Evolutionary markers in the \((\beta/\alpha)_8\)-barrel fold
M Cristina Vega, Esben Lorentzen, Anni Linden and Matthias Wilmanns*

Enzymes with the \((\beta/\alpha)_8\)-barrel fold are involved in the catalysis of a wide variety of biochemical reactions. The active sites of these enzymes are located on the C-terminal face of the central \(\beta\)-barrel. Conserved amino acid sequence, as well as secondary, tertiary and quaternary structure patterns are providing a rich body of data to support the premise of a common ancestry of many members of the \((\beta/\alpha)_8\)-barrel fold family of enzymes. Recent data indicate that there is at least one example of a bienzyme that functions as an ammonia channel, adding a new level of functional diversity to the \((\beta/\alpha)_8\)-barrel fold. These proteins have become ideal tools that can be used in conjunction with directed evolution techniques to engineer novel catalytic activities.

Addresses
EMBL-Hamburg c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany
*e-mail: wilmanns@embl-hamburg.de

Abbreviations
HisA phosphoribosylformimino-5-aminolimidazole carboxamide ribonucleotide isomerase
HisF imidazole glycerol phosphate synthase
HisH glutaminase of the imidazole glycerol phosphate synthase
KGPDC 3-keto-L-gulonate 6-phosphate decarboxylase
OMPDC orotidine 5'-monophosphate decarboxylase
PDB Protein Data Bank
SCOP Structural Classification of Proteins (database)
TrpA \(\alpha\)-subunit of tryptophan synthase
TrpF N-(5'-phosphoribosyl)anthranilate isomerase

Introduction
Approximately 10% of all enzymes with known molecular structure are folded as \((\beta/\alpha)_8\)-barrels, exceeding any other known fold in terms of overall number and functional diversity. A canonical \((\beta/\alpha)_8\)-barrel consists of an inner ring of eight parallel \(\beta\)-strands that is wrapped by an outer wheel comprising typically eight \(\alpha\)-helices. While the \(\beta\)-strands of the inner barrel are connected via a parallel \(\beta\)-sheet H-bond pattern, the interactions between the outer helices are less regular. These secondary structural elements generate an alternating \((\beta\alpha)\) pattern that facilitates reliable predictions of their fold from corresponding sequence patterns [1]. The Structural Classification of Proteins (SCOP) database (http://scop.mrc-lmb.cam.ac.uk/scop/, version 1.63, May 2003) lists 26 \((\beta/\alpha)_8\)-barrel fold families. In the Class, Architecture, Topology and Homologous superfamily (CATH) database (http://www.biochem.ucl.ac.uk/bsm/cath/, version 2.4, January 2002), the class-3 \((\alpha/\beta)\) superfamily contains 28 \((\beta/\alpha)_8\)-barrel families. Some proteins display deviations from the classical \((\beta/\alpha)_8\)-barrel fold, including variations in the number of \(\beta\)-strands, their orientation (typically all parallel), local structural distortions and lack of barrel closure [2**.3]. A representative set of \((\beta/\alpha)_8\)-barrel topology variants is shown in Figure 1.

Most \((\beta/\alpha)_8\)-barrel proteins function as enzymes, with a few exceptions, such as narbonin and concanavalin B [4]. These enzymes catalyze a diverse array of biochemical reactions [2**,5**]. Members of the \((\beta/\alpha)_8\)-barrel fold are subsumed within five of the six general classes of catalytic activities according to the Enzyme Commission classification scheme, excluding only the ligases. Because of their repetitive amino acid sequence and structural features, it has been difficult to track a possible common evolutionary ancestry on the basis of sequence and structure comparative analyses only, triggering more than a decade-long debate on the convergent versus divergent general evolutionary roadway of these proteins [5**,6,7]. A general argument to support the premise of a broad, common ancestry may originate from the observation that the active sites of all \((\beta/\alpha)_8\)-barrels are located on the C-terminal face of the central \(\beta\)-barrel, as there is no inherently obvious structural basis for preference of the C-terminal face over the N-terminal face.

