The TIM barrel fold

Deepesh Nagarajan* and Neha Nanajkar

1Department of Biochemistry, Indian Institute of Science, Bangalore - 560012, India
*corresponding author: 1337deepesh@gmail.com

ABSTRACT

Proteins are biological polymers composed of linear chains of 20 different amino acids. The sequence of amino acids for every protein is unique, and guides its folding into intricate 3-dimensional shapes, known as protein folds. The TIM barrel is one such fold, and is characterized by an interior 8-stranded β-barrel, surrounded and enclosed by 8 α-helices. TIM barrels are named after triose phosphate isomerase (TIM), an enzyme first structurally characterized in 1975, which lends its name to the fold. TIM barrels are prevalent in all forms of life, and across diverse metabolic pathways, with over 10% of all enzymes adopting this fold. The majority of TIM barrels are thought to have evolved from a common ancestor through gene duplication and domain fusion processes. TIM barrels have been created by protein engineers using preexisting half-barrel templates and de novo, without an existing template. This review will discuss the topological, structural, evolutionary, and design characteristics of TIM barrels in detail.

Introduction

The TIM barrel is a structurally conserved protein fold, named after triose phosphate isomerase, a constituent enzyme of the glycolysis pathway whose structure was first solved in 19751. TIM barrels contain 200-250 amino acid residues2, which form 8 β-strands and 8 α-helices. The β-strands are arranged into a parallel β-barrel, and are surrounded by the 8 α-helices. The inner β-barrel is in many cases stabilized by intricate salt-bridge networks3. Loops at the C-terminal ends of the β-barrel are responsible for catalytic activity4, 5. Structural inserts ranging from extended loops to independent domains may be inserted in place of these loops or at the N/C-terminals. TIM barrels are ubiquitous, with approximately 10% of all enzymes adopting this fold6, 7. Further, 5 of 6 enzyme commission (EC) enzyme classes include TIM barrel proteins8, 9. The TIM barrel fold is evolutionarily ancient, with many of its members possessing little similarity today10, instead falling within the ‘twilight zone’ of sequence similarity11, 12. TIM barrels appear to have evolved through gene duplication and domain fusion events of half-barrel proteins13, with a majority of TIM barrels originating from a common ancestor. This lead many TIM barrels to possess internal symmetries14. Further gene duplication events of this ancestral TIM barrel lead to diverging enzymes possessing the functional diversity observed today. TIM barrels have also been a longstanding target for protein designers. Successful TIM barrel designs include both domain fusions of existing proteins and de novo designs. Domain fusions experiments have resulted in many successful designs15-21, whereas de novo designs only yielded successes after 28 years of incremental development22.

Structure

Topology

The X-ray crystallographic structure of triose phosphate isomerase (TIM) isolated from chicken muscles was first solved in 19751, lending its name to the TIM barrel fold (Figure 1A). TIM barrels contain 200-250 amino acid residues2, folded into 8 α-helices and 8 β-strands. The β-strands are arranged into a parallel β-barrel. The defining property of TIM β-barrels is that they always possess a shear number of 82, due to their staggered nature. The shear number is determined by picking a residue in a β-strand, and moving along the β-barrel until the original strand is reached. The number of residues separating the start and end positions is the shear number. Due to this stagger, the interior β-barrel residues (pore residues) are stacked with 4-fold geometric symmetry. The α-helices surround and completely enclose the inner β-barrel. Short loops typically connect the α and β secondary structures, forming a (βα)8 repeat topology. In some cases, structures ranging from extended loops to independent domains may be inserted in place of these loops, or may be attached to the N/C-terminals. All TIM barrel enzymes...
Figure 1. Triose phosphate isomerase (TIM) isolated from chicken muscles (PDB ID: 1TIM), the archetypal TIM barrel enzyme. (A) Cartoon representation of the TIM barrel structure. α-helices are colored teal, β-strands are colored orange, and loops are colored green. Note that the C-terminal ends of β-strands are depicted with arrowheads. (B) Core and pore regions are highlighted. Amino acid residues belonging to the pore are colored blue. Amino acid residues belonging to the core are colored orange. Note that the TIM barrel is depicted in a top-down view, where the C-terminal ends of the β-barrel are pointed towards the reader.

possess catalytic sites at the C-terminal end of the β-barrel and structural inserts present close to this end may aid in catalytic activity. The overall topology of the TIM barrel is illustrated in Figure 2.

