doi:10.1016/j.jmb.2007.06.036 J. Mol. Biol. (2007) 372, 114–129

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Stabilisation of a $(\beta\alpha)_8$-Barrel Protein Designed from Identical Half Barrels

Tobias Seitz, Marco Bocola, Jörg Claren and Reinhard Sterner*

Institute of Biophysics and Physical Biochemistry
University of Regensburg
Universitätsstrasse 31
D-93053 Regensburg, Germany

It has been suggested that the common $(\beta\alpha)_8$-barrel enzyme fold has evolved by the duplication and fusion of identical $(\beta\alpha)_4$-half barrels, followed by the optimisation of their interface. In our attempts to reconstruct these events *in vitro* we have previously linked in tandem two copies of the C-terminal half barrel HisF-C of imidazole glycerol phosphate synthase from *Thermotoga maritima* and subsequently reconstituted in the fusion construct HisF-CC a salt bridge cluster present in wild-type HisF. The resulting recombinant protein HisF-C*C, which was produced in an insoluble form and unfolded with low cooperativity at moderate urea concentrations has now been stabilised and solubilised by a combination of random mutagenesis and selection *in vivo*. For this purpose, *Escherichia coli* cells were transformed with a plasmid-based gene library encoding HisF-C*C variants fused to chloramphenicol acetyltransferase (CAT). Stable and soluble variants were identified by the survival of host cells on solid medium containing high concentrations of the antibiotic. The selected HisF-C*C proteins, which were characterised *in vitro* in the absence of CAT, contained eight different amino acid substitutions. One of the exchanges (Y143C) stabilised HisF-C*C by the formation of an intermolecular disulfide bond. Three of the substitutions (G245R, V248M, L250Q) were located in the long loop connecting the two HisF-C copies, whose subsequent truncation from 13 to 5 residues yielded the stabilised variant HisF-C*C Δ. From the remaining substitutions, Y143H and V234M were most beneficial, and molecular dynamics simulations suggest that they strengthen the interactions between the half barrels by establishing a hydrogen-bonding network and an extensive hydrophobic cluster, respectively. By combining the loop deletion of HisF-C*C Δ with the Y143H and V234M substitutions, the variant HisF-C**C was generated. Recombinant HisF-C**C is produced in soluble form, forms a pure monomer with its tryptophan residues shielded from solvent and unfolds with similar cooperativity as HisF. Our results show that, starting from two identical and fused half barrels, few amino acid exchanges are sufficient to generate a highly stable and compact $(\beta\alpha)_8$-barrel protein with wild-type like structural properties.

Keywords: protein design; enzyme evolution; $(\beta\alpha)_8$-barrel; half barrel; stability screen

*Corresponding author

Introduction

It has been suggested that complex contemporary protein folds have evolved by the association and fusion of small polypeptide fragments.[1,2] The internal symmetry observed in a number of enzymes suggests that identical polypeptide segments first assembled to homo-oligomers, followed by the covalent linkage of the fragments as a consequence of gene duplication and fusion.[3–5] The $(\beta\alpha)_8$ (or

Abbreviations used: AdoMet, *S*-adenosyl-L-methionine; CAT, chloramphenicol acetyltransferase; MD, molecular dynamics; GdmCl, guanidinium chloride.

E-mail address of the corresponding author: Reinhard.Sterner@biologie.uni-regensburg.de

TIM) barrel is the most common enzyme fold, which currently contains about 400 identified members listed in the SCOP database†, and therefore is ideally suited for the study of enzyme evolution.[6–11] The canonical $(\beta\alpha)_8$-barrel consists of eight $(\beta\alpha)$ units, each of which comprises a β-strand and a subsequent α-helix which are connected by a βα-loop; the individual units are linked by αβ-loops.[12–14] The eight β-strands form a central parallel β-sheet (the barrel), which is surrounded by the eight α-helices. Within the central β-barrel, the residues are forming four layers, which are ordered perpendicular to the barrel axis and are numbered from the C- to the N-terminal barrel face. Each layer consists of residues from the odd (1, 3, 5, 7) or the even (2, 4, 6, 8) numbered strands, which generates a 4-fold symmetry. The active site of all known $(\beta\alpha)_8$-barrel enzymes is harboured by the C-terminal ends of the β-strands and by the βα-loops.