It has been noted that a large proportion of the enzymatic reactions of some metabolic pathways, such as glycolysis or tryptophan biosynthesis, are catalyzed by \((\beta/\alpha)_8\)-barrel proteins [8,9]. These observations have enabled the testing and validation of previous evolutionary models, including the ‘retrograde pathway hypothesis’ by Horowitz (1945), the ‘patchwork evolution hypothesis’ by Jensen (1976) and modifications thereof that have been reviewed recently [10,11]. These pathway-oriented considerations have been complemented by a series of investigations in an effort to derive common ancestry from mechanistic, biochemical associations [5**]. In their recent review, Gerlt and Rauschel have summarized data providing evidence for the evolution of ‘new’ enzymes based on a common chemical mechanism, a common ligand specificity as well as a common active site architecture. Recent sequence, structural and biochemical analyses of two \((\beta/\alpha)_8\)-barrel enzymes of the histidine biosynthesis
pathway have led to the prediction that they may have evolved from smaller fragments such as (β/α)_4-barrel units [12,13]. In the following contribution, we provide an overview of recent data revealing an association between (β/α)_8-barrel enzymes with respect to their amino acid sequences, as well as secondary, tertiary and quaternary structures.

**Sequence relations in (β/α)_8-barrels**

Comparative analyses of the amino acid sequences of known (β/α)_8-barrel folds do not reveal significant sequence similarity in general, precluding detection of an association by this criterion alone. However, there are several exceptions, mostly from small clusters of homologous (β/α)_8-barrels that are associated with the catalysis of specific reaction mechanisms or are known to be members of the same or related pathways [14–18,19*]. Knowledge of their three-dimensional structures (see below), however, not only permits unbiased, structure-based sequence alignments, but also allows detection of conserved structural motifs not found in sequence comparisons.

**Secondary structure relations in (β/α)_8-barrels**

The eightfold repeat of (βα) units in the canonical (β/α)_8-barrel fold generates a highly symmetrical arrangement of secondary structural elements. The structure of the inner parallel β-strand barrel is more constrained, with a few notable exceptions [2**], than the arrangement of the outer α-helical ring. The architecture of the inner β-barrel can be described geometrically as hyperboloid that, because of an approximate 36-degrees tilt with respect to...
the barrel major axis, generates a shear number $S = 8$ [20] (Figure 2a). Furthermore, the eightfold $\beta$-barrel forms an arrangement of side chains in three to four layers within the core of the central $\beta$-barrel (Figure 2b). Each layer is perpendicular to the principal barrel axis and is established by an alternating pattern of residues from the four odd-numbered (1,3,5,7) or even-numbered (2,4,6,8) $\beta$-strands, imposing a fourfold symmetry pattern. Although rare in nature, circular permutated ($\beta/\alpha$)$_8$-barrel versions have been detected [21,22] and could be engineered [23]. The repetitive and symmetrical arrangement of secondary structural elements has facilitated several predictions of the ($\beta/\alpha$)$_8$-barrel fold from amino acid sequences with no known structure [1,24–26].
Common motifs in the 3D-structures of \( (\beta/\alpha)_8 \)-barrels

The general consensus is that the detection of a common \( (\beta/\alpha)_8 \)-barrel fold alone is insufficient evidence to prove a possible evolutionary association. Although the apparent restriction of the active sites to the C-terminal face of the inner parallel \( \beta \)-barrel does serve as an appealing argument for the general evolutionary ancestry of many, if not all, \( (\beta/\alpha)_8 \)-barrel enzymes, three-dimensional structural analyses of these enzymes are providing an abundance of common structural motifs, providing an impetus toward the detection of further evolutionary relations.

Common phosphate binding site

Approximately two-thirds of the established \( (\beta/\alpha)_8 \)-barrel enzyme families are utilizing substrates or cofactors that contain at least one phosphate group. Initially, it was recognized that three \( (\beta/\alpha)_8 \)-barrel enzymes from the tryptophan biosynthesis pathway use the same motif, comprising the loops connecting \( \beta_7-\alpha_7 \) and \( \beta_8-\alpha_8 \), to bind the phosphate groups of the respective substrates [9] (Figure 3a). The latter loop contains a short additional helix (\( \alpha_{8'} \)) whose N-terminal macrodipole moment provides additional binding potential. The phosphate binding site is considered to be the most conserved structural motif among \( (\beta/\alpha)_8 \)-barrel enzymes [2**].