Core and pore regions

TIM barrels contain two distinct buried regions, where amino acid residues are completely enveloped by their neighbors and lack access to solvent. These regions have been historically referred to as the 'core' and 'pore', and are depicted in Figure 1B. It should be noted that the term 'pore' is a misnomer, as no solvent channels exist within this region. The core region consists of all residues constituting the α-β interface, and lies exterior to the central β-barrel. The pore region consists of all interior β-barrel residues, which are surrounded and enclosed by the β-barrel backbone.

Due to the pleated nature of β-strands, alternate residues along a strand are almost evenly split between the pore (53%) and core (47%). For β-barrels, 95% of their core residues are buried. Consequently, only 11% of their core residues are polar, possessing an affinity for water, and possessing the ability to form hydrogen bonds or salt bridges. Similarly, 84% of β-strand pore residues are buried. However, 42% of their pore residues are polar. These residues form intricate salt bridge networks to compensate for their lack of solvent accessibility.

β-barrel salt bridge networks

Salt bridges within TIM barrel pores are thought to contribute to the overall stability of the fold. An example of a large salt bridge network network can be found in 2-deoxyribose-5-phosphate aldolase (Figure 3). This network was found to be conserved across the Class I aldolase family.

The exact reason for the overrepresentation of polar residues and salt bridges within the pore remains unclear. One study
Figure 2. TIM barrel topology. α-helices are colored teal, loops are colored green, and β-strands are colored in two shades of orange. Lighter shades indicate residues pointing inward, towards the barrel pore. Darker shades indicate residues pointing outward, towards the barrel core. Cyan lines depict an example backbone β-barrel hydrogen bonding network. Note that side-chain hydrogen bonding networks are not depicted here. Interior β-barrel residues (pore residues) display a 4-fold geometric symmetry, despite emerging from an 8-strand β-barrel. This symmetry is illustrated in red and blue. The shear number for TIM barrels is always 8, and is illustrated in magenta. Some TIM barrels naturally adopt, or are designed to adopt, two or four-fold symmetry. Example asymmetric units are also highlighted. This figure has been adapted with permission from previously published work.23
proposes that they improve foldability rather than thermodynamic stability of TIM barrels. During the folding process, inner pore residues on β-strands would be exposed to water. Partially-folded βαβα modules, called foldons, would be energetically stabilized by polar pore residues during this stage of folding. In another study involving the *S. solfataricus* indole-3-glycerol phosphate synthase TIM barrel protein, a conserved βαβαβ module was found to be an essential folding template, which guided the folding of other secondary structures. β-barrel closure only occurred at the end of the folding process.

Figure 3. Example salt bridge network in 2-deoxyribose-5-phosphate aldolase (PDB ID: 1P1X). Interactions are shown as cyan dashed lines. Polar residues are colored green. Polar amino acids aspartate (D), glutamate (E), lysine (K), and arginine (R), are shown here.

**Structural inserts**

The N/C-terminal and loop regions on TIM barrel proteins are capable of hosting structural inserts ranging from simple secondary structural motifs to complete domains. These domains aid in substrate recognition and catalytic activity. Four diverse examples of TIM barrels containing additional motifs and domains are discussed below.

**Bacillus subtilis** Orotidine 5-monophosphate decarboxylase (PDB ID: 1DBT, Figure 4A) is a TIM barrel protein displaying 4 α-helices in place of the βα loops (at the C-terminal of the β-barrel. One of these helices (R215→K219) contains a conserved arginine residue (R215) required for interacting with a phosphate moiety on orotidine 5-monophosphate. The other helices were not found to host residues critical for catalytic activity, and may serve in structural roles.

