the FEBS Journal

In vitro selection and characterization of a stable subdomain of phosphoribosylanthranilate isomerase

Wayne M. Patrick[1,2] and Jonathan M. Blackburn[1,3,4]

1 Department of Biochemistry, University of Cambridge, UK
2 Department of Chemistry, Emory University, GA, USA
3 Department of Biotechnology, University of the Western Cape, Cape Town, South Africa
4 Department of Molecular and Cell Biology, University of Cape Town, South Africa

Keywords
$(\beta\alpha)_8$-barrel; *in vitro* selection; phosphoribosylanthranilate isomerase; plasmid display; subdomain

Correspondence
J.M. Blackburn, Department of Biotechnology, University of the Western Cape, Bellville 7535, Cape Town, South Africa
Fax: +27 21 9591432
Tel: +27 21 9592817
E-mail: jblackburn@uwc.ac.za

(Received 8 April 2005, revised 16 May 2005, accepted 26 May 2005)

doi:10.1111/j.1742-4658.2005.04794.x

The $(\beta\alpha)_8$-barrel is the most common enzyme fold and it is capable of catalyzing an enormous diversity of reactions. It follows that this scaffold should be an ideal starting point for engineering novel enzymes by directed evolution. However, experiments to date have utilized *in vivo* screens or selections and the compatibility of $(\beta\alpha)_8$-barrels with *in vitro* selection methods remains largely untested. We have investigated plasmid display as a suitable *in vitro* format by engineering a variant of phosphoribosylanthranilate isomerase (PRAI) that carried the FLAG epitope in active-site-forming loop 6. Trial enrichments for binding of mAb M2 (a mAb to FLAG) demonstrated that FLAG-PRAI could be identified from a 10^6-fold excess of a FLAG-negative competitor in three rounds of *in vitro* selection. These results suggest PRAI as a useful scaffold for epitope and peptide grafting experiments. Further, we constructed a FLAG-PRAI loop library of $\approx 10^7$ clones, in which the epitope residues most critical for binding mAb M2 were randomized. Four rounds of selection for antibody binding identified and enriched for a variant in which a single nucleotide insertion produced a truncated $(\beta\alpha)_8$-barrel consisting of $(\beta\alpha)_{1-5}\beta_6$. Biophysical characterization of this clone, trPRAI, demonstrated that it was selected because of a 21-fold increase in mAb M2 affinity compared with full-length FLAG-PRAI. Remarkably, this truncated barrel was found to be soluble, structured, thermostable and monomeric, implying that it represents a genuine subdomain of PRAI and providing further evidence that such subdomains have played an important role in the evolution of the $(\beta\alpha)_8$-barrel fold.

An oft-quoted estimate is that there are 1000 structurally distinct protein families in nature [1]. Of all these families, the $(\beta\alpha)_8$-barrels are particularly prominent in terms of both their sheer abundance and also their remarkable functional diversity. The $(\beta\alpha)_8$-barrel is the most commonly occurring enzyme fold in the RCSB Protein Data Bank (PDB) and it has been estimated that 10–12% of all enzymes include a $(\beta\alpha)_8$-barrel domain [2,3]. Proteins possessing this architecture are widespread in the central pathways of metabolism and populate five of the six primary classes of enzymes (as defined by the Enzyme Commission) [4]. Archetypal examples include a perfect catalyst (triosephosphate isomerase) [5], an extremely proficient enzyme (orotidine 5′-monophosphate decarboxylase) [6] and the most abundant protein on

Abbreviations
CdRP, 1′-(2′-carboxyphenylamino)-1′-deoxyribulose 5′-phosphate; IGPS, indoleglycerol-phosphate synthase; PRA, *N*-(5′-phosphoribosyl)anthranilate; PRAI, phosphoribosylanthranilate isomerase; SPR, surface plasmon resonance; trPRAI, truncated PRAI variant consisting of $(\beta\alpha)_{1-5}\beta_6$.

earth (ribulose 1,5-bisphosphate carboxylase/oxygenase, Rubisco) [7].

The diversity of function among $(\beta\alpha)_8$-barrel proteins is ascribable to the apparently modular construction of the fold. It is characterized by secondary structure consisting of eight β-strand–α-helix units which are closed into a cylindrical topology by hydrogen-bonding between the first and last β-strands. The α-helices therefore pack around the central, parallel β-barrel (Fig. 1). This arrangement effectively partitions those parts of the $(\beta\alpha)_8$-barrel important for catalysis (the C-terminal residues of each β-strand and the loops that connect each strand, β_n, to the following helix, α_n) from those that stabilize the overall fold (the core β-barrel and the loops connecting α_n to β_{n+1}) [8].

In delineating so clearly the structurally and functionally important parts of the $(\beta\alpha)_8$-barrel molecule, nature has arrived at a mechanism for altering catalytic activity by mutation, without compromising stability. Functional groups delivered from the eight β-strand–loop units can be positioned in the active site at virtually any position relative to the bound substrate, and, importantly, these functional groups and associated units of secondary structure can evolve with some degree of independence [9]. This combinatorial complexity, introduced by the ability to 'mix and match' active-site-forming units, is thought to have been central to the functional diversification of the $(\beta\alpha)_8$-barrels throughout evolution.

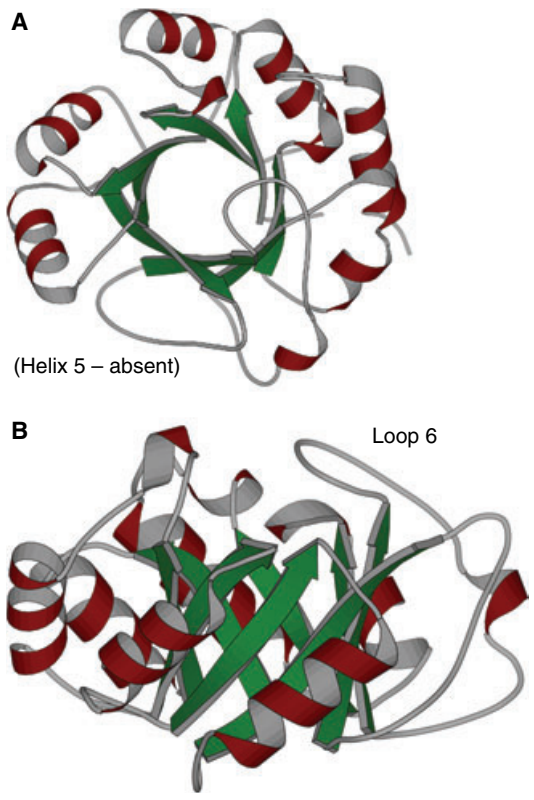

Fig. 1. The *E. coli* $(\beta\alpha)_8$-barrel protein PRAI, viewed from (A) the C-terminal face of the β-barrel and (B) side on (with helix 4 nearest the viewer). Note that, unlike the archetypal $(\beta\alpha)_8$-barrel structure, helix 5 is absent from PRAI. Loop 6 forms a flexible lid over the active site.

It follows that the $(\beta\alpha)_8$-barrel scaffold should be an ideal starting point for engineering novel enzymatic activities by rational redesign or directed evolution. In particular, it has long been hypothesized that varying the residues of the active-site-forming loops might alter enzymatic function without affecting the stability of the fold [10]. A number of recent reports appear to bear out this assertion [11–14]. However, in each case only a small number of variants were assessed by whole cell-based screening or selection, and only slight improvements in the desired activities were observed. It seems apparent that more ambitious loop replacement and randomization strategies will be required to realize the full potential of the $(\beta\alpha)_8$-barrel architecture for engineering new enzymes. However, the very combinatorial complexity that has been so critical in evolution also ensures that any directed evolution experiment involving the randomization of multiple loops will require the interrogation of vast libraries of variants. Moreover, it is recognized that many of the properties targeted by directed evolution are not those that can be easily linked to *in vivo*, life or death selection [15].

The limitations of *in vivo* screens and selections could be overcome by the development of an effective *in vitro* selection methodology. Significantly, however, the compatibility of $(\beta\alpha)_8$-barrel enzymes with *in vitro* systems remains largely untested, and the absence of a robust system somewhat limits the potential for redesigning these proteins. The $(\beta\alpha)_8$-barrel proteins are predominantly cytoplasmic, often co-ordinate cofactors or metal ions, and can be sensitive to oxidative inactivation through nonspecific disulfide formation, all of which complicate their production and selection *in vitro*. To our knowledge, the only examples of *in vitro* selection on the scaffold are the display of a secreted, thermostable α-amylase from *Bacillus licheniformis* on the surface of phage fd [16], and the selection of phosphotriesterase variants from a microbead-displayed library using *in vitro* compartmentalization [17].