Metal-assisted catalysis

Approximately one-half of the characterized \( (\beta/\alpha)_8 \)-barrel enzymes require the presence of metal ions for catalysis [2**]. Although the positions of catalysis-assisting metals are highly constrained in some enzymes from the enolase family [27,28], overall, there appears to be more structural versatility in terms of positions, types and number of metal centres [2**] than, for instance, for the conserved phosphate binding site (Figure 3b). A striking, recent example illustrates how modest changes in the active site architecture in a pair of closely related \( (\beta/\alpha)_8 \)-barrel enzymes, 3-keto-L-gulonate 6-phosphate decarboxylase (KGPDC) and orotidine 5′-monophosphate decarboxylase (OMPDC), lead to unrelated mechanisms of catalysis [19*]. These changes include

-the tryptophan biosynthesis pathway, triose phosphate isomerase (light gray, PDB code: 1HG3) and fructose 1,6-bisphosphate aldolase (gray, PDB code: 1OK4) from the glycolysis pathway. In each of the three structures there is a substrate analog bound to the active site. They are colored in green, cyan and red, respectively. The phosphate groups of these compounds bind into the same binding pocket that is formed by the loops \( \beta_7-\alpha_7 \) and \( \beta_8-\alpha_8 \); (b) Superposition of 2-phosphoglycerate dehydratase (light gray, PDB code: 1E9I) and glucarate dehydratase (dark gray, PDB code: 1EC8). These two structures do not exhibit significant sequence similarities. Three structurally conserved residues and the active site magnesium ions are shown in red and blue, respectively. (c) Superposition of an eukaryotic (dark gray) and an archaeal class I aldolase (PDB codes: 1J4E and 1OK4, respectively). The loop connecting \( \beta_3 \) and \( \alpha_3 \) is involved in their oligomerization into tetramers and pentamers (forming a decamer with a second pentamer).
the conversion of a negatively charged residue into a positively charged one (KGPD, E33; OMPDC, K33), thereby replacing the function of a magnesium ion in the active site of KGPD with the side-chain amino group of K33 in OMPDC. Another interesting example is found in family-13 a-amylases with the (β/α)8-barrel fold. They contain a conserved calcium ion site in the close vicinity of the active site, probably for the purpose of stabilizing the active site conformation [29]. Interestingly, a-amylases from hyperthermophilic bacteria and archaea have developed distinct bi/tri-metal centres that superimpose, but are unrelated with respect to metal preference [30]. Comparative analyses of these enzymes indicate that structural adaptation to specific environmental conditions may have evolved independently of a (β/α)8-barrel progenitor.

Conserved active sites
Although (β/α)8-barrel enzymes are involved in a diverse array of catalyzed reaction mechanisms, there are groups among these enzymes that share remarkable active site similarities. Aldolases that are folded as (β/α)8-barrels exhibit a range of different substrate specificities and display, in part, unrelated catalytic mechanisms. Class I aldolases from eucary and archaea, specifically catalyzing the cleavage of fructose 1,6-bisphosphate, have diverged considerably in sequence, precluding the detection of significant sequence similarity [31]. However, superposition of the structures of an eukaryotic and archaeal fructose 1,6-bisphosphate aldolase has revealed that they share at least six active site residues located at structurally invariant positions, thereby indicating common ancestry [32*]. One of the most widely studied groups of (β/α)8-barrel enzymes is the enolase family, which, according to present data, is associated with 11 known different catalytic activities and numerous others yet to be defined [5**]. While the canonical enolase (β/α)8-barrel comprises a conserved pair of acid/base catalysts plus an active site magnesium ion, several modified active site topologies have evolved to contend with specific mechanistic requirements of other members of the same family [33]. Other examples of enzymes with conserved active sites have been well illustrated in a recent review [5**].

