Retracing Enzyme Evolution in the $(\beta\alpha)_8$-Barrel Scaffold

James D. Stevenson, Stefan Lutz, and Stephen J. Benkovic*

The structure–function relationship and evolutionary origin of enzymes containing the $(\beta\alpha)_8$-barrel fold has been extensively studied in recent years.[1] The high abundance of this particular fold in nature (found in approximately 10% of all soluble proteins), coupled with an apparent flexibility in the chemistry carried out by these enzymes, has made the $(\beta\alpha)_8$-barrel an interesting target for protein engineers. Two recent papers[2, 3] report the successful engineering of the chemistry and binding properties of $(\beta\alpha)_8$-barrel enzymes using rational and nonrational design approaches. In conjunction with structural[4] and computational studies,[1a] these results support the theory of widespread recruitment of ancestral $(\beta\alpha)_8$-barrel scaffolds in the evolution of multiple metabolic pathways.

The core $(\beta\alpha)_8$-barrel structure itself is composed of an eightfold repeat of β-strand–loop–α-helix–loop secondary structure elements.[5] These assemble with the eight β-strands forming a central cylindrical parallel β-sheet and with the α-helices forming the outer surface of the barrel (see Figure 1).

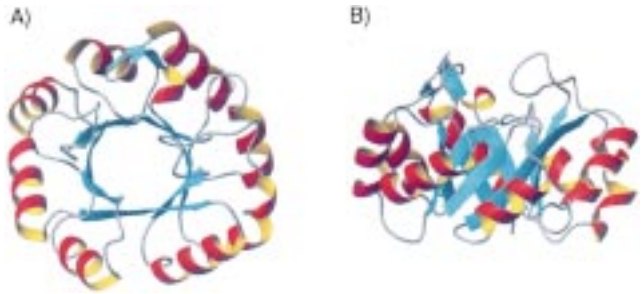


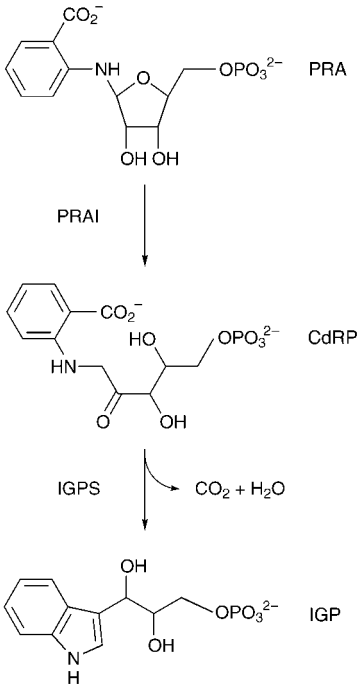
Figure 1. Ribbon representation of the structure of the $(\beta\alpha)_8$-barrel enzyme PRAI. A) Top view from the C-terminal face of the barrel showing the active site of the enzyme. B) Side view of the barrel structure.

Particularly appealing to protein engineers is the apparent spatial separation of the different elements responsible for binding and catalysis. Typically the residues involved in catalysis are located on the loop regions, while the substrate-binding residues are located within the barrel itself. This enables the elements of binding and catalysis to be separately addressed in mutagenesis approaches to modifying enzyme function.

[*] Prof. S. J. Benkovic, J. D. Stevenson, S. Lutz
Department of Chemistry
The Pennsylvania State University
414 Wartik Laboratory
University Park, PA, 16802 (USA)
Fax: (+1) 814-865-2973
E-mail: sjb1@psu.edu

Changing Chemistry, Retaining Binding

Altamirano et al. reported the conversion of one $(\beta\alpha)_8$-barrel into a functional equivalent of another using a combination of rational design and directed evolution.[2] The two enzymes, *N*-(5′-phosphoribosyl)anthranilate isomerase (PRAI) and indole-3-glycerol phosphate synthase (IGPS), catalyze consecutive steps in the tryptophan biosynthetic pathway (see Scheme 1) and are thought to have evolved from



Scheme 1. The consecutive reactions in the tryptophan biosynthetic pathway catalyzed by enzymes PRAI and IGPS. PRAI catalyzes the Amadori rearrangement of *N*-(5′-phosphoribosyl)anthranilate (PRA) to 1′-(2′-carboxyphenylamino)-1′-deoxyribulose 5′-phosphate (CdRP), and IGPS catalyzes the ring closure of CdRP to indole-3-glycerol phosphate (IGP) with release of CO_2 and H_2O.

an ancestral IGPS enzyme through gene duplication.[6] The authors sought to reproduce this evolutionary divergence by converting IGPS from *E. coli* into a functional PRAI. In so doing, they would not only support the evolutionary link

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4010-1854 $ 17.50+.50/0

between these two enzymes, but would also demonstrate the conversion of one enzyme activity into that of another.