In this study, we have investigated plasmid display [18] as an *in vitro* display format with general applicability for the directed evolution of $(\beta\alpha)_8$-barrel proteins. In this approach (Fig. 2), the polypeptides of

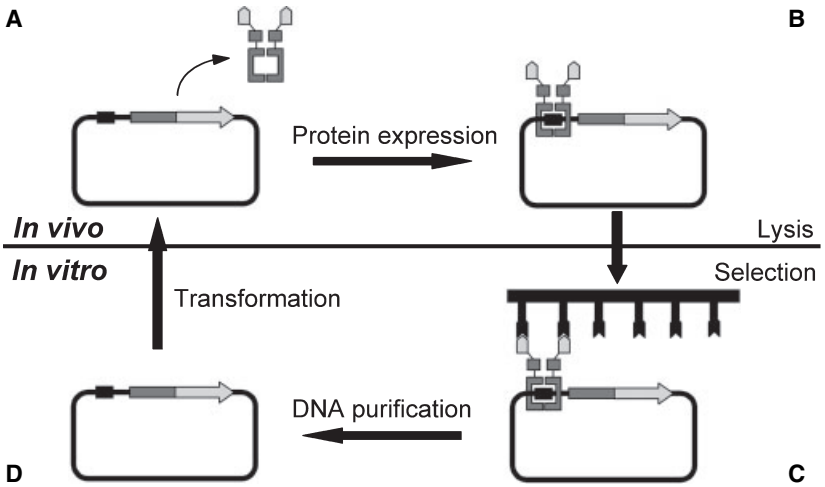

Fig. 2. One round of selection by plasmid display. (A) The protein of interest (light grey) is expressed in the cytoplasm of *E. coli*, fused to NF-κB p50 (dark grey). (B) The fusion protein binds to the p50 recognition sequence on the plasmid (black box), in turn repressing further transcription. (C) On cell lysis, specific fusion protein–plasmid complexes are selected *in vitro* by binding to immobilized ligands. (D) Selected plasmids are recovered and characterized, or used as the substrate for further rounds of enrichment.

a library are expressed from a plasmid vector fused to the DNA binding protein NF-κB p50. Inclusion of an idealized p50 target site on the plasmid establishes a phenotype–genotype linkage within the cell. Polypeptides are folded in the cytoplasm (rather than the periplasmic space, as in filamentous phage display), increasing the likelihood of correct folding in a reducing environment while minimizing the risk of proteolytic degradation. The association of plasmid and fusion protein can be maintained on cell lysis, and selection is carried out *in vitro*.

We selected *N*-(5′-phosphoribosyl)anthranilate isomerase (PRAI, EC 5.3.1.24) from *Escherichia coli* as our target for validating plasmid display, and for addressing the hypothesis that the active-site-forming loops of $(\beta\alpha)_8$-barrel proteins could be regarded as modular with respect to the rest of the scaffold. PRAI catalyzes the Amadori rearrangement of *N*-(5′-phosphoribosyl)anthranilate (PRA) to 1′-(2′-carboxyphenylamino)-1′-deoxyribulose 5′-phosphate (CdRP) [19], which is the third step in the synthesis of tryptophan from chorismic acid. CdRP is in turn the substrate for indoleglycerol-phosphate synthase (IGPS). Although PRAI is part of a bifunctional IGPS–PRAI enzyme in *E. coli*, the two domains have been separated genetically and expressed as stable, monomeric proteins with virtually full catalytic activity [20]. The PRAI enzymes from *E. coli* and *Saccharomyces cerevisiae* were also the targets of a number of pioneering protein engineering experiments undertaken by Kirschner and colleagues. The yeast enzyme was modified by circular permutation [21], duplication of the final two (βα) units [22], and fragmentation into $(\beta\alpha)_{1-6}$ and $(\beta\alpha)_{7-8}$ substructures [23]; *E. coli* PRAI was subjected to an internal duplication of the fifth (βα) unit [24]. Retention of at least trace activity in all cases underlined the apparent thermodynamic advantage inherent in the folding of the $(\beta\alpha)_8$-barrel scaffold. More recently, *S. cerevisiae* PRAI has also been explored as a novel, cytoplasmic split-protein sensor for the detection of protein–protein interactions [25].

In *E. coli* PRAI, the loop connecting β-strand 6 with helix 6 ('loop 6') forms a long and flexible lid over the top of the active-site pocket (Fig. 1). We have investigated the mutability of this loop by the insertion of the FLAG epitope, an antibody-selectable marker [26]. The selection of PRAI proteins carrying a functional FLAG epitope from an excess of FLAG-negative competitors and from a large library of random variants was also undertaken by plasmid display, to confirm the efficacy of this method for engineering $(\beta\alpha)_8$-barrel proteins.

Results

Stable display of the FLAG epitope

Sequence encoding the FLAG epitope and six linker amino acids (AGSDYKDDDDKGSA, FLAG sequence underlined) was introduced into the *trpF* gene for PRAI by overlap extension PCR, replacing three loop 6 codons (for Ser385–Gln387, numbered according to their positions in the bifunctional IGPS–PRAI enzyme). FLAG-PRAI and PRAI itself were overexpressed in *E. coli* strain XL1-Blue. Both proteins accumulated in the soluble, intracellular fractions of induced cultures and were purified to near homogeneity by using C-terminal His_6 tags. Final yields of purified protein were 30–50 mg per litre of induced culture.

FLAG-PRAI showed no detectable catalytic activity (i.e. conversion of PRA into CdRP; data not

shown). This was in contrast with a variant carrying the insertion of a duplicated 24-residue (βα) module in loop 5 [24], but consistent with a proposed critical role for loop 6 (which is mobile and thought to adopt different conformations in unliganded and ligand-bound states) in binding PRA. More importantly for this study, its soluble over-expression suggested that PRAI was able to accommodate insertion of the FLAG epitope without significantly perturbing the folding of the underlying $(\beta\alpha)_8$-barrel. This was investigated by comparing the far-UV and near-UV CD spectra of PRAI and FLAG-PRAI in a buffer in which PRAI retains catalytic activity (Fig. 3). The spectra are effectively superimposable in both cases. The far-UV spectra are also consistent with those observed previously for PRAI with and without a loop 5 insertion [24], albeit at an increased resolution in the present study.

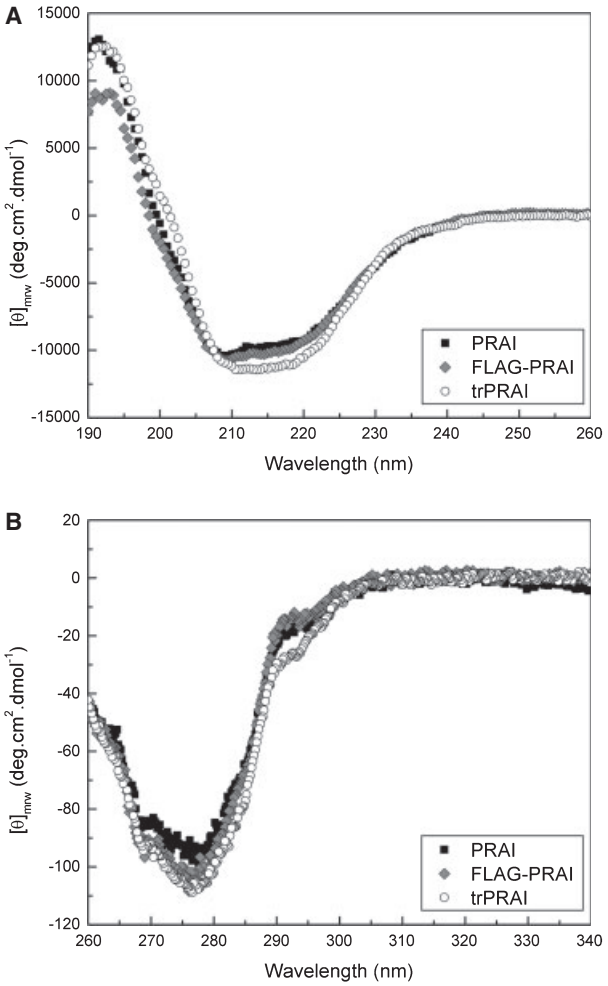

Fig. 3. CD analyses. (A) Far-UV CD spectra of the two full-length proteins PRAI and FLAG-PRAI, and the subdomain trPRAI-His. (B) Near-UV CD spectra of the same proteins. All spectra represent the mean of eight traces.

