and we expect that other proteins will also fit the profile.

## Example 1: antibodies

Antibody affinity maturation is the most rapid form of adaptive protein evolution. The bone marrow produces $\sim 10^8$ new lymphocytes every day, each displaying a unique receptor (antibody precursor) produced from random recombination of VDJ gene segments. This finite set of receptors must recognize a virtually infinite number of possible pathogen epitopes. Each of these germline precursor antibodies is thought to bind multiple antigens with modest affinities. When a B cell encounters a cognate antigen and receives a signal from a helper T cell, it proliferates and hypermutates the variable regions of the antibody. Daughter cells that display antibody variants with increased affinity for the antigen proliferate and secrete antibodies more quickly than the parental cell. The low affinity germline antibodies, evolve into high affinity, mature antibodies within days.

Nature appears to have selected antibodies that are evolvable, rather than those that are specific towards any particular subset of epitopes. The immunoglobulin fold is apparently robust to mutation, as it is modular in design and thermostable (the unfolding temperature of the constant region is $\sim 70^{\circ}$C; Vermeer and Norde, 2000). The crystal structures of five germline antibodies and their corresponding mature forms are generally consistent with the generalist to specialist transition described in our hypothesis. Four of the five germline antibodies exhibited significant conformational changes upon antigen binding (induced fit). In contrast, the corresponding mature antibodies bound their antigens in a 'lock and key' mechanism; the somatic mutations that were fixed during affinity maturation generally stabilized the binding sites in their most productive conformations (Schultz *et al*., 2002). Exceptional cases, including a germline antibody with a polyspecific but apparently inflexible binding site (Romesberg *et al*., 1998) and a mature but promiscuous antibody (James *et al*., 2003), suggest mechanistic diversity but are consistent with the postulated capacity for evolving narrowed specificity.

## Example 2: HIV protease

The Human Immunodeficiency Virus-1 protease is essential for viral replication. The HIV proteome is initially produced as a long Gag-Pol polyprotein; the individual viral proteins are inactive until HIV protease catalyzes their hydrolysis from this polyprotein (Kohl *et al*., 1988). The enzyme functions within the cytoplasm of the host cell and is, therefore, continually challenged with potential inhibitors and alternative substrates (Shoeman *et al*., 1991). The protease is somewhat specific as it must catalyze the hydrolysis of amide bonds between the individual viral proteins without recognizing sites within these proteins, but it is broad in the sense that the sites it does recognize do not conform to any consensus sequence. The cleavage of host proteins late in the viral life cycle (just prior to apoptosis of the host cell), when the protease exerts its enzymatic activity (Strack *et al*., 1996), is not deleterious to the virus. Absolute specificity is not necessarily advantageous and is not maintained. The heterologous expression of HIV protease is toxic to bacterial, yeast and human cells (Blanco *et al*., 2003), presumably owing to its broad substrate specificity and the resulting subsequent cleavage of essential cellular proteins.

The genomes of the Simian Immunodeficiency Viruses (SIV) are amongst the most rapidly evolving in the world (Wain-Hobson, 1993). SIV variants have infected 20 primate species and have adapted to human hosts at least twice in the past 70 years (Rambaut *et al*., 2004). Each time the virus invades a new host, the protease must adapt to a new proteome, replete with potential inhibitors and alternative substrates. HIV and other rapidly diversifying viruses express proteases that retain the ability to adapt quickly to a variety of new hosts. Unfortunately for human hosts, the administration of synthetic protease inhibitors usually leads to the rapid evolution of inhibitor-resistant forms; resistance is associated with a variety of amino acid replacements throughout the entire protein (Miller, 2001).

The natural history of HIV protease shows that it is particularly evolvable, with regard to both inhibitor resistance and substrate specificity. As predicted by theory (*vide ante*), it is broad in specificity and robust to mutations. Its overall structural simplicity (two monomers, each containing only 99 amino acids) suggests an absence of specialized structures for substrate recognition. Its substrate-binding cleft is long—the total area of contact between the enzyme and substrate is >1000 Å$^2$. HIV protease can change the shape of the cleft by forming slightly different homodimers (Prabu-Jeyabalan *et al*., 2000), thereby enabling reactions with a very broad range of substrates (Beck *et al*., 2000). The structural simplicity and accommodating nature of the active site confer evolvability but impose a significant cost in terms of catalytic efficiency (the $K_M$ of the enzyme is only $\sim$3 mM and the $k_{cat}$ is between 0.25 and 43 $s^{-1}$; Maschera *et al*., 1996). Therefore, we would predict that HIV protease possesses considerable potential for specialization; indeed, our initial experimental results are consistent with this postulation (O'Loughlin *et al*., 2006).