The experimental design employed demonstrates an impressive combination of rational and nonrational design (directed evolution) and exploits the power of in vivo metabolic selection (see Figure 2). The authors initially

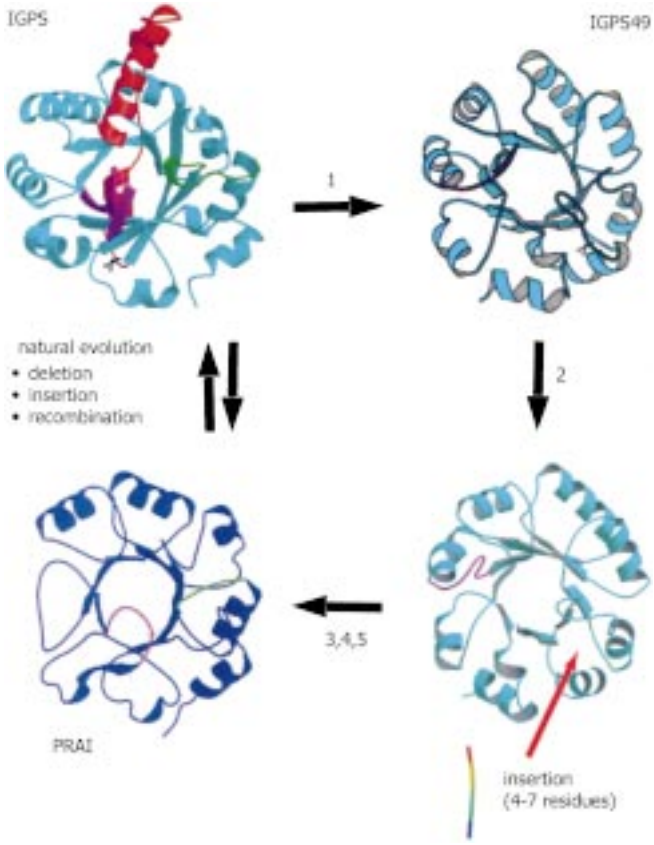


Figure 2. Stepwise-directed evolution approach employed to convert IGPS into a functional PRAI. 1) Deletion of the 48-residue N-terminal extension (shown in red). 2) Substitution of IGPS loop $\beta1\alpha1$ (15 residues in length) with the loop library sequence $GK(X)_{2\text{-}5}$, which ranges from 4–7 residues in length. 3) Replacement of IGPS loop $\beta6\alpha6$ with the PRAI consensus sequence DGXGGQG. 4) In vivo metabolic selection. 5) Two rounds of DNA shuffling and StEP polymerase chain reaction recombination mutagenesis, followed by in vivo metabolic selection to generate in vitro evolved PRAI (ivePRAI). Reproduced with permission from ref. [2]. Copyright (2000) Macmillan Magazines Ltd.

identified the regions in the tertiary structure of IGPS which differ significantly from that of PRAI; they reasoned that these regions would primarily account for the differences in the chemistry carried out by the two enzymes. In a series of structural modifications, the authors removed the N-terminal helix from IGPS (which is not present in PRAI) and introduced loop library sequences into the equivalent loop regions in PRAI (loops $\alpha1\beta1$ and $\alpha6\beta6$). This strategy essentially switches the active sites between the two enzymes.

In order to select functional members from this combinatorial library with different loops, an auxotrophic *E. coli* strain (JA300) deficient in PRAI activity was employed. In the absence of exogenous tryptophan only cells which express a functional PRAI enzyme will grow. Although none of the library members were able to grow in the absence of tryptophan, at very low concentrations (0–25 $\mu g\,mL^{-1}$ of tryptophan), a small proportion (1.7%) of the library was able to grow. This was taken to indicate the presence of weak PRAI activity in some of the library members. To further enhance this activity, the authors employed a directed evolution approach to generate random diversity in the selected library members. Two rounds of DNA shuffling[7] and staggered extension procedure[8] (StEP) were carried out to generate a large library of random mutants. In vivo library selection identified 360 colonies that were able to grow in the complete absence of tryptophan. The most active mutant protein, termed ivePRAI (in vitro evolved PRAI), displayed sixfold-improved PRAI activity in vitro over the wild-type enzyme, mainly due to an increase in substrate binding, and completely lacked the original IGPS activity.

Sequence analysis revealed that ivePRAI had 90% sequence identity to wild-type IGPS, but only 28% to PRAI, clearly demonstrating the central role of the loop regions to the activity of this $(\beta\alpha)_8$-barrel enzyme. That the function can be so readily altered by targeting these regions for mutagenesis indicates a clear functional plasticity in the $(\beta\alpha)_8$-barrel scaffold. This may account for the prevalence of this structure in nature.

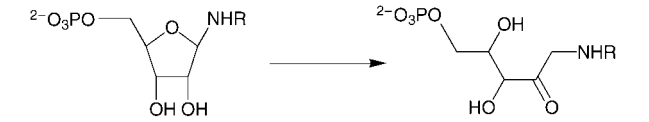
In terms of protein engineering, a number of lessons can be learned from this study. Firstly, that rational and nonrational design may be used in tandem. The initial approach of rational loop grafting was unsuccessful in obtaining a functional enzyme. Only when the diversity-generating power of in vitro recombination was applied, were functional enzymes obtained. To say that this step merely fine-tuned the mutagenesis already in place is possibly understating the importance of this technology. That said, it is unlikely that DNA shuffling alone would have been capable of generating PRAI activity in the IGPS scaffold.