Repeated motifs
Structural comparison of two (β/α)8-barrel enzymes from the histidine biosynthesis (HisA, HisF) pathway has revealed a common twofold repeat structure, each comprising a (β2α4) unit [12] that show native-like biophysical properties if purified separately [13]. The twofold repetitive structure is reflected by structure-based sequence similarities, two phosphate binding motifs (loops 3/4 and loops 7/8), a symmetric arrangement of catalytic residues on loops 1 and 5, and a twofold repetitive loop structure on the C-terminal face of each of the two (β/α)8-barrels (Figure 2c). Interestingly, a twofold repetitive arrangement of catalytic acid/base residues has also been found in class I aldolases [32*]. A recent study was carried out on two differentially regulated isoforms of the DHAP synthase (β/α)8-barrel, yielding catalytically active chimeras [34**].

Related quaternary structural arrangements
Many of the known (β/α)8-barrel structures are oligomeric. Recently, it has been shown that the structural dimers of two decarboxylases (OMPDC, KGPD) of the ribulose phosphate binding family (SCOP nomenclature) can be superimposed, demonstrating that a common quaternary arrangement may serve as further indication for common ancestry [19*]. Interestingly, even OMPDC and other dimeric (β/α)8-barrels, such as triose phosphate isomerase, display related quaternary arrangements that form an active site interface. In other families, such as fructose 1,6-bisphosphate aldolases, different decameric and tetrameric arrangements have been found. Although they lead to unrelated quaternary structures, it has been noticed that their assembly is generated by a topographically identical loop [32*] (Figure 3c). In another group of oligomeric (β/α)8-barrels, enzymes first observed in phosphoenolpyruvate mutase [35], swapping of the C-terminal barrel helix α20H appears to be critical for oligomeric assembly. The latest addition to this class of enzymes has been the structure of ketopantoate hydroxymethyltransferase from the pantotheine biosynthesis pathway [28].

(β/α)8-Barrels as part of molecular machines
The structure of the tryptophan synthase ββσ complex, the terminal enzyme complex of tryptophan biosynthesis,so provides a potential example of how a (β/α)8-barrel enzyme (TrpA) may be involved in a complex reaction requiring compound tunneling [36,37]. While the active site loop structure of the (β/α)8-barrel TrpA subunit is critical for submitting its product indole into the funnel linking the TrpA and TrpB active sites, TrpA does not provide the structural framework for tunneling. Recent molecular structures of the glutaminase-imidazole glycerol phosphate synthase (HisH-HisF) complex, however, indicated, for the first time, how a (β/α)8-barrel (HisF) may directly provide the scaffold for compound (ammonia) tunneling [38*,39*,40]. In these complexes, the glutaminase subunit (HisH) docks onto the N-terminal face of the HisF (β/α)8-barrel, thereby suggesting a model for ammonia tunneling along the central β-barrel axis of the HisF (β/α)8-barrel towards the synthase active site on the C-terminal β-barrel face of HisF. Recent structural data indicate that at least one conserved residue on the N-terminal face of the HisF (β/α)8-barrel (Q123) is essential for glutaminase catalysis [38*], MC Vega and M Wilmanns, unpublished data), supporting previous mutagenesis data [41]. The presence of catalytic residues and a cluster of charged residues providing the gate for the suggested ammonia tunnel may explain previous observations of an unusually high number of conserved
residues on the N-terminal face of the HisF (β/α)8-barrel [12]. The FMN binding domain of glutamate synthase, which uses ammonia that is generated by glutamine hydrolysis as well, is also folded as (β/α)8-barrel. In contrast to the glutaminase-imidazole glycerol phosphate synthase complex, the ammonia channel in glutamate synthase seems to be formed by the interface from different domains not directly involved in the (β/α)8-barrel core [42].