Trial enrichments demonstrate *in vitro* selection

To demonstrate that plasmid display could be used for *in vitro* selection of $(\beta\alpha)_8$-barrel proteins, the enrichment of FLAG-PRAI from a large excess of a FLAG-negative competitor was undertaken. Selection was based on affinity for mAb M2 (a mAb to FLAG). The competitor used was identical with FLAG-PRAI except that it included the sequence LGLDDADK in place of the FLAG epitope; this was shown to be unreactive in western blots with the antibody.

PRAI forms the C-terminal domain of the bifunctional IGPS–PRAI enzyme in *E. coli*; this arrangement was mimicked by fusing p50 to the N-terminus of the displayed proteins. The vector used for plasmid display was pRES112 [27], in which the p50 DNA binding site (5′-GGGAATTCCC-3′) is located in the −10 region of the *lac* promoter used to drive fusion protein expression. This insertion, which is essential for association of protein and plasmid during selection, has been shown not to affect the intrinsic strength of the promoter, although it does disrupt the LacI binding site and therefore make induction with isopropyl β-D-thiogalactoside unnecessary [28]. Moreover, this design feature effectively regulates expression: translated p50 acts as a repressor of its own synthesis, preventing the production of excess protein molecules that may bind nonself plasmids during *in vitro* selection.

Two trial enrichments were carried out, in which cells expressing p50–FLAG-PRAI were diluted 10^3-fold and 10^6-fold in a background of the FLAG-negative competitor. The number of FLAG-negative cells used in the enrichments was fixed at 10^{10}; the 10^{-3} dilution therefore contained $\approx 10^7$ cells carrying FLAG-PRAI, and the 10^{-6} dilution contained a mere 10 000 FLAG-positive cells. Multiple rounds of affinity selection for mAb M2 were carried out using a 96-well plate format adapted from the basic plasmid display methodology [18]. The use of anti-mouse IgG as an intermediary in the immobilization process (see Experimental procedures for details) was found to increase the yield of selected plasmids, presumably by facilitating a uniform presentation of active mAb M2 molecules available for FLAG epitope recognition. Each cycle of selection was completed in less than 24 h, and successive rounds of selection and re-transformation were assessed by colony western blotting using mAb M2. The results are summarized in Table 1, and representative blots for the 10^{-6} dilution are shown in Fig. 4.

Table 1. Colony blotting demonstrates selection for FLAG-PRAI to near homogeneity over successive rounds of plasmid display. NA, not applicable.

Selection round	No. of colonies recovered	No. FLAG-positive	% FLAG-positive	Enrichment factor
1 in 10^3		0 of 48	(0.1)	NA
1	35 000	26 of 172	15	150-fold
2	1500 000	138 of 172	80	5.3-fold
1 in 10^6		0 of 166	(0.0001)	NA
1	240 000	0 of 172	< 0.6	340-fold[a]
2	360 000	20 of 172	12	340-fold[a]
3	5800 000	158 of 172	92	7.9-fold

[a] Enrichment factors per round estimated by taking the square root of the total enrichment (116 000-fold) observed between Round 0 and Round 2 (when positive colonies were first detected).

Enrichment of the FLAG-positive clone to near homogeneity was achieved in three rounds of selection or less for each dilution. Enrichments of up to an estimated 340-fold per round of plasmid display were observed, consistent with previously reported results [18]. Indeed, because no positive clones were observed after a single round of selection from the 10^{-6} dilution, it is possible that the actual enrichment factor was much higher here.

Finally, six FLAG-positive clones from each of rounds 1 and 2 (10^{-3} dilution) and rounds 2 and 3 (10^{-6} dilution) were sequenced. In all cases the sequence obtained was identical with that of the input clone from the original dilution, confirming functional epitope display and *in vitro* selection for a full-length $(\beta\alpha)_8$-barrel protein.

Selection from a FLAG-PRAI loop library

The results from trial enrichments demonstrated that plasmid display was suitable for *in vitro* selection on the $(\beta\alpha)_8$-barrel scaffold. To explore the limits of the system further, a *bona fide* FLAG-PRAI loop library was constructed in which the codons for residues D1, Y2, K3 and D6 of the FLAG epitope were randomized. Data from the analysis of alternate FLAG epitopes had previously determined that these amino acids are the most critical for binding mAb M2 [29–31]. However, the same reports also suggested that some variability at these positions was tolerated with retention of antibody binding. For example, a short, linear peptide epitope containing threonine at position 1 was selected by CIS display [31], while the D1E and D6E mutations led to sixfold and 1.3-fold decreases in affinity for mAb M2, respectively [30]. In contrast with the binary mixtures of the trial enrichments, the epitope library was therefore expected to contain variants spanning a spectrum of affinities for the selection matrix (i.e. mAb M2). Consequently, this represented a stringent test of plasmid display, particularly as a notable feature of other *in vitro* selection methodologies is an apparent inability to discriminate the highest affinity (or activity) variant in the presence of similar but less effective competitors.

To avoid the possibility of contaminating the library with previously constructed plasmids encoding selectable fusion proteins, the template used for epitope randomization was the FLAG-negative variant from the trial enrichments (Table 2). The effective size of the FLAG-PRAI loop library (from which a vector-derived background of < 1% had been subtracted) was 7.3×10^6 clones. DNA sequence information was obtained for 24 randomly selected variants. Each sequence was unique and none contained more than one parental codon, indicating that the library was suitably diverse. Randomizing four amino acid positions with NNS codons (N = G/A/T/C; S = G/C) generates approximately one million DNA sequence variants ($32^4 = 1\,048\,576$). In the absence of nucleotide bias, our library completeness statistic [32] indicated that the FLAG-PRAI library therefore contained sufficient degeneracy to include > 99.9% of these possible sequences.

The FLAG-PRAI library was subjected to four rounds of selection for a regenerated epitope using the

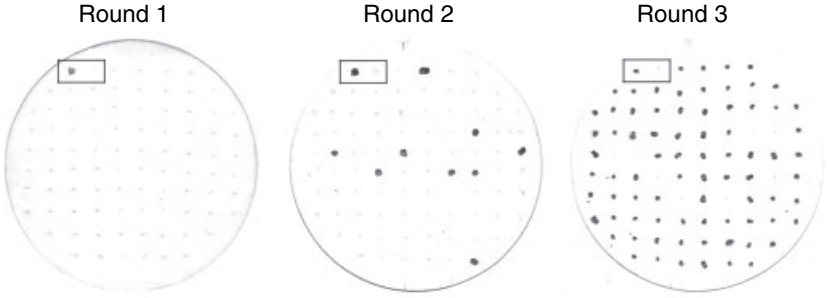

Fig. 4. Colony blots of 86 clones from each round of selection from a 1 in 10^6 dilution of FLAG-PRAI in a background of a FLAG-negative competitor. The two clones used in the enrichment were also used as controls for the colony blot (boxed): FLAG-PRAI (left) and FLAG-negative competitor (right).

Table 2. Summary of epitope sequences in the FLAG-PRAI loop library. Insertion of a thymine nucleotide (bold, underlined) led to an altered epitope and a truncated $(\beta\alpha)_8$-barrel.