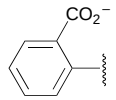
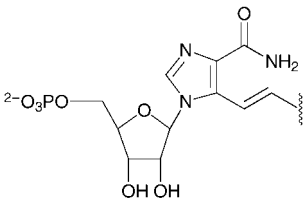
Retaining Chemistry, Changing Binding

In addition to the use of existing enzymes as structural scaffolds, nature has also employed the strategy of tailoring the binding specificity of existing activity to recognize different substrates. This is again apparent with the enzyme PRAI, which has a related homologue in the histidine biosynthetic pathway. The enzyme *N*′-[(5′-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase (ProFARI) catalyzes the Amadori rearrangement of ProFAR to *N*′-[(5′-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (PRFAR; see Table 1). This is the same chemical transformation catalyzed by PRAI, but using a different substrate. Like PRAI, ProFARI is also a $(\beta\alpha)_8$-barrel enzyme, and, while the two enzymes share only 10% sequence identity, this similar structure and chemistry suggests a clear evolutionary link.

Jürgens et al. set out to retrace the divergent evolution of PRAI and ProFARI from a proposed ancestral enzyme displaying broader substrate specificity, by generating variants of ProFARI that were capable of catalyzing both the PRAI and ProFARI reactions.[3] Such a mutant would support the evolutionary link between PRAI and ProFARI. The authors

Table 1. The substrate specificity of ProFARI and PRAI.



R	Substrate	Product	Enzyme
	ProFAR	PRFAR	ProFARI
	PRA	CdRP	PRAI

employed a combination of random mutagenesis and in vivo selection to obtain ProFARI mutants with PRAI activity. In contrast to the elaborate mutagenesis procedure carried out by Altamirano et al.,[2] Jürgens et al. needed only to perform a single round of DNA shuffling in order to achieve their aims. Two variant ProFAR enzymes capable of isomerizing PRA to CdRP were identified from a randomly mutated ProFAR library, by in vivo selection in an auxotrophic *E. coli* strain deficient in PRAI activity. Kinetic analysis showed that both mutants displayed significant PRAI activity, although $3-11 \times 10^4$ times less than wild-type PRAI in terms of the specificity constant ($k_{cat.}/K_M^{PRA}$). In addition, significant ProFARI activity was retained in one of the mutants (26-fold lower than wild-type), while the other displayed much lower levels of activity. It would be expected that carrying out additional rounds of mutagenesis and in vivo selection would result in the further enhancement of PRAI activity at the expense of ProFARI activity. This is due to the focus of the selection criterion purely on PRAI activity.

Of the seven point mutations identified in the two ProFARI variants, only the single point mutation D127V conveyed PRAI activity when reproduced by site-directed mutagenesis. However, this single-site mutant was significantly less active in either reaction than the two selected mutants; this indicates that multiple mutations were acting cooperatively to alter the activity of ProFARI. In turn, the fact that a single amino acid substitution was capable of altering the specificity of ProFARI supports the hypothesis of an evolutionary connection between these two enzyme activities.

In conclusion, while the de novo creation of efficient tailor-made protein catalysts is currently an unattainable goal for protein engineers, this approach of "refitting" an existing enzyme scaffold offers a powerful route to new enzyme activities.[9] As shown by the cases of the two $(\beta\alpha)_8$-barrel enzymes discussed here, either the chemistry or the binding properties can be altered to obtain variant enzymes with the desired new activity. The $(\beta\alpha)_8$-barrel structure is thus being used as a generic scaffold, into which new activities can be introduced. This essentially copies nature itself which, in the course of natural evolution, has generated new enzyme functionality by recruiting existing protein structural and functional folds.

[1] a) R. R. Copley, P. Bork, *J. Mol. Biol*. **2000**, *303*, 627–640; b) N. Nagano, E. G. Hutchinson, J. M. Thornton, *Protein Sci*. **1999**, *8*, 2072–2084, and references therein.

[2] M. M. Altamirano, J. M. Blackburn, C. Aguayo, A. R. Fersht, *Nature* **2000**, *403*, 617–622.

[3] C. Jürgens, A. Strom, D. Wegener, S. Hettwer, M. Wilmanns, R. Sterner, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9925–9931.

[4] D. Lang, R. Thoma, M. Henn-Sax, R. Sterner, M. Wilmanns, *Science* **2000**, *289*, 1546–1550.

[5] D. Reardon, G. K. Farber, *FASEB J.* **1995**, *9*, 497–503.

[6] K. Kirschner, H. Szadkowski, T. S. Jardetzky, V. Hager, *Methods Enzymol.* **1987**, *142*, 386–397.

[7] W. P. Stemmer, *Nature* **1994**, *370*, 389–391.

[8] H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat. Biotechnol.* **1998**, *16*, 258–261.

[9] I. P. Petrounia, F. H. Arnold, *Curr. Opin. Biotechnol*. **2000**, *11*, 325–330.

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 2001 1433-7851/01/4010-1856 $ 17.50+.50/0