The apparent modularity of the $(\beta\alpha)_8$-barrel suggests that it has evolved by the duplication, fusion and recombination of fragments, which comprised either a single primordial $(\beta\alpha)$ element, or an even numbered multiple thereof. This hypothesis has been supported by protein fragmentation and folding studies.[11] For example, the observation that many $(\beta\alpha)_8$-barrels contain a conserved Gly-X-Asp motif in the loops α1β2, α3β4, α5β6, and α7β8 but not in the loops α2β3, α4β5, and α6β7 implies that $(\beta\alpha)_2$-units might constitute the smallest evolving entity.[10] Along these lines, the members of the *S*-adenosyl-L-methionine (AdoMet) radical protein family have three putative architectures, which comprise four, six and eight $(\beta\alpha)$-units.[15] The four or six N-terminal units contain the elements for radical generation, whereas the remainder of the protein is crucial for substrate binding. These findings suggest that AdoMet proteins might have evolved by the stepwise fusion of $(\beta\alpha)_2$ units, starting from a $(\beta\alpha)_4$-half barrel to a $(\beta\alpha)_6$-three-quarter to a $(\beta\alpha)_8$-full barrel structure. Along the same lines, fragmentation studies with the α-subunit of tryptophan synthase are compatible with a 6 (4+2) + 2 folding mechanism.[16,17] Moreover, the phosphoribosyl anthranilate isomerase (PRAI) from yeast forms a stable fragment consisting of the N-terminal $(\beta\alpha)_{1\text{-}6}$ units, which associates with the unstructured $(\beta\alpha)_{7–8}$ unit to a functional complex *in vitro*.[18] However, out of three fragment combinations of PRAI from *Escherichia coli* co-expressed *in vivo*, only the non-covalent complex between $(\beta\alpha)_{1\text{-}4}$ and $(\beta\alpha)_{5\text{-}8}$ yielded a functional enzyme.[19] This finding is in favour of a "4+4" folding and evolution mechanism, which has also been supported by studies on triosephosphate isomerase from chicken and rabbit muscle.[20,21]

Although the mechanistic details appear to be more complex than suggested by such a simple model,[22,23] the analysis of the enzymes *N*′-[(5′-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase (HisA) and imidazole glycerol phosphate synthase (HisF) has provided ample evidence for the crucial role of $(\beta\alpha)_4$-half barrels in the folding and evolution of $(\beta\alpha)_8$-barrel enzymes.[11] The amino acid sequences and X-ray structures of HisA and HisF from *Thermotoga maritima*, which catalyse two successive transformations of similar bisphosphorylated substrates within histidine biosynthesis, show an internal 2-fold symmetry (Figure 1(a)).[24,25] The pairs of N-terminal halves (HisA-N and HisF-N) that consist of the first four units ($(\beta\alpha)_{1–4}$) and the pairs of the C-terminal halves (HisA-C and HisF-C) that consist of the last four units ($(\beta\alpha)_{5–8}$) show sequence identities between 16 and 26% and root-mean-square (rms) deviations of their main-chain non-hydrogen bonds between only 1.4 Å and 2.1 Å, respectively. When produced separately, HisF-N and HisF-C form folded and mainly homo-dimeric proteins, which upon co-expression *in vivo* or joint refolding *in vitro* assemble to the non-covalent hetero-dimeric HisF–NC complex that displays wild-type catalytic activity.[9] Remarkably, four of the five $(\beta\alpha)$ units of several members of the flavodoxin-like fold family show striking similarities to the four $(\beta\alpha)$ units of HisF-N and HisF-C.[26] Moreover, the catalytically crucial N-terminal $(\beta\alpha)_4$-half barrels of eukaryotic and prokaryotic phosphoinositide-specific phospholipases C (PI-PLC) superimpose with rms deviations of 1.85 Å for 104 equivalent C^{α}-atoms.[27] Taken together, these findings suggest that $(\beta\alpha)_4$-half barrels and presumably also $(\beta\alpha)_2$-quarter barrels are independently evolving units. This conclusion would be further strengthened if stable and functional $(\beta\alpha)_8$-barrels could be generated by the fusion and recombination of existing half barrels.

Along these lines, we recently generated and characterised the stable chimeric $(\beta\alpha)_8$-barrels HisAF and HisFA, in which the N-terminal half of HisA was linked to the C-terminal half of HisF, and *vice versa*.[28] Moreover, we duplicated and fused in tandem the *his*F-C gene, yielding the HisF-CC protein (Figure 1(b)). Using rational protein design, the interface between the two identical halves in HisF-CC was stabilised by reconstituting a salt bridge cluster, which is conserved in all known HisF proteins. Although the resulting construct HisF-C*C containing the mutations A124R and A220K (Figure 1(b)) is a predominantly monomeric and a more compact protein than HisF-CC, it was produced in *E. coli* as inclusion bodies, had a slight tendency to aggregate, and differed in its spectroscopic properties from wild-type HisF.[28] To further increase the stability and solubility of HisF-C*C, we have now applied random mutagenesis followed by selection *in vivo* using a chloramphenicol acetyltransferase (CAT) reporter system.[29] The combination of several beneficial mutations identified by this approach allowed us to generate the HisF-C**C protein (Figure 1(c)), which can be produced in soluble form in *E. coli*, forms a homogeneous and compact monomer in gel filtration experiments, shows an identical fluorescence spectrum as wild-

† scop.mrc-lmb.cam.ac.uk/scop/