Clone	Epitope sequence (5′ → 3′)							
FLAG epitope	GAC	TAC	AAG	GAT	GAC	GAT	GAT	AAG
	D	Y	K	D	D	D	D	K
Library template	TTG	GGG	CTG	GAT	GAC	GCG	GAT	AAG
(FLAG-negative)	L	G	L	D	D	A	D	K
Randomization	NNS	NNS	NNS	GAT	GAC	NNS	GAT	AAG
	X	X	X	D	D	X	D	K
Selected variant	GAC	TAC	AAG	GAT	GAC	GA**T**	CGA	TAA
(FLAG-positive)	D	Y	K	D	D	D	R	*

same experimental protocol as the trial enrichments. Colony western blotting demonstrated the identification and continued enrichment of FLAG-positive library variants over successive rounds of selection (Fig. 5; Table 3). DNA sequence information was obtained for 33 of the positive clones (all positive variants identified in rounds 2 and 3, and 20 of the 38 identified in round 4). In every case, we observed a frameshift caused by insertion of a single thymine nucleotide into the fourth randomized codon of an otherwise wild-type epitope (Table 2). The frameshift produced a novel epitope (DYKDDDR), truncated the p50–PRAI fusion protein immediately after the

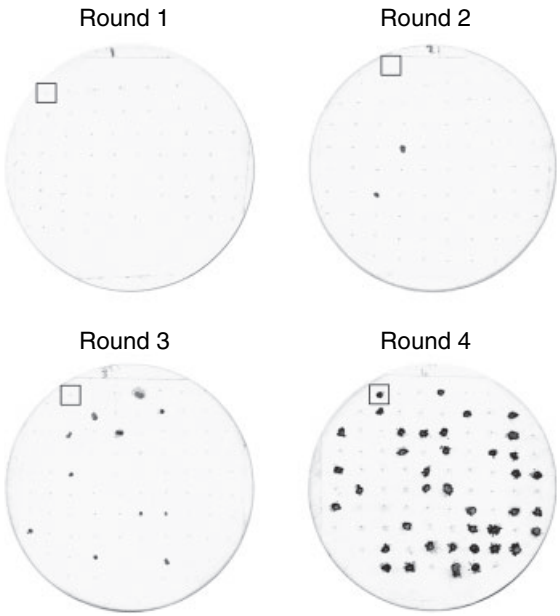

Fig. 5. Colony blots demonstrate the selection and continued enrichment of a FLAG-positive library variant. The boxed clone on each filter is a control: *E. coli* carrying the FLAG-negative library template (rounds 1–3); and one of the previously selected positive clones (round 4).

Table 3. Colony blotting data for selection from the FLAG-PRAI loop library. Selection round 0, unselected library; NA, not applicable.

Selection round	No. of colonies recovered	No. FLAG-positive	% FLAG-positive
0	NA	0 of 87	< 1.1
1	15 000	0 of 68	< 1.5
2	22 000	2 of 87	2.3
3	46 000	11 of 87	13
4	230 000	38 of 87	44

arginine of the new epitope, and removed the fragment of the PRAI $(\beta\alpha)_8$-barrel corresponding to $\alpha_6(\beta\alpha)_{7-8}$. The remaining fragment of PRAI consisted of 130 residues (Gly255–Gly384) and included the first five $(\beta\alpha)$ units of the $(\beta\alpha)_8$-barrel, β-strand 6 and the first four residues of loop 6, before terminating with the altered epitope (a further 10 residues).

Determination of affinities for mAb M2

The nature of the observed insertion suggested that the full-length FLAG epitope was likely to be present in the starting library; however, it is noteworthy that plasmid display did not select it, instead continuing to enrich for the truncated variant through multiple rounds of selection. Positive selection pressure for a $(\beta\alpha)_{1-5}\beta_6$ 'part barrel' at the expense of full-length FLAG-PRAI was hypothesized to reflect an increased affinity for the mAb M2 selection matrix. To test this directly, the truncated PRAI variant (trPRAI) was subcloned without the p50 fusion partner that had been required for plasmid display. Over-expression yielded a protein of the same predicted mass as trPRAI (15.2 kDa), ≈ 50% of which was found in the soluble fraction after cell lysis. The trPRAI deletion removed the C-terminal His_6 tag, making it necessary to purify trPRAI from the soluble cell lysate by its affinity for mAb M2 agarose. Although a rather low binding capacity was observed for this agarose, trPRAI was recovered in sufficient quantities for affinity measurements by surface plasmon resonance (SPR).

On acquiring SPR data for FLAG-PRAI and trPRAI binding to mAb M2, it became apparent that the latter displayed increased binding at any given concentration (Fig. 6). Affinities for the antibody were quantified by analyzing binding data at five concentrations of each protein; as expected, trPRAI displayed a higher affinity for mAb M2 than FLAG-PRAI (Table 4). The equilibrium dissociation constant for trPRAI is ≈ 5.1 nM, a 21-fold improvement over the measured affinity of FLAG-PRAI for the antibody (K_d ≈ 110 nM). The

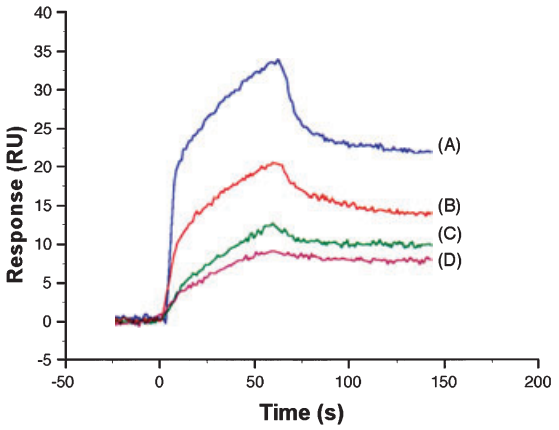

Fig. 6. Sensorgram illustrating increased binding by trPRAI to mAb M2. (A) 1200 nM trPRAI; (B) 1200 nM FLAG-PRAI; (C) 150 nM trPRAI; (D) 150 nM FLAG-PRAI. RU, response units. A baseline response corresponding to nonspecific binding to immobilized BSA has been subtracted from each curve.

Table 4. Kinetic and affinity constants for the binding of FLAG-PRAI and trPRAI to mAb M2. Standard errors for all values are less than 10%.

Protein	k_a ($M^{-1}\cdot s^{-1}$)	k_d (s^{-1})	K_d (M)
FLAG-PRAI	8.7×10^3	9.6×10^{-4}	1.1×10^{-7}
trPRAI	6.0×10^4	3.1×10^{-4}	5.1×10^{-9}

major contribution to this increase in affinity is an approximately sevenfold increase in the second order association rate constant k_a, although the k_d data demonstrate that trPRAI also dissociates from mAb M2 threefold more slowly than FLAG-PRAI. The SPR data therefore confirmed that trPRAI was selected on the basis of its greater affinity for the mAb M2 selection matrix.

Biophysical characterization of trPRAI

Soluble expression, nondenaturing purification and SPR analysis of trPRAI all provided strong circumstantial evidence that the truncated variant was structured in solution, and could therefore be considered an autonomously folding subdomain of PRAI. To confirm this, the CD spectra of the His_6-tagged truncated protein, trPRAI-His, were compared with those obtained for PRAI and FLAG-PRAI (Fig. 3). The far-UV spectrum was of the same form as those of the two full-length $(\beta\alpha)_8$-barrels, consistent with retention of a mixed α/β structure. Further, the similar signal intensities across all three spectra implied not only that trPRAI displays secondary structure, but also that it is approximately as structured as full-length PRAI on a per-residue basis. This is in contrast with analogous $(\beta\alpha)_6$ fragments of both *S. cerevisiae* PRAI and the α subunit of tryptophan synthase, which show spectra of a similar shape but of much reduced intensity compared with the full-length protein [23,33]. In even starker contrast with the $(\beta\alpha)_6$ fragment from the yeast enzyme [23], the trPRAI-His near-UV spectrum is also of a similar form and magnitude to that of full-length PRAI. The only major difference is the absence of a shoulder at 291 nm, which is probably attributable to the removal of a tryptophan residue (Trp391) in the truncation.

The stability of trPRAI-His to thermal denaturation was investigated by monitoring ellipticity at 219 nm (Fig. 7A). As observed previously [34], a sharp, symmetric unfolding transition was observed for PRAI, with the midpoint at 43 °C. The unfolding of trPRAI-His was more gradual, although with a very similar midpoint (T_m = 42 °C).

PRAI contains two tryptophan residues (Trp356 in β-strand 5 and Trp391 in helix 6), the second of which is absent from trPRAI-His. As expected, then, comparison of the fluorescence emission spectra of the two proteins (Fig. 7B) shows a decrease of ≈50% in the total relative fluorescence of the latter. Interestingly, the emission maximum of trPRAI-His is also blue-shifted by 5 nm, from 340 nm to 335 nm. This is consistent with the more solvent-exposed of the two tryptophans (i.e. Trp391) being deleted; however, the implication is also that Trp356 remains in a buried, hydrophobic environment.

Given the nature of the deletion and the presumed energetic advantage in shielding hydrophobic core residues such as Trp356 from the solvent, it seemed unlikely that trPRAI-His could exist as a monomer without dramatic repacking of its secondary structural elements. The oligomeric states of PRAI and trPRAI-His were therefore compared using size exclusion chromatography. As expected, PRAI (molecular mass 22.1 kDa) was eluted as a single peak with a predicted mass of 22.8 kDa, corresponding in size to a monomer (Fig. 7C). Rather more surprisingly, trPRAI-His (molecular mass 15.1 kDa) was also found in a single fraction, eluting with a predicted mass of 18.3 kDa (Fig. 7C). The combined data suggest, then, that trPRAI-His adopts a unique, compact and monomeric conformation in solution.