## Example 3: GroEL/GroES

The GroEL/GroES complex catalyzes the folding of unstable proteins in prokaryotes. The partially folded polypeptide binds the substrate recognition site (apical domain) of GroEL. The apical domain is very flexible and can bind many different polypeptide substrates ($\sim$250 proteins *in vivo*; Kerner *et al*., 2005) in different conformations. The GroEL/substrate complex binds GroES and ATP, inducing a conformational change that sequesters the polypeptide substrate within the cage-like GroEL/GroES structure, preventing aggregation with other partially folded intermediates (Gomez-Puertas *et al*., 2004). Both the N- and C-termini of GroEL are intrinsically unstructured and project into the central cavity; the interaction between these segments and the substrate promote folding through solubilization, local unfolding of misfolded domains and proximal positioning (Tompa and Csermely, 2004).

Wang *et al*. (2002) directed the evolution of GroEL/GroES variants that enhance the fluorescence of *Aequorea victoria*

Green Fluorescent Protein (GFP) co-expressed in *E.coli*. The fluorescence of GFP is absolutely dependent upon proper folding, and the wild-type GroEL/GroES complex catalyzes the folding of GFP *in vivo*. The evolved 'substrate-optimized' GroE variant was neither over-expressed nor more active than the wild-type; it proved more specific for the GFP, improving the folding of that substrate ~8-fold *in vivo*. The evolved chaperonin was less efficient than the wild-type at folding a variety of other substrates. The selected amino acid replacements did not map to the flexible or intrinsically unstructured domains of GroEL/GroES. It appears that substrate recognition was not rate-limiting in the reaction of the wild-type GroEL/GroES and GFP, and that changes in the chemical environment of the folding cavity somehow accelerated the reaction cycle. Owing to the design of the high throughput screen, GroEL/GroES variants unable to fold essential *E.coli* proteins could not have evolved in this experiment (Wang *et al*., 2002). Elimination of this constraint might enable further specialization, including adaptation of the substrate-binding domains.

## Summary

The design of directed evolution experiments is presently an art, rather than a science. The absence of quantitative models of adaptive enzyme evolution hinders the prediction of experimental outcomes. Here we reviewed some of the prevailing hypotheses, which suggest that substrate range, catalytic promiscuity and mutational robustness are determinants of protein evolvability. These parameters are important to consider, but difficult to measure, so we recommend that protein engineers consider natural history when choosing their starting templates for directed evolution. Nature may have selected some proteins for evolvability, rather than for specificity or catalytic efficiency. The identification and careful study of these exceptional proteins should enable the formulation of a proper theoretical framework and improved efficiency at the lab bench.

## Acknowledgments

We thank Dr Dan Tawfik and Mr Jesse Bloom for their insightful comments prior to the editing of this manuscript. T.L.O. and I.M. were supported by NIH/NIGMS 1 R01 GM074264-01. W.M.P. was supported by NSF/CHE-0404677.

