FOR THE RECORD

On the evolution of alternate core packing in eightfold β/α -barrels

ANDREW R.C. RAINE,¹ NIGEL S. SCRUTTON,¹ AND F. SCOTT MATHEWS²

¹ Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom

² Department of Biochemistry, Washington University School of Medicine, St. Louis, Missouri 63110

(RECEIVED July 14, 1994; ACCEPTED August 8, 1994)

Abstract: Two sequence-related subfamilies of flavin-binding β/α -barrels have been identified (the type I and type II proteins) that differ in the nature of residue packing in the core of the barrel domain. Similar observed differences in the packing of internal amino acid side chains in β/α -barrels have previously been used to argue that these domains have evolved convergently toward a stable structural framework. Using structural alignments of flavin-binding barrel proteins, we demonstrate that simple genetic alterations may be responsible for switching the nature of side-chain packing observed in β/α -barrels. The implication is that the 2 structural classes of β/α -barrel cores can arise divergently from an ancestral barrel framework and that convergent evolution to a stable fold need not be invoked to account for the emergence of 2 classes of β/α -barrel core.

Keywords: β/α -barrel; core structure; evolution; flavin

The evolution of β/α -barrels has been vigorously debated in recent years (for a review, see Farber, 1993) and an understanding of their pedigree is central to our knowledge of the mechanisms by which enzyme molecules containing β/α -barrel domains have evolved. The 8-fold β/α -fold was first observed in triose phosphate isomerase (Phillips et al., 1978), and it is presently found in about 10% of all structurally determined enzyme molecules. Despite this wealth of information, analyses of sequence and structure have fueled the debate as to whether these domains have arisen convergently to a stable fold or divergently from a common ancestor. Poor sequence relationships between different β/α -barrel proteins have been used to argue in favor of their convergent evolution toward a stable structural framework. On the other hand, sequence analyses have also suggested that small subregions of some barrel domains may be related (Wilmanns et al., 1991), but the sequence similarity is restricted to only a small C-terminal region of the barrel domain.

Farber and Petsko (1990) have used a structural approach to illustrate possible divergent pathways of evolution and they have classified barrel domains into 4 groups on the grounds of geometrical relationships. In contrast, Lesk and colleagues (1989), in a study of residue packing in selected barrel cores and in the absence of extensive sequence homologies, favor convergent pathways of barrel evolution. Clearly, the evolutionary origins of barrel frameworks are ill-defined, and the evidence in favor of divergent or convergent evolution is not substantive.

Studies of the β/α -barrel flavin oxidase/dehydrogenase family are beginning to provide new insights into the enigma of barrel descent (Scrutton, 1994). The flavin oxidase/dehydrogenase family comprises a series of redox proteins that catalyze a wide range of chemical reactions. Members range from simple redox proteins (containing a single flavin cofactor) to complex multidomain proteins (containing flavin and other cofactors). The recent acquisition of sequence information for members of the β/α -barrel flavin oxidase/dehydrogenase family indicates that these proteins can be divided into 2 sequence-related subfamilies, the type I and type II proteins (Scrutton, 1994). The type I and type II barrel proteins have provided the first compelling evidence in favor of barrel domains sharing a common ancestry. Although members of each subfamily exhibit substantial sequence similarity throughout the entire barrel domain, no discernible sequence similarities exist between the 2 subfamilies of the flavin oxidase/dehydrogenase family or with other β/α -barrel structures (Scrutton, 1994). The question remains, therefore, as to whether the 2 subfamilies are related. Extensive sequence scrambling may be responsible for the loss of any significant sequence homology between the 2 subfamilies. Alternatively, the 2 subfamilies may have arisen independently of each other and converged on a common structural framework.