Discussion

In vitro selection by plasmid display

This study underlines the modularity and mutability of the active-site-forming loops of $(\beta\alpha)_8$-barrel proteins

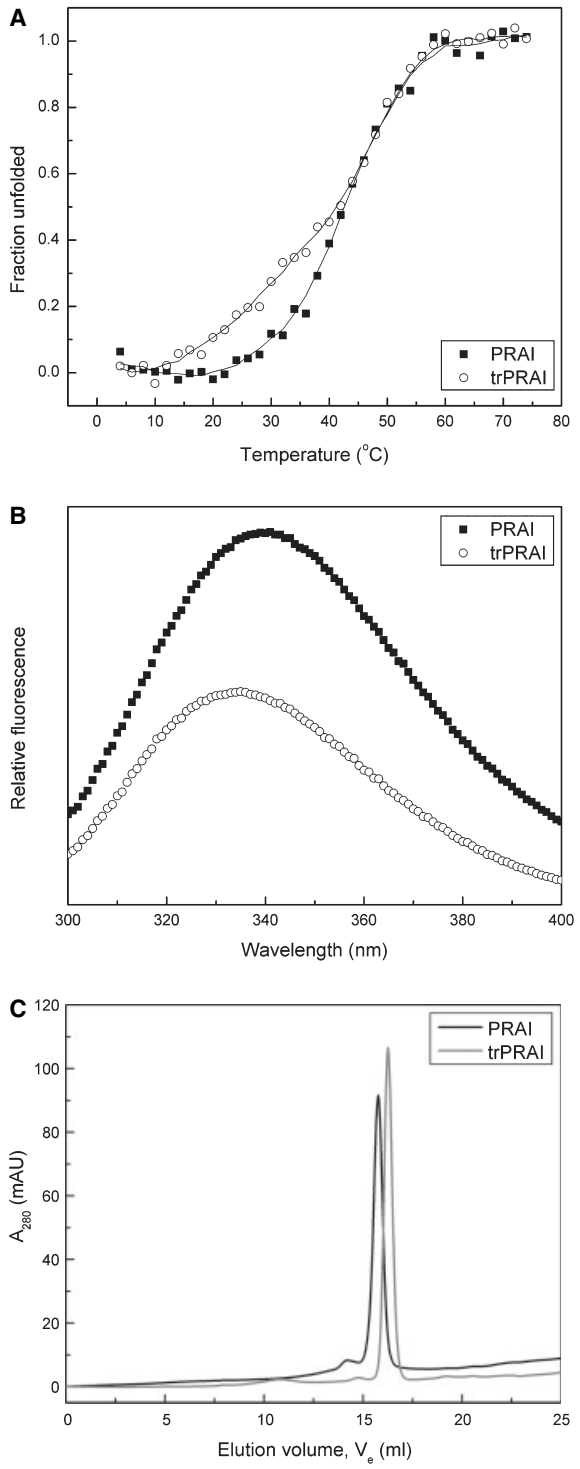

Fig. 7. Biophysical characterization of the PRAI subdomain. (A) Thermal denaturation of PRAI and trPRAI-His as monitored by CD at 219 nm. Raw data and smoothed curves are shown. (B) Fluorescence emission spectra of PRAI and trPRAI-His. The excitation wavelength was 280 nm and the emission signals have been normalized for protein concentration. (C) Profiles of purified PRAI and trPRAI-His eluted from a Superdex 200 gel filtration column.

such as PRAI. In particular, the CD spectra of PRAI and FLAG-PRAI were almost superimposable (Fig. 3), providing strong evidence that all elements of secondary and tertiary structure, and by implication the $(\beta\alpha)_8$-barrel architecture itself, remained intact. This is in spite of an insertion that doubled the length of loop 6 (from 11 to 22 residues) and contained potentially disruptive, charged residues (five aspartates and two lysines).

Although the $(\beta\alpha)_8$-barrel of FLAG-PRAI remained unperturbed, trial enrichments and SPR analysis demonstrated functional presentation of an epitope with nanomolar affinity for its cognate antibody. Further, all the variants selected from our FLAG-PRAI loop library encoded residues of the parental FLAG epitope at the randomized positions, confirming that these residues (D1, Y2, K3 and D6) are the most important for antibody recognition, both in the context of synthetic [30] or displayed [31] peptides, and for the protein scaffold analyzed here. Interestingly though, either a mispriming event during the PCRs and overlap extension used to construct the loop library, or a subsequent point mutation within a bacterium during the first two rounds of selection, transformation and clonal amplification gave rise to an insertion in what otherwise would have been the unmutated FLAG epitope. The preferential selection of the resulting, truncated part-barrel has provided further proof of the maxim that 'you get what you select for' [15] – in this case, the epitope that has the highest affinity for mAb M2. Removing helix 6 and the two final $(\beta\alpha)$ units of the PRAI $(\beta\alpha)_8$-barrel concomitantly removed any structural constraints imposed on the FLAG epitope by being tethered at both ends within loop 6. Presumably it was this new-found conformational freedom that accounted for the 21-fold increase in affinity for mAb M2 of trPRAI over FLAG-PRAI.

Statistical analysis of our library showed that it was > 99.9% complete, so it seemed reasonable to assume that it contained full-length FLAG-PRAI. The observation that this protein was selectable (*viz*. the trial enrichment data described above) but that it was not actually selected therefore confirmed the ability of the plasmid display system to enrich selectively the highest affinity species in the presence of other closely related, but lower affinity, species. This result has often appeared difficult to achieve with other display systems. For example, three rounds of phage display [29] or five rounds of CIS display [31] identified diverse ranges of low-affinity FLAG derivatives, and selection for phosphotriesterase activity from oil-in-water emulsions yielded 35 clones, each with different sequences [17]. We suggest that the greater discriminatory power of plasmid

display may be a unique and advantageous feature of this display format.

A stable subdomain of PRAI

Loop 6 was chosen as the site of FLAG epitope insertion because of its expected tolerance to mutation (*vide supra*). The discovery of the $(\beta\alpha)_{1-5}\beta_6$ subdomain, trPRAI, through *in vitro* selection was therefore a serendipitous result of our engineering strategy. It is not immediately clear from our data whether, had another loop of PRAI been chosen as the original point of epitope insertion, an analogous truncation at that loop would have led to the expression of a selectable variant. However, the requirement for any truncated variant to remain folded, soluble and free from degradation in order to survive multiple rounds of *in vitro* selection suggests that this result is unlikely to be common to the other loops. Experiments to test the mutability of the remaining active-site loops in PRAI have now been initiated.

The biophysical characterization of trPRAI has demonstrated the remarkable robustness of the $(\beta\alpha)_8$-barrel architecture. Despite deletion of one quarter of the strands that make up the β-barrel core of the protein, CD and fluorescence data suggest that trPRAI retains the same degree of α/β structure as PRAI and that it is almost as thermostable as the full length protein (Figs 3 and 7). Moreover, size exclusion chromatography demonstrated that trPRAI is exclusively monomeric in solution (Fig. 7C), albeit with a Stokes radius slightly larger than that expected for a tightly packed, globular protein of the same mass.

Basic principles of protein folding suggest that, to remain monomeric, trPRAI must repack its secondary structural elements to shield newly exposed hydrophobic surfaces, while simultaneously disfavouring the formation of higher-order multimers or aggregates. Further, the possibility that trPRAI-His exists in a molten globule state is precluded by its near-UV CD spectrum. Examination of the high-resolution structure of PRAI (PDB code 1PII [35]) suggests that only three of the 14 residues contributing to the interior of the β-barrel – Leu403, Ala405 and Asp425 – are absent from trPRAI. Perhaps importantly, the residues contributing to one of the hydrogen bonds in the core of the barrel, Lys258 (β1) and Gln332 (β4), remain in trPRAI. It is tempting to speculate that, in the absence of the salt bridge linking Lys258 and Asp425 (β8), the hydrogen bond donor Lys258 could instead be involved in closing a new, six-stranded structure. This would involve contacts with a now-skewed strand β6; a candidate hydrogen bond acceptor could be Asp379. Ultimately though, further structural studies will be required to reveal the true nature of this PRAI subdomain.