## References

Aharoni,A., Amitai,G., Bernath,K., Magdassi,S. and Tawfik,D.S. (2005a) *Chem. Biol*., **12**, 1281–1289.
Aharoni,A., Gaidukov,L., Khersonsky,O., Gould,S. McQ., Roodveldt,C. and Tawfik,D.S. (2005b) *Nat. Genet*., **37**, 73–76.
Axe,,D.D., Foster,N.W. and Fersht,A.R. (1998) *Biochemistry*, **37**, 7157–7166.
Beck,Z.Q., Hervio,L., Dawson,P.E., Elder,J.H. and Madison,E.L. (2000) *Virology*, **274**, 391–401.
Blanco,R., Carrasco,L. and Ventoso,I. (2003) *J. Biol. Chem*., **278**, 1086–1093.
Bloom,J.D., Labthavikul,S.T., Otey,C.R. and Arnold,F.H. (2006) *Proc. Natl Acad. Sci. USA*, **103**, 5869–5874.
Bornberg-Bauer,E., and Chan,H.S. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 10689–10694.
Bridgham,J.T., Carroll,S.M. and Thornton,J.W. (2006) *Science*, **312**, 97–101.
Chen,Z., and Zhao,H. (2005) *J. Mol. Biol*., **348**, 1273–1282.
Earl,D.J., and Deem,M.W. (2004) *Proc. Natl Acad. Sci. USA*, **101**, 11531–11536.
Elena,S.F., and Sanjuan,R. (2003) *Science*, **302**, 2074–2075.
England,J.L., and Shakhnovich,E.I. (2003) *Phys. Rev. Lett*., **90**, 218101.
Gaucher,E.A., Thomson,J.M., Burgan,M.F. and Benner,S.A. (2003) *Nature*, **425**, 285–288.
Gomez-Puertas,P., Martin-Benito,J., Carrascosa,J.L., Willison,K.R. and Valpuesta,J.M. (2004) *J. Mol. Recognit*., **17**, 85–94.
Guo,H.H., Choe,J. and Loeb,L.A. (2004) *Proc. Natl Acad. Sci. USA*, **101**, 9205–9210.
Jacobson,R.H., Zhang,X.J., DuBose,R.F. and Matthews,B.W. (1994) *Nature*, **369**, 761–766.
James,L.C., Roversi,P. and Tawfik,D.S. (2003) *Science*, **299**, 1362–1367.
James,L.C., and Tawfik,D.S. (2003) *Trends Biochem. Sci*., **28**, 361–368.
Jensen,R.A. (1976) *Annu. Rev. Microbiol*., **30**, 409–425.
Kerner,M.J. *et al*. (2005) *Cell*, **122**, 209–220.
Kimura,M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
Kirschner,M., and Gerhart,J. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 8420–8427.
Kohl,N.E., Emini,E.A., Schleif,W.A., Davis,L.J., Heimbach,J.C., Dixon,R.A., Scolnick,E.M. and Sigal,I.S. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 4686–4690.
Li,H., Helling,R., Tang,C. and Wingreen,N. (1996) *Science*, **273**, 666–669.
Maschera,B., Darby,G., Palu,G., Wright,L.L., Tisdale,M., Myers,R., Blair,E.D. and Furfine,E.S. (1996) *J. Biol. Chem*., **271**, 33231–33235.
Matsumura,I., and Ellington,A.D. (2001) *J. Mol. Biol*., **305**, 331–339.
Miller,V. (2001) *J. Acquir Immune. Defic. Syndr*., **26(Suppl 1)**, S34–50.
O'Brien,P.J., and Herschlag,D. (1999) *Chem. Biol*., **6**, R91–R105.
O'Loughlin,T.L., Greene,D.N. and Matsumura,I. (2006) *Mol. Biol. Evol*., **23**, 764–772.
Olsen,M.J., Stephens,D., Griffiths,D., Daugherty,P., Georgiou,G. and Iverson,B.L. (2000) *Nat. Biotechnol*., **18**, 1071–1074.
Prabu-Jeyabalan,M., Nalivaika,E. and Schiffer,C.A. (2000) *J. Mol. Biol*., **301**, 1207–1220.
Rambaut,A., Posada,D., Crandall,K.A. and Holmes,E.C. (2004) *Nat. Rev. Genet*., **5**, 52–61.
Rennell,D., Bouvier,S.E., Hardy,L.W. and Poteete,A.R. (1991) *J. Mol. Biol*., **222**, 67–88.
Romesberg,F.E., Spiller,B., Schultz,P.G. and Stevens,R.C. (1998) *Science*, **279**, 1929–1933.
Schmidt,D.M., Mundorff,E.C., Dojka,M., Bermudez,E., Ness,J.E., Govindarajan,S., Babbitt,P.C., Minshull,J. and Gerlt,J.A. (2003) *Biochemistry*, **42**, 8387–8393.
Schultz,P.G., Yin,J. and Lerner,R.A. (2002) *Angew. Chem. Int. Ed. Engl*., **41**, 4427–4437.
Shoeman,R.L., Kesselmier,C., Mothes,E., Honer,B. and Traub,P. (1991) *FEBS Lett*., **278**, 199–203.
Strack,P.R., Frey,M.W., Rizzo,C.J., Cordova,B., George,H.J., Meade,R., Ho,S.P., Corman,J., Tritch,R. and Korant,B.D. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 9571–9576.
Taverna,D.M., and Goldstein,R.A. (2002) *Proteins*, **46**, 105–109.
Thomson,J.M., Gaucher,E.A., Burgan,M.F., De Kee,D.W., Li,T., Aris,J.P. and Benner,S.A. (2005) *Nat. Genet*., **37**, 630–635.
Thornton,J.W. (2004) *Nat. Rev. Genet*., **5**, 366–375.
Tiana,G., Broglia,R.A. and Provasi,D. (2001) *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys*., **64**, 011904.
Tocchini-Valentini,G.D., Fruscoloni,P. and Tocchini-Valentini,G.P. (2005) *Proc. Natl Acad. Sci. USA*, **102**, 8933–8938.
Tompa,P., and Csermely,P. (2004) *FASEB J*., **18**, 1169–1175.
Varadarajan,N., Gam,J., Olsen,M.J., Georgiou,G. and Iverson,B.L. (2005) *Proc. Natl Acad. Sci. USA*, **102**, 6855–6860.
Vendruscolo,M., Maritan,A. and Banavar,J.R. (1997) *Phys. Rev. Lett*., **78**, 3967–3970.
Vermeer,A.W., and Norde,W. (2000) *Biophys. J*., **78**, 394–404.
Voigt,C.A., Mayo,S.L., Wang,Z.-G. and Arnold,F.H. (2004) In Jen,E. (ed.), *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press, pp. 105–134.
Wagner,A. (2005) *FEBS Lett*., **579**, 1772–1778.
Wain-Hobson,S. (1993) *Curr. Opin. Genet. Dev*., **3**, 878–883.
Wang,J.D., Herman,C., Tipton,K.A., Gross,C.A. and Weissman,J.S. (2002) *Cell*, **111**, 1027–1039.
Yano,T., Oue,S. and Kagamiyama,H. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 5511–5515.
Zhang,J.H., Dawes,G. and Stemmer,W.P. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 4504–4509.

**Received May 2, 2006; revised June 19, 2006; accepted June 22, 2006**

**Edited by Daniel Tawfik**