npg © 1996 Nature Publishing Group http://www.nature.com/naturebiotechnology

That would certainly appear to be so when set within a brief history of protein engineering for solvent resistance. First came some theoretical proposals. Arnold[4] suggested that enzyme stability in organic solvents would be improved by the removal of charged groups from the protein surface, because they would become less favored as the medium dielectric reduced. Conversely, Alvaro and Russell[5] suggested increasing the number of charged groups, on the equally plausible grounds that they would better bind essential water molecules to the protein. Arnold reported[6] the first experimental finding, that the stability of a protease in 84% DMF improved after removal of a surface arginine; however, the best single replacement was found to be Glu, reversing rather than removing the charge. With this unpredictability, it is perhaps not surprising that Arnold's subsequent studies have concentrated on random mutagenesis.

The "directed evolution" approach does still have its limits, at least using current methods. It is fundamentally a step-by-step uphill improvement strategy. The enormous space of possible protein sequences will clearly have multiple optima for a given function. While an uphill climb will eventually find a local optimum, it may not find a better optimum (the global optimum is an unrealistic hope). This would require "downhill" movement. The use of random mutagenesis is also extremely unlikely to deliver amino acid substitutions requiring two-base changes (in one round).

Of course, natural evolution faces exactly the same problems, and it can get round them by, for instance, mutating temporarily silent genes that later are reexpressed with a step-improved function. There is, however, a general problem in attempting to emulate natural evolution too closely. There is an enormous disparity in the numbers of mutants that can be screened (with any current technology) and in the timescales available. But Arnold's approach is clearly still able to make significant progress towards targets where there has been little or no natural selection.

1. Moore, J.C. and Arnold, F.H. 1996. *Nature Biotechnology* **14**:458-467.
2. You, L. and Arnold, F.H. 1995. *Protein Eng.* **9**:77-83.
3. van Tol, J.B.A., Kraayveld, D.E., Jongejan, J.A. and Duine, J.A. 1995. *Biocatalysis. Biotrans.* **12**:119-136.
4. Arnold, F.H. 1990. *Trends Biotechnol.* **8**:244-249.
5. Alvaro, G. and Russell, A.J. 1991. *Methods Enzymol.* **202**:620-643.
6. Martinez, P. and Arnold, F.H. 1991. *J. Amer. Chem. Soc.* **113**:6336-6337.

Unexpected pathways to protein stabilization

Jonathan King

Two papers in this issue provide cogent examples of both the maturation of protein engineering for industrial processes, and its role in revealing the subtleties of protein structure and function. Stempfer et al.[1] show that polycationic hexaArg protein fusions provide a quick and easy route to immobilizing enzymes for reactor use. Lehle et al.[2] show that the disulfide bonds that one might have expected to be essential are in fact dispensable, even though their absence may generate folding problems in bacteria.

The papers are among the many practical fruits emerging from the seminal studies on protein folding and aggregation carried out by Rainier Jaenicke and his students over the past 30 years. The engineering of polyhistidine or polyarginine sequences at protein N-termini or C-termini was introduced as a method for the rapid purification of proteins whose properties might not be well known or well defined. Though these additional sequences may interfere with the biological activities of some proteins, in many cases the activities are unperturbed. Stempfer et al.[1] report a natural next stage in the utility of this form of engineering.

Using α-glucosidase as their model immobilized enzyme reactor, they show that activity of the immobilized form is indistinguishable from wild type. However, it maintains its activity much longer under process conditions: after 10 days at 30°C the preparation retained essentially 100% activity.

The fusion protein route to immobilization does not require covalent modification, which frequently results in loss of activity. This method also offers flexibility with respect to the choice and packing of the matrix material.

One of the important pathways for the noncovalent inactivation of proteins is through aggregation of partially unfolded species. The aggregated state is a kinetic trap that generally prevents renaturation under reaction conditions. One of the clues to the stability of the bound glucosidase is described in a previous paper from Rudolph's group[3]. Using the same construct, α-glucosidase with a polyarginine tail, they showed that if denatured chains were allowed to bind to the column, and then exposed to renaturing buffer while still bound, the refolding yields were two orders of magnitude higher than for the free enzyme. Presumably tethering the chains prevents the partially folded intermediates from associating and aggregating. The methodology thus offers the promise of generating highly efficient refolding and immobilized product in a closely coupled process.

Disulfide bonds have been a focus in the study of protein folding and stability since Christian Anfinsen's studies on the ribonuclease and Thomas Creighton's (EMBL, Heidelberg) elucidation of the disulfide rearrangement pathway in BPTI. An important member of the Kunitz family of trypsin inhibitors is derived from the seeds of the South African legume *Erythrina*. The protein is used in the affinity purification of tissue plasminogen activator. All members of the family have four conserved cysteines that form two disulfide bonds in the native active inhibitors. In contrast to many other disulfide-bonded proteins, *Erythina* trypsin inhibitor (ETI) retains its activity when the protein is fully reduced.

Though disulfides stabilize proteins, the presence of reactive cysteines at the protein surface can cause problems in industrial settings[5]. Lehle et al. substituted alanines for each cysteine in ETI. Given the stability of the reduced state, one might have expected the ETI chains to fold up efficiently in *E. coli*. In fact, chains with the various Cys to Ala substitutions ended up in inclusion bodies, underscoring the necessity of distinguishing folding pathways from the stabilities of folded state. Nonetheless, the Regensburg group was able to refold the ETI chains and achieve high yields of active ETI. Proteins lacking both disulfide bonds, though slightly reduced in their resistance to denaturation, were as active as the wildtype in terms of protease inhibition. This unexpected result points out that it is dangerous to assume that disulfide bonds in a protein from the seed of a bean are performing the same roles as in proteins from mammalian tissues.

Jonathan King is in the department of biology, Massachusetts Institute of Technology, Building 16-535, Cambridge, MA 02139 (jaking@MIT.EDU).

1. Stempfer, G., Holl-Neugebauer, B., Kopetzki, E, and Rudolph, R. 1996. *Nature Biotechnology* **14**:481-484
2. Lehle, K., Kohnert, U., Stern, A., Popp, F., and Jaenicke, R. 1996. *Nature Biotechnology* **14**:476-480.
3. Stempfer, G., Holl-Neugebauer, B., and Rudolph, R. 1996. *Nature Biotechnology* **14**: 329-334.
4. Creighton, T. E. 1988. *Proc. Nat. Acad. Sci. USA* **85**: 5082.