Knowledge-based engineering of novel catalytic activities in (β/α)8-barrels

The present sequence, structural and biochemical data provide an increasingly emerging body of unambiguous evidence for a common evolutionary ancestry for enzymes that are folded as (β/α)8-barrels. Several possible evolutionary scenarios have been widely discussed [5**,43], although available data on their phylogenetic association with implications for family classifications are not yet fully consistent [2**,18] and, therefore, must be interpreted with caution [5**]. Recent support for a common progenitor for two (β/α)8-barrel enzymes comes from a genetic study in Mycobacterium tuberculosis and Streptomyces coelicolor, providing evidence that in some microbial organisms, two similar isomerase reactions in histidine biosynthesis (HisA) and tryptophan biosynthesis (TrpF) [44] may be catalyzed by a bifunctional common ancestor-like enzyme [45]. In a previous directed evolution experiment in Escherichia coli, it was shown that one amino acid replacement is sufficient to swap the substrate specificities of these two enzymes [46]. Directed evolution techniques have also been applied to (β/α)8-barrel members of the enolase family, generating enzymes with indiscriminate substrate specificities but requiring different reaction mechanisms [47], and similar experiments for future studies have been proposed [48]. Using the same technique, the enantioselectivity of the reaction catalyzed by N-acetylneuraminic acid aldolase from E. coli could be converted, demonstrating the potential of the (β/α)8-barrel fold to tinker specific chemical reactions [49].

Conclusions

Proteins with the (β/α)8-barrel architecture present the largest fold family of enzymes with known molecular structure. There is an overwhelming and increasingly emerging body of evidence that most of them, if not all, share one common evolutionary ancestor. Proteins with the (β/α)8-barrel fold provide an ideal framework to study almost any conceivable biochemical reaction, including reactions that are coupled with other enzymatic activities. Hence, they represent excellent tools that can be used in conjunction with directed evolution techniques to explore novel catalytic activities.

Update

Recently, the yeast cytosine deaminase crystal structure has been reported [50]. Unlike the bacterial enzyme with (β/α)8-barrel topology, the yeast cytosine deaminase shows a mixed (α/β) topology, suggesting independent evolution. In addition, the structure of ketopantoate hydroxymethyl transferase from E. coli in a decameric arrangement of (β/α)8-barrels has been published [51].

Acknowledgements

We thank Areti Malapetsas for editing the manuscript. MW acknowledges support of the Deutsche Forschungsgemeinschaft in relation to the topics discussed in this review (Wi 1058/5-3).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

☆ of special interest
** of outstanding interest


The sequence, structural and functional diversity of (β/α)8-barrels is thoroughly reviewed, mostly from a bioinformatics perspective. The review contains excellent illustrations on various conserved (β/α)8-barrel motifs.


5. Gert JA, Rauschel RM: Evolution of function in (β/α)8-barrel enzymes. Curr Opin Chem Biol 2002, 7:252-264. Several scenarios for the evolution of (β/α)8-barrels are reviewed mostly from a biochemical viewpoint, particularly considering preservation of reactions mechanisms, ligand (substrate) specificity and active site architecture. Data from the enolase, amidohydrolase, OMP deaminase and aldolase superfamilies are summarized and discussed. The review contains numerous key references.


Model systems


20. The structures of KGPD and OMPDC are comparing, revealed close relations of two (beta/alpha)-barrels that catalyze unrelated reactions, thus presenting a case study for ‘opportunistic’ divergent evolution.


33. Henn-Sax M, Thoma R, Schmidt S, Henning M, Kirschner K, Sterner R: Two (beta/alpha)-barrel enzymes of histidine and...
tryptophan biosynthesis have similar reaction mechanisms and common strategies for protecting their labile substrates. Biochemistry 2002, 41:12032-12042.


Orthologous genes from M. tuberculosis and S. coelicolor can complement E. coli auxotrophs for HisA and TrpF, thus indicating that their gene products are bifunctional, and suggesting the presence of an ancient-like HisA/TrpF isomerase in M. tuberculosis and S. coelicolor.


Using an α-succinylbenzoate synthase auxotroph E. coli strain, directed evolution experiments were carried out using the genes of the related L-Ala-D/-Glu epimerase and muconate lactonizing enzyme II, resulting in (β/α)8-barrel enzyme versions with relaxed substrate specificities.