Lesk et al. (1989) identified 2 modes of residue packing within the core region of β/α -barrel domains. They made a study of 3 seemingly unrelated β/α -barrel proteins, viz., triose phosphate isomerase, ribulose biphosphate carboxylase, and glycolate oxidase, and identified 3 planes of amino acid side chains in the barrel core with 4 amino acids present in each plane. This internal packaging of side chains gives rise to 2 types of barrel core. In the first, strands 1, 3, 5, and 7 provide residues for the

Reprint requests to: Nigel S. Scrutton, Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK; e-mail: ns101@mole.bio.cam.ac.uk.

middle plane and strands 2, 4, 6, and 8 provide residues for the outer planes. In the second case, the order is reversed, and strands 1, 3, 5, and 7 make up the outer 2 planes and strands 2, 4, 6, and 8 the middle plane. We refer to the first mode of packing as a $(1:2)_4$ structure and the second mode as a $(2:1)_4$ structure.

For the flavin barrel proteins, both types of core structure can be identified. Structural information is available for 3 members of the β/α -barrel flavin oxidase/dehydrogenase family from a total of 10 members. As originally reported by Lesk et al. (1989), glycolate oxidase conforms to the $(2:1)_4$ structure. Flavocytochrome b_2 is related to glycolate oxidase (both are type II proteins; Lindqvist et al., 1991) and we show here (Fig. 1 and Kinemage 1) that flavocytochrome b_2 also possesses a $(2:1)_4$ core structure. The only type I member of the flavin barrel family for which there is high-resolution structural information available is trimethylamine dehydrogenase (Lim et al., 1986). Inspection of the core region of this protein reveals that the 3 planes of core residues can be readily recognized, but in this case the core packing conforms to the $(1:2)_4$ structure (Fig. 1 and Kinemage 2). On the basis of sequence homology, one might expect the remaining type I proteins to be constructed on a $(1:2)_4$ core structure and type II proteins centered around a $(2:1)_4$ core structure. Homology modeling of the cores of other members of the family suggests this to

be so (A.R.C. Raine & N.S. Scrutton, unpubl. results). This difference between the 2 families represents an important structural variation between the type I and type II proteins.

Lesk and colleagues (1989) argue that β/α -barrel domains that display different modes of core packing may have arisen by convergent evolution toward a common structural fold. Their reasoning is founded on the difficulties one encounters in identifying a simple genetic change that can switch a $(1:2)_4$ to a $(2:1)_4$ core structure or vice versa. They state that a circular permutation event is the only reasonable genetic change that could give rise to a switch in core structure, but this strategy is perhaps too elaborate to account for the structural change. However, laboratory-based circular permutations have given rise to stable variants of the β/α -barrel of yeast phosphoribosyl anthranilate isomerase (Luger et al., 1989), and, in the absence of a structural solution for these variants, one must surmise that the core packing has changed as a result of the engineering exercise.

We have observed that, for the 2 types of flavin-binding barrel domain, the positions of the N-terminal barrel cap and the flavin cofactor relative to the 8 strands of the barrel are equivalent in the type I and type II proteins (see below). This observation is difficult to reconcile with a circular permutation event having given rise to a switch in core structure.

Fig. 1. Stereo projections of the 3 planes making up the barrel cores of flavocytochrome b_2 , glycolate oxidase, and trimethylamine dehydrogenase. Coordinates for flavocytochrome b_2 (Xia et al., 1987; Brookhaven code 1FCB) and glycolate oxidase (Lindqvist, 1989; Brookhaven code 1GOX) were taken from the Brookhaven database. Coordinates for trimethylamine dehydrogenase are the refined coordinates for trimethylamine dehydrogenase (F.S. Mathews & S.A. White, unpubl. work; Brookhaven code 1TMD) that were obtained using the recently determined sequence of trimethylamine dehydrogenase (Boyd et al., 1992). A general description of the structure of trimethylamine dehydrogenase (lacking sequence information) is given by Lim et al. (1986). **A:** Flavocytochrome b_2 , plane 1 (Y193, W250, V347, F407); plane 2 (M226, F278, V369, G430); plane 3 (S195, Q252, K349, D409). **B:** Glycolate oxidase, plane 1 (M74, F125, L228, F283); plane 2 (T104, A153, I250, F306); plane 3 (A76, Q127, K230, D285). **C:** Trimethylamine dehydrogenase, plane 1 (A55, I165, M253, I 317); plane 2 (Y25, G101, A220, L295); plane 3 (N57, Y167, D255, G319). The first, second, and third planes are colored red, magenta, and yellow, respectively.