Evolution of $(\beta\alpha)_8$-barrels

Gerlt and others [9,36,37] have suggested that loop modularity would have been a convenient device in the evolution of $(\beta\alpha)_8$-barrel enzyme superfamilies, as the semiautonomous evolution of critical functional groups could have allowed the generation of novel binding and catalytic activities in a combinatorial manner. In the case of PRAI, it is now apparent that loops 5 [24] and 6 (this work) satisfy this requirement for evolvability. Moreover, although the active-site-forming loops of PRAI [and indeed, other $(\beta\alpha)_8$-barrel enzymes] undoubtedly require some degree of co-operativity to pack and to confer enzymatic activity, the mutability of two of these loops in isolation offers broad scope for further engineering of multiple loops simultaneously.

In the last five years, a substantial body of evidence has accumulated for the existence of autonomously folding subdomains in $(\beta\alpha)_8$-barrel proteins including triosephosphate isomerase [38,39], the $(\beta\alpha)_8$-barrels of histidine biosynthesis [40–42], IGPS [43], and the α subunit of tryptophan synthase [33,44]. Protein folding studies have suggested that PRAI folds through an intermediate consisting of $(\beta\alpha)_{1-5}\beta_6$ [34]. However, a PRAI $(\beta\alpha)_{1-6}$ part barrel was found to be structured [23], and fragment complementation demonstrated that $(\beta\alpha)_{1-4}$ and $(\beta\alpha)_{5-8}$ could associate to yield a functional enzyme *in vivo* [45], obfuscating somewhat the interpretation of the folding result. The experimental selection and characterization of trPRAI therefore constitutes support for the identity of the putative $(\beta\alpha)_{1-5}\beta_6$ folding intermediate in PRAI and perhaps suggests that the $(\beta\alpha)_{1-4}$, $(\beta\alpha)_{1-6}$ and $(\beta\alpha)_{5-8}$ fragments are of lesser evolutionary significance.

Our data lend weight to the hypothesis that $(\beta\alpha)_8$-barrel proteins may not have evolved through divergent evolution from a single ancestor as commonly assumed. Instead, the existence of part barrels such as trPRAI seems to support an alternative scenario in which (re)combinatorial mixing and matching of minigene encoded, autonomously folding subdomains initially gave rise to multiple, ancestral $(\beta\alpha)_8$-barrels by convergent evolution, each of which later underwent more gradual divergent evolution. One advantage of this route to diversification is that it could have given rise to a greater range of functions early in $(\beta\alpha)_8$-barrel evolution. Interestingly, the most comprehensive global analysis to date grouped 889 $(\beta\alpha)_8$-barrels from the PDB into 21 structurally homologous superfamilies, between 17 of which were found 'hints of a common

ancestry' [4]. However, the same study was unable to find evidence for a single common ancestor, nor was it able to rule out convergent evolution to generate multiple lineages of $(\beta\alpha)_8$-barrel proteins, perhaps in accord with an 'ancient convergence, recent divergence' evolutionary model.

A corollary of such a model might be the survival of intermediate, 'subdomain-like' proteins. Two recent reports suggest that these have indeed persisted, albeit with additional elements of secondary structure recruited to provide substrate specificity and/or catalytic competence. In the first, structural homology was observed between the half-barrels of histidine biosynthesis and members of the $(\beta\alpha)_5$ flavodoxin-like fold [46]. In the second, a comprehensive structure-based alignment suggested that members of the *S*-adenosyl-L-methionine radical protein superfamily adopt $(\beta\alpha)_4$, $(\beta\alpha)_6$ and $(\beta\alpha)_8$ architectures, all based around a common, cofactor-binding $(\beta\alpha)_4$ subdomain [47]. It therefore seems likely that the recruitment and assembly of subdomains such as trPRAI has played a critical role in the evolution of the $(\beta\alpha)_8$-barrel fold; experiments are now underway to explore this hypothesis.

Experimental procedures

Materials

Oligonucleotides were obtained from the Protein and Nucleic Acid Chemistry Facility, Department of Biochemistry, University of Cambridge and, in the case of primer Lib2.for, from Gibco BRL (Paisley, UK). Details of all primers are available on request. The construction of all plasmids was verified by DNA sequencing, which was carried out at the DNA Sequencing Facility, Department of Biochemistry, University of Cambridge. *E. coli* XL1-Blue (Stratagene, La Jolla, CA, USA) was used for all cloning and expression. All antibodies for *in vitro* selection, colony western blotting and SPR analyses were from Sigma Chemical Co (St Louis, MO, USA).

Construction, expression and purification of FLAG-PRAI

The template for inserting the FLAG epitope into loop 6 of PRAI by overlap extension PCR [48] was pMS401. This derivative of pJB122 [49] encodes His_6-tagged *E. coli* PRAI and had been tested previously (M. Samaddar and J. M. Blackburn, unpublished data). The mutagenic primers also encoded linker amino acids; the complete insertion into *trpF* was therefore AGSDYKDDDDKGSA. Ligation of the assembled product with pJB122 yielded the new plasmid pWP101.

PRAI and FLAG-PRAI were purified from *E. coli* cultures harbouring pMS401 and pWP101, respectively. After isopropyl thio-β-D-galactoside-induced expression and lysis by sonication, the recombinant proteins were purified using the Talon metal affinity chromatography system (Clontech, Mountain View, CA, USA). Microcon (Amicon Bioseparations, Billerica, MA, USA) or VivaSpin (Vivascience, Hannover, Germany) centrifugal filter devices were used to exchange the purified proteins into filtered, degassed CD buffer (10 mM Tris/HCl, 100 mM NaCl, 400 μM dithiothreitol, pH 8.5). Protein concentrations were quantified by measuring A_{280}; molar absorption coefficients for each protein were calculated as described by Pace *et al.* [50].

CD

Far-UV and near-UV CD spectra were measured on a Jasco (Great Dunmow, Cambs, UK) J-810 spectropolarimeter at 20.0 °C. Far-UV CD spectra were recorded from 260 to 190 nm (0.5 nm increments), using a 0.1 mm pathlength cell, a 2 nm bandwidth, a 4 s response time and a 20 nm·min^{-1} scan rate. Near-UV spectra were collected from 340 to 260 nm (0.2 nm increments), with a 1 cm pathlength cell, a 1 nm bandwidth, a 2 s response time and a 10 nm·min^{-1} scan rate. Proteins were analyzed at concentrations of 0.7–1.1 mg·mL^{-1}, and each spectrum represents the mean of eight accumulation scans. Spectra were corrected for blank absorption and converted into mean residue ellipticity ($[\theta]_{mrw}$).

Plasmid display

The plasmid display vector pRES112 [27] was modified by inserting a 1272 bp DNA fragment at the unique *Sal*I restriction site, to allow its digestion to be monitored to completion. A *Sal*I restriction site was introduced at the 5′ end of the gene encoding FLAG-PRAI by PCR and the product was subcloned, generating pWP103(+) for use in trial enrichments. The gene for the FLAG-negative competitor used in the enrichments had been identified in a previous randomization experiment and was similarly subcloned, producing pWP103(–). Spheroplasts for the trial enrichments were prepared as described [18], immediately after mid-exponential phase *E. coli* carrying pWP103(+) had been diluted in the appropriate culture volume of *E. coli* [pWP103(–)]. The resulting pellets were stored at −80 °C.

The plasmid display selection matrix was prepared by first adsorbing anti-mouse IgG, diluted 1 : 100 in NaCl/P_i (50 mM potassium phosphate, 50 mM NaCl, pH 7.2), to the wells of a MaxiSorp™ microtiter plate (Nalge Nunc International, Rochester, NY, USA) by incubation at room temperature for 3–6 h. The wells were then washed three times in NaCl/P_i with 0.05% (v/v) Tween 20 (NaCl/P_i-T) and three times in NaCl/P_i. mAb M2 was diluted 1 : 500 in

$NaCl/P_i$ and added to each well in a 200-μL aliquot. Immobilization of mAb M2 by its affinity for the adsorbed anti-mouse IgG was by incubation at 4 °C for 12–16 h. After extensive washing in $NaCl/P_i$-T and $NaCl/P_i$, the plate surface was blocked by incubation with 5% (w/v) BSA in $NaCl/P_i$ (room temperature 1–2 h) before excess BSA was removed by washing in $NaCl/P_i$-T, $NaCl/P_i$ and finally in p50 binding buffer [10 mM Tris/HCl, 50 mM potassium glutamate, 10% (v/v) glycerol, 0.02% (v/v) Triton X-100, 3 mM dithiothreitol, 0.1 $mg{\cdot}mL^{-1}$ herring sperm DNA, pH 7.4).