**Mycobacterium tuberculosis** bifunctional histidine/tryptophan biosynthesis isomerase (PriA) (PDB ID: 2Y85, Figure 4B) possesses the ability to catalyse the conversion of both N-[(5-phosphoribosyl) formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (ProFAR) and N-(5’-phosphoribosyl)-anthranilate (PRA) into 1-(O-carboxyphenylamino)-1’-deoxyribose-5’-phosphate (CdRP). PriA is a TIM barrel enzyme that accommodates both substrates using active site loops (loops 1, 5, and 6, extended βα loops at the C-terminal end of the β-barrel) that change conformation depending on the reactant present. Loop 1 wraps over the active site only in the presence of ProFAR. Loop 5 wraps over the active site, adopting a β-sheet conformation in the presence of CdRP, or a knot-like conformation in the presence of ProFAR. Loop 6 wraps over the active site for all reactants.

**Lactococcus lactis** dihydroorotate dehydrogenase A (DHODA) (PDB ID: 2DOR, Figure 4C) is an example of a TIM barrel possessing β-sheets and extended loops over the C-terminal end of the β-barrel. DHODA catalyzes the oxidation of dihydroorotate
Figure 4. Examples of structural inserts at TIM barrel loop and N/C-terminal regions. (A) The *Bacillus subtilis* Orotidine 5-monophosphate decarboxylase (PDB ID: 1DBT). Orotidine 5-monophosphate is colored green. α-helical inserts are colored teal. The catalytic arginine residue (R215) is displayed as sticks. (B) *Mycobacterium tuberculosis* bifunctional histidine/tryptophan biosynthesis isomerase (PriA) (PDB ID: 2Y85). The product CdRP is colored green. β-strand/loop interchangeable structures are colored orange. (C) *Lactococcus lactis* dihydroorotate dehydrogenase A (DHODA) (PDB ID: 2DOR). β-strands forming a sheet are colored orange. Extended loops are colored green. The cavity formed by these structures is displayed as a blue mesh. The product orotate is colored magenta. the cofactor FMN is colored pink. (D) *Methylophilus methylotrophus* trimethylamine dehydrogenase (PDB ID: 2TMD). The Rossmann fold domain is colored according to secondary structural elements. Cofactor FMN is colored magenta. The [4Fe-4S]$^+$ is colored red. Note that substrate/product were not crystallized.
to orotate\textsuperscript{30,31}, which is part of the \textit{de novo} uridine 5'-monophosphate (UMP) synthesis pathway. This oxidation is mediated by flavin mononucleotide (FMN). Here, $\beta$-sheets and extended loops enclose the active site forming a cavity, while also hosting several catalytic residues.

The \textit{Methylophilus methylotrophus} trimethylamine dehydrogenase\textsuperscript{26} (PDB ID: 2TMD, Figure 4D) TIM barrel is an example of a complete domain insertion. Here, a Rossmann fold domain is inserted at the C-terminal end of the TIM-barrel. Trimethylamine dehydrogenase catalyzes the conversion of trimethylamine to formaldehyde\textsuperscript{32}. This reaction requires both a reduced 6-S-cysteinyl flavin mononucleotide (FMN) cofactor and a reduced iron-sulphur ([4Fe-4S]\textsuperscript{2+}) center. FMN is covalently bound within the C-terminal region of the $\beta$-barrel. The [4Fe-4S]\textsuperscript{2+} center is too large to be accommodated within the TIM barrel, and is instead placed in close proximity, 7 Å away, at the interface between the TIM barrel and Rossmann fold domains.

\section*{Evolution and origins}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig5}
\caption{Model for the evolution of TIM barrels through gene duplication and domain fusion, as proposed by Lang \textit{et al.}\textsuperscript{13}. This model described the evolution of enzymes HisA and HisF of the histidine biosynthesis pathway. Two gene duplication steps are thought to have occurred. The first gene duplication resulted in two half-barrels that later fused and evolved into an ancestral TIM barrel. The second gene duplication event lead to diversification, and the evolution of different TIM barrel enzymes catalyzing different reactions.}
\end{figure}

TIM barrels have evolved through gene duplication and fusion, starting with a half-barrel and eventually forming an enzymatically active TIM barrel. Multiple studies support the theory of divergent evolution from a single ancestor, and are discussed
Evolution from a common ancestor

In the early 1990s, Farber et al. suggested that all TIM barrel structures solved at the time were enzymes, indicating divergence from a common ancestor. Bränden et al. further suggested that a common phosphate binding site, formed by a small α-helix and TIM barrel loops-7/8, strongly indicated divergent evolution. Copley et al. further studied these phosphate groups, concluding that 12 of 23 SCOP (structural classification of proteins) TIM barrel families diverged from a common ancestor. Based on these reports, it is plausible that the majority of TIM barrel proteins evolved from a common ancestor.