A structural analysis of the flavin-binding barrel domains of glycolate oxidase, flavocytochrome b_2 , and trimethylamine dehydrogenase suggests that a simpler genetic change may account for a switch in core packing and, by implication, that the 2 types of core *can* arise divergently from a common ancestral protein. This provides an attractive alternative to the "convergent pathway" hypothesis developed by Lesk and colleagues. Although trimethylamine dehydrogenase displays no significant sequence similarity with the 2 type II proteins, we have been able to perform a structural alignment of trimethylamine dehydrogenase with glycolate oxidase and flavocytochrome b_2 (Fig. 2 and Kinemage 3). Common to the 3 structurally determined flavin barrel domains is the presence of a β -hairpin positioned over the N-terminal end of the β/α -barrel domain. Also, at the C-terminal end of each domain, flavin mononucleotide is found bound to the loops that interconnect the C-terminal ends of the internal β -strands and the N-terminal ends of the surrounding α -helices. The orientation of the bound flavin with respect to the β/α -domain is identical in each of the type I and type II

proteins, reinforcing the structural similarity seen across the subfamily divide. These 2 conserved structural features (the N-terminal cap and the position of the flavin) allow barrel alignments to be conducted using the cap and flavin as internal reference positions. From Figure 2 and Kinemage 3, it is plainly evident that the degree of structural conservation within and between the 2 subfamilies is very high.

Within the framework of the structural alignment, it is possible to identify those residues (Fig. 1) that comprise the 3 planes in each of the barrel structures. For flavocytochrome b_2 and glycolate oxidase, the positions of the backbone atoms of those residues making up the 3 planes are spatially equivalent. For trimethylamine dehydrogenase, however, the positions of the 3 planes are translated along the barrel axis by 1 residue toward the N-termini of the β -strands compared with the planes identified for the type II proteins. Consequently, plane 1 of the 2 class II proteins is equivalent to plane 2 of trimethylamine dehydrogenase, and plane 2 of the type II proteins is equivalent to plane 3 of trimethylamine dehydrogenase, etc. In the class II

Fig. 2. A: Structural alignment of the barrel domains of trimethylamine dehydrogenase, flavocytochrome b_2 , and glycolate oxidase. Coordinates were obtained as described in the legend to Figure 1. Unweighted mean structural alignments of the C α atoms in the β -strands and N-terminal cap were performed using the program "uwmm" (M.J. Hartshorn & L.S.D. Caves, unpubl.). The graphics illustration was generated using the MOLSCRIPT software (Kraulis, 1991). Only the protein backbone is illustrated (carbonyl oxygens are not shown). Trimethylamine dehydrogenase is depicted in red, glycolate oxidase in yellow, and flavocytochrome b_2 in blue. The RMS deviations from the average structure are: flavocytochrome b_2 , 0.69 Å; glycolate oxidase, 0.7 Å; trimethylamine dehydrogenase, 1.25 Å. **B:** The 3 core planes of flavocytochrome b_2 . Red, first plane; magenta, second plane; yellow, third plane. **C:** The 3 core planes of glycolate oxidase. Red, first plane; magenta, second plane; yellow, third plane. **D:** The 3 core planes of trimethylamine dehydrogenase. Red, first plane; magenta, second plane; yellow, third plane.

proteins, plane 0 is prevented from forming by the 2 aromatic residues of plane 1 (Fig. 1). In trimethylamine dehydrogenase, plane 4 is not formed; the backbone of the first β -strand projects slightly into the barrel axis such that the side chain of residue Pro-28 (C-terminal to plane 4) squeezes the seventh β -strand out of the barrel core to form the early part of the loop that interconnects strand 7 with the surrounding helix.