A single spheroplast pellet was resuspended in 1 mL sterile water, causing osmotic lysis. The lysate was diluted to 6 mL with p50 binding buffer, and insoluble debris was removed by centrifugation at 2500 ***g*** for 10 min. *In vitro* selection for mAb M2 binding was by adding 200-μL aliquots of the supernatant to the selection matrix and incubating at room temperature for 30 min. Unselected protein–plasmid complexes were removed by washing five times in p50 binding buffer and twice in p50 binding buffer from which herring sperm DNA had been omitted. Selected plasmids were eluted by the addition of high-salt buffer (10 mM Tris/HCl, 500 mM NaCl, pH 7.4) and incubation at room temperature for 20 min. The plasmids in the selected lysates were pooled, desalted using buffer N3 (Qiagen, Valencia, CA, USA) and the QIAprep Spin Miniprep kit, and used to retransform *E. coli* by electroporation. After overnight growth on Luria–Bertani plates containing carbenicillin (100 $\mu g{\cdot}mL^{-1}$), colonies were harvested by scraping the plates with liquid Luria–Bertani medium. Finally, the pool of selected clones was used to make fresh aliquots of spheroplasts for the next round of selection. Individual colonies from each round of selection were also analyzed for the presence of FLAG-PRAI by western blotting. This was carried out as described previously [18], using mAb M2 as the primary antibody.

FLAG-PRAI loop library construction

The loop library was constructed by overlap extension PCR, using the nonselectable template pWP103(–) and the partially randomized primer Lib2.for (5′-CCAGGGTGG AGCGGGATCCNNSNNSNNSGATGACNNSGATAAG GGTAGTGCAC-3′). The assembled library was amplified in a secondary PCR with outside primers, cloned into the modified pRES112 vector, desalted and used to transform *E. coli*. The library was recovered by scraping the plates in Luria–Bertani liquid medium, yielding 40 mL cell suspension ($D_{600} = 25$), aliquots of which were used to prepare spheroplast pellets for plasmid display. *In vitro* selection and colony western blotting using mAb M2 followed protocols identical with those described for the trial enrichments.

Expression and purification of trPRAI

Subcloning trPRAI for expression (without its p50 fusion partner) required the re-introduction of an initiator methionine codon and was achieved by PCR with primers incorporating the necessary sequence. The product was re-inserted into pMS401, and the resulting plasmid was named pWP107. Induction of trPRAI expression and cell lysis were as described above for His_6-tagged PRAI and FLAG-PRAI. Purification was on anti-FLAG M2 agarose (Sigma), with specific elution by addition of lysis buffer containing the competitor FLAG peptide (100 μM).

Surface plasmon resonance

The affinities of FLAG-PRAI and trPRAI for mAb M2 were analyzed directly using a Biacore 2000 optical biosensor with a streptavidin-coated sensor chip SA (both Biacore AB, Uppsala, Sweden). The first two flow cells of the chip were left underivatized to control for bulk refractive index changes and signal drift. Flow cell 3 was loaded to saturation with biotinylated BSA, to control for non-specific protein–protein interactions. Biotinylated mAb M2 (10 nM, diluted in 50 mM sodium phosphate, pH 5.5) was used to load flow cell 4 with ≈900 response units, to ensure a maximal detectable response on injection of FLAG-labelled proteins.

For each experiment, purified protein was diluted to 150–4800 nM (FLAG-PRAI, five concentrations) or 75–1200 nM (trPRAI, five concentrations) in Tris-buffered saline (50 mM Tris/HCl, 150 mM NaCl, pH 7.4) and injected for 60 s at a flow rate of 40 $\mu L{\cdot}min^{-1}$ over each flow cell. The protein sample was then replaced by Tris-buffered saline alone at the same flow rate, and the PRAI–mAb M2 complexes were allowed to dissociate for 500 s. The chip surface was regenerated with an injection of 10 μL glycine buffer (10 mM, pH 2.0) for 15 s. All solutions were filtered through a 0.22-μm membrane (GS type; Millipore, Billerica, MA, USA) and degassed before use. Assays were carried out at 25.0 °C.

Sensorgrams were analyzed using the BIAEVALUATION version 3.1 software package (Biacore AB). The software's model for 1 : 1 binding (i.e. of the form $A + B \rightarrow AB$, corresponding to the stoichiometry of PRAI–mAb M2 complex formation) was used to fit the data. In the first iteration, rate constants for the dissociation phase at each protein concentration were calculated independently according to a first-order rate equation. These values for k_d were then used to begin a second iteration in which the global, second-order association rate constant (k_a) for all protein concentrations under consideration was estimated. Finally, the global k_a was input as the starting point for iterations to fit the full data sets at all concentrations simultaneously, generating the k_a, k_d and K_d values in Table 4.

Construction of pWP107His

To facilitate high-yield purification, the C-terminal FLAG epitope of trPRAI was replaced with a His_6 tag. This was

achieved by amplifying pWP107 with a reverse primer that introduced the tag and an *ochre* stop codon; after subcloning of the insert, the resulting plasmid was named pWP107His. Expression and purification of trPRAI-His using this plasmid was then as described for PRAI and FLAG-PRAI. The final yield of soluble, purified trPRAI-His was 20–30 mg per litre of induced culture.

Characterization of trPRAI-His

The far-UV and near-UV CD spectra of trPRAI-His were recorded and analyzed in a manner identical with those for PRAI and FLAG-PRAI (*vide supra*). In addition, the thermal melting curves of trPRAI-His and PRAI were compared by monitoring their CD signals at 219 nm. Each protein was diluted to 0.2 mg·mL^{-1} in CD buffer and heated from 4 °C to 76 °C at 1 °C·min^{-1}, in a 1-mm pathlength cell. Data were collected at 2 °C intervals.

The tryptophan fluorescence emission spectra of PRAI and trPRAI-His were measured using a Jasco FMO-427S monochromator fitted to the J-810 spectropolarimeter. The proteins were diluted to 0.4 mg·mL^{-1} (PRAI) and 0.2 mg·mL^{-1} (trPRAI-His) and excited at 280 nm in a 1-cm pathlength cell. Emission data were accumulated at 20.0 °C over the range 300–400 nm (1-nm increments), with a response time of 1 s. Each spectrum represents the mean of four scans, normalized for protein concentration and for fluorescence of a buffer-only control.

The oligomeric states of PRAI and trPRAI-His at room temperature were determined by size exclusion chromatography with a Superdex 200 10/300 GL column and the Äkta FPLC system (both Amersham Pharmacia Biotech). The column was equilibrated and all runs were performed in 50 mM potassium phosphate, 200 mM NaCl, 1 mM dithiothreitol, pH 7.2, at a flow rate of 0.5 mL·min^{-1}. The column was calibrated with acetone for the total volume (V_t = 20.72 mL), blue dextran for the void volume (V_o = 7.86 mL), and the Bio-Rad Gel Filtration Standard set. The partition coefficients of the standards, $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e is the elution volume, were linearly related to log molecular mass. The K_{av} values of PRAI and trPRAI-His were determined by loading 100 μL purified protein (0.9 mg·mL^{-1}) on to the column.

Acknowledgements

We thank Dr Ichiro Matsumura for his critique of this manuscript and Monica Gerth for her assistance in analyzing biophysical data. We also thank Dr Stefan Lutz (Department of Chemistry, Emory University) for the use of facilities. W.M.P. gratefully acknowledges financial support from the Cambridge Commonwealth Trust's Prince of Wales Scholarship and an Overseas Research Student Award. J.M.B. thanks the Royal Society for a University Research Fellowship.