Origin through gene duplication and domain fusion

Many TIM barrel proteins possess 2-fold, 4-fold or 8-fold internal symmetry, suggesting that TIM barrels evolved from ancestral (βα)₄, (βα)₂, or βα motifs through gene duplication and domain fusion. A good example of 2-fold internal symmetry is observed in the enzymes ProFAR isomerase (HisA) and imidazole glycerol phosphate synthase (HisF) of the *Thermotoga maritima* histidine biosynthesis pathway. They catalyze 2 successive reactions in the pathway, possess 25% sequence homology, and possess root-mean-square deviations (RMSDs) between 1.5-2Å, suggesting divergence from a common ancestor. More interestingly, the loops on the C terminal ends of both HisA and HisF showed a twofold repeated pattern, suggesting that their common ancestor also possessed 2-fold internal symmetry. Using these observations, Lang et al. constructed a model for the evolution of the TIM barrels (Figure 5). An ancestral half-barrel would have undergone a gene duplication and fusion event, resulting in a single protein containing two half-barrel domains. Structural adaptations would have occurred, resulting in the merging of these domains to form a closed β-barrel, and forming an ancestral TIM barrel. Functional adaptations would have also occurred, resulting in the evolution of new catalytic activity at the C terminal end of the β-barrel. At this point, the common ancestor of HisA and HisF would have undergone a second gene duplication event. Divergent evolution of the duplicated genes of the ancestral TIM barrel would have resulted in the formation of HisA and HisF.

Interestingly, this evolutionary model has been experimentally validated using directed evolution and protein design techniques. Seitz et al. constructed proteins HisF-C*C and HisF-C***C from C-terminal HisF half-barrels. A salt-bridge cluster present in wild-type HisF was reconstructed, and random mutagenesis was performed to stabilize and solubilize the construct. The crystal structure of HisF-C***C revealed a 2-fold symmetric TIM barrel, validating the possibility of natural domain fusion. Similar experiments were performed by Höcker et al. using HisA and HisF half-barrels, resulting in the successful creation of a chimeric HisA-HisF TIM barrel. These experiments lead Höcker et al. to propose a novel means of diversification and evolution of TIM-barrel enzymes through the exchange of (βα)₄ half-barrel domains amongst preexisting TIM barrels. Other fusion experiments to generate new TIM barrels have been performed. A chimeric Phosphoribosylanthranilate isomerase (TrpF)/HisA TIM barrel with wild-type catalytic activity, chimeric (βα)₅-flavodoxin-like fold (CheY)/HisF TIM barrels, and a perfectly 2-fold symmetric HisF-based TIM barrel have all been created.

The existence of 4/8-fold internal symmetry was suggested by Söding et al. based on a computational analysis of TIM barrel sequences. For example, *Escherichia coli* KDPG aldolase (PDB ID: 1FQ0) was suggested to possess a distinct 4-fold symmetry, with discernible 8-fold symmetry. The design of a 4-fold symmetric TIM barrel confirmed the possibility of higher orders of internal symmetry in natural TIM barrels, and will be discussed in detail in the next section. It should be noted that no experimental evidence for the existence of 8-fold symmetric TIM barrels has been reported to date.
Figure 6. sTIM-11, the first successful de novo TIM barrel design. The asymmetric \((\alpha\beta)_2\) units are colored distinctly, highlighting the internal 4-fold symmetry.

De novo TIM barrel design

The TIM barrel fold has been a long-standing target for de novo protein designers. As previously described, numerous TIM barrels have been successfully designed based on preexisting natural half-barrels. In contrast, the de novo design of TIM barrels occurred in incremental steps over a period of 28 years.

The Octarellin series of proteins (Octarellin I → VI) were the first attempts to create a de novo TIM barrel. As the field of protein design was still in its infancy, these design attempts were only met with limited success. Although they displayed circular dichroism spectra consistent with \(\alpha\beta\) proteins and some cooperative folding characteristics, all Octarellin series peptides were insoluble, and