In looking at the structural relatedness of barrel domain cores between different proteins, the strict adherence to the "three-plane" description can be misleading. The boundaries of the plane definitions need to be extended to include those residues at the ends of the barrel that do not form planes in the strict sense of the word but nevertheless need to be considered when aligning barrel protein structures. This is especially so when barrel proteins that contain no internal positional reference (e.g., N-terminal barrel cap and flavin) are aligned.

The observed difference in the location of the barrel core planes between the 2 classes of flavin-binding barrels suggests a divergent mechanism for switching between the 2 types of barrel core. For example, for the class II proteins ((2:1)₄ structure), a switch to the (1:2)₄ structure could arise by a "loosening" of the packing in plane 3 by single amino acid substitutions and a "tightening" of the packing in plane 0. Consequently, in a (2:1)₄ to (1:2)₄ switch, the old plane 0 would become the new plane 1 and the old plane 3 would become the new plane 4, and the planes would be translated along the barrel axis by 1 residue. The mutational changes involved in switching between the 2 types of barrel core would be minimal single-residue changes, and the mechanism of switching is therefore highly plausible.

In summary, the structural alignment has revealed a translational slip in the positions of the 3 planes that constitute the hydrophobic core in the 2 sequence-related classes of flavin-binding barrel proteins. This structural plasticity in the core of the domain suggests that switching between the 2 types of core structure can be accomplished via single mutational events along a divergent evolutionary pathway. This type of evolutionary

mechanism provides an attractive and perhaps more plausible alternative to the convergent pathway hypothesis previously used to account for the existence of both types of core structure.

References

- Boyd G, Mathews FS, Packman LC, Scrutton NS. 1992. Trimethylamine dehydrogenase of bacterium W3A1: Molecular cloning, sequence determination and overexpression of the gene. *FEBS Lett* 308:271-276.
- Farber GK. 1993. An α/β barrel full of evolutionary trouble. *Curr Opin Struct Biol* 3:409-412.
- Farber GK, Petsko GA. 1990. The evolution of α/β barrel enzymes. *Trends Biochem Sci* 15:228-234.
- Kraulis PJ. 1991. MOLSCRIPT - A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946-950.
- Lesk AM, Branden CI, Chothia C. 1989. Structural principles of α/β barrel proteins: The packing of the interior of the sheet. *Proteins Struct Funct Genet* 5:139-148.
- Lim LW, Shamala N, Mathews FS, Steenkamp DJ, Hamlin R, Xuong NH. 1986. Three-dimensional structure of the iron-sulfur flavoprotein trimethylamine dehydrogenase at 2.4 Å resolution. *J Biol Chem* 261:15140-15146.
- Lindqvist Y. 1989. Refined structure of glycolate oxidase at 2 Å resolution. *J Mol Biol* 209:151-166.
- Lindqvist Y, Branden CI, Mathews FS, Lederer F. 1991. Spinach glycolate oxidase and yeast flavocytochrome *b*₂ are structurally homologous and evolutionarily related enzymes with distinctly different function and flavin mononucleotide binding. *J Biol Chem* 266:3198-3207.
- Luger K, Hommel U, Herold M, Hofsteenge J, Kirschner K. 1989. Correct folding of circularly permuted variants of a $\beta\alpha$ barrel enzyme in vivo. *Science* 243:206-210.
- Phillips DC, Sternberg MJE, Thornton JM, Wilson IA. 1978. Structure of triose phosphate isomerase and its comparison with lactate dehydrogenase. *J Mol Biol* 119:329-351.
- Scrutton NS. 1994. α/β Barrel evolution and the modular assembly of enzymes: Emerging trends in the flavin oxidase/dehydrogenase family. *BioEssays* 16:115-122.
- Wilmanns M, Hyde CC, Davies DR, Kirschner K, Jansonius JN. 1991. Structural conservation in parallel β/α barrel enzymes that catalyse three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry* 30:9161-9169.
- Xia ZX, Shamala N, Bethge PH, Lim LW, Bellamy HD, Xuong NH, Lederer F, Mathews FS. 1987. Three-dimensional structure of flavocytochrome *b*₂ from baker's yeast at 3.0 Å resolution. *Proc Natl Acad Sci USA* 84:2629-2633.