References

1 Chothia C (1992) One thousand families for the molecular biologist. *Nature* **357**, 543–544.
2 Reardon D & Farber GK (1995) The structure and evolution of α/β barrel proteins. *FASEB J* **9**, 497–503.
3 Pujadas G & Palau J (1999) TIM barrel fold: structural, functional and evolutionary characteristics in natural and designed molecules. *Biologia (Bratislava)* **54**, 231–254.
4 Nagano N, Orengo CA & Thornton JM (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J Mol Biol* **321**, 741–765.
5 Albery WJ & Knowles JR (1976) Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* **15**, 5631–5640.
6 Radzicka A & Wolfenden R (1995) A proficient enzyme. *Science* **267**, 90–93.
7 Ellis RJ (1979) The most abundant protein in the world. *Trends Biochem Sci* **4**, 241–244.
8 Höcker B, Jürgens C, Wilmanns M & Sterner R (2001) Stability, catalytic versatility and evolution of the $(\beta\alpha)_8$-barrel fold. *Curr Opin Biotechnol* **12**, 376–381.
9 Babbitt PC & Gerlt JA (1997) Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. *J Biol Chem* **272**, 30591–30594.
10 Brändén C-I (1991) The TIM barrel: the most frequently occuring folding motif in proteins. *Curr Opin Struct Biol* **1**, 978–983.
11 Jürgens C, Strom A, Wegener D, Hettwer S, Wilmanns M & Sterner R (2000) Directed evolution of a $(\beta\alpha)_8$-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc Natl Acad Sci USA* **97**, 9925–9930.
12 Schmidt DM, Mundorff EC, Dojka M, Bermudez E, Ness JE, Govindarajan S, Babbitt PC, Minshull J & Gerlt JA (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry* **42**, 8387–8393.
13 Cheon YH, Park HS, Kim JH, Kim Y & Kim HS (2004) Manipulation of the active site loops of D-hydantoinase, a $(\beta/\alpha)_8$-barrel protein, for modulation of the substrate specificity. *Biochemistry* **43**, 7413–7420.
14 Leopoldseder S, Claren J, Jürgens C & Sterner R (2004) Interconverting the catalytic activities of $(\beta\alpha)_8$-barrel enzymes from different metabolic pathways: sequence requirements and molecular analysis. *J Mol Biol* **337**, 871–879.

15 Zhao H & Arnold FH (1997) Combinatorial protein design: strategies for screening protein libraries. *Curr Opin Struct Biol* **7**, 480–485.
16 Verhaert RM, Beekwilder J, Olsthoorn R, van Duin J & Quax WJ (2002) Phage display selects for amylases with improved low pH starch-binding. *J Biotechnol* **96**, 103–118.
17 Griffiths AD & Tawfik DS (2003) Directed evolution of an extremely fast phosphotriesterase by *in vitro* compartmentalization. *EMBO J* **22**, 24–35.
18 Speight RE, Hart DJ, Sutherland JD & Blackburn JM (2001) A new plasmid display technology for the *in vitro* selection of functional phenotype-genotype linked proteins. *Chem Biol* **8**, 951–965.
19 Hommel U, Eberhard M & Kirschner K (1995) Phosphoribosyl anthranilate isomerase catalyzes a reversible Amadori reaction. *Biochemistry* **34**, 5429–5439.
20 Eberhard M, Tsai-Pflugfelder M, Bolewska K, Hommel U & Kirschner K (1995) Indoleglycerol phosphate synthase-phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from *Escherichia coli* with engineered monofunctional domains. *Biochemistry* **34**, 5419–5428.
21 Luger K, Hommel U, Herold M, Hofsteenge J & Kirschner K (1989) Correct folding of circularly permuted variants of a βα barrel enzyme *in vivo*. *Science* **243**, 206–210.
22 Luger K, Szadkowski H & Kirschner K (1990) An 8-fold (βα)-barrel protein with redundant folding possibilities. *Protein Eng* **3**, 249–258.
23 Eder J & Kirschner K (1992) Stable substructures of eightfold βα-barrel proteins: fragment complementation of phosphoribosylanthranilate isomerase. *Biochemistry* **31**, 3617–3625.
24 Urfer R & Kirschner K (1992) The importance of surface loops for stabilizing an eightfold βα barrel protein. *Protein Sci* **1**, 31–45.
25 Tafelmeyer P, Johnsson N & Johnsson K (2004) Transforming a $(\beta/\alpha)_8$-barrel enzyme into a split-protein sensor through directed evolution. *Chem Biol* **11**, 681–689.
26 Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Cerretti DP, Urdal DL & Conlon PJ (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* **6**, 1204–1210.
27 Speight RE (2000) *The development of a new plasmid display technology for functional genomics*. PhD Thesis, University of Cambridge, Cambridge, UK.
28 Hart DJ, Speight RE, Sutherland JD & Blackburn JM (2001) Analysis of the NF-κB p50 dimer interface by diversity screening. *J Mol Biol* **310**, 563–575.
29 Miceli RM, DeGraaf ME & Fischer HD (1994) Two-stage selection of sequences from a random phage display library delineates both core residues and permitted structural range within an epitope. *J Immunol Methods* **167**, 279–287.
30 Slootstra JW, Kuperus D, Plückthun A & Meloen RH (1997) Identification of new tag sequences with differential and selective recognition properties for the anti-FLAG monoclonal antibodies M1, M2 and M5. *Mol Divers* **2**, 156–164.
31 Odegrip R, Coomber D, Eldridge B, Hederer R, Kuhlman PA, Ullman C, FitzGerald K & McGregor D (2004) CIS display: *in vitro* selection of peptides from libraries of protein–DNA complexes. *Proc Natl Acad Sci USA* **101**, 2806–2810.
32 Patrick WM, Firth AE & Blackburn JM (2003) User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. *Protein Eng* **16**, 451–457.
33 Zitzewitz JA, Gualfetti PJ, Perkons IA, Wasta SA & Matthews CR (1999) Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a β/α barrel protein. *Protein Sci* **8**, 1200–1209.
34 Jasanoff A, Davis B & Fersht AR (1994) Detection of an intermediate in the folding of the $(\beta\alpha)_8$-barrel *N*-(5′-phosphoribosyl)anthranilate isomerase from *Escherichia coli*. *Biochemistry* **33**, 6350–6355.
35 Wilmanns M, Priestle JP, Niermann T & Jansonius JN (1992) Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from *Escherichia coli* refined at 2.0 Å resolution. *J Mol Biol* **223**, 477–507.
36 Gerlt JA & Babbitt PC (2001) Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annu Rev Biochem* **70**, 209–246.
37 Gerlt JA & Raushel FM (2003) Evolution of function in $(\beta/\alpha)_8$-barrel enzymes. *Curr Opin Chem Biol* **7**, 252–264.
38 Silverman JA, Balakrishnan R & Harbury PB (2001) Reverse engineering the $(\beta/\alpha)_8$ barrel fold. *Proc Natl Acad Sci USA* **98**, 3092–3097.
39 Silverman JA & Harbury PB (2002) The equilibrium unfolding pathway of a $(\beta/\alpha)_8$ barrel. *J Mol Biol* **324**, 1031–1040.
40 Lang D, Thoma R, Henn-Sax M, Sterner R & Wilmanns M (2000) Structural evidence for evolution of the β/α barrel scaffold by gene duplication and fusion. *Science* **289**, 1546–1550.
41 Höcker B, Beismann-Driemeyer S, Hettwer S, Lustig A & Sterner R (2001) Dissection of a $(\beta\alpha)_8$-barrel enzyme into two folded halves. *Nat Struct Biol* **8**, 32–36.
42 Höcker B, Claren J & Sterner R (2004) Mimicking enzyme evolution by generating new $(\beta\alpha)_8$-barrels from $(\beta\alpha)_4$-half-barrels. *Proc Natl Acad Sci USA* **101**, 16448–16453.
43 Forsyth WR & Matthews CR (2002) Folding mechanism of indole-3-glycerol phosphate synthase from

Sulfolobus solfataricus: a test of the conservation of folding mechanisms hypothesis in $(\beta\alpha)_8$ barrels. *J Mol Biol* **320**, 1119–1133.

44 Zitzewitz JA & Matthews CR (1999) Molecular dissection of the folding mechanism of the α subunit of tryptophan synthase: an amino-terminal autonomous folding unit controls several rate-limiting steps in the folding of a single domain protein. *Biochemistry* **38**, 10205–10214.

45 Soberón X, Fuentes-Gallego P & Saab-Rincón G (2004) *In vivo* fragment complementation of a $(\beta/\alpha)_8$ barrel protein: generation of variability by recombination. *FEBS Lett* **560**, 167–172.

46 Höcker B, Schmidt S & Sterner R (2002) A common evolutionary origin of two elementary enzyme folds. *FEBS Lett* **510**, 133–135.

47 Nicolet Y & Drennan CL (2004) AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Res* **32**, 4015–4025.

48 Horton RM, Cai ZL, Ho SN & Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* **8**, 528–535.

49 Altamirano MM, Blackburn JM, Aguayo C & Fersht AR (2000) Directed evolution of new catalytic activity using the α/β-barrel scaffold. *Nature* **403**, 617–622.

50 Pace CN, Vajdos F, Fee L, Grimsley G & Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4**, 2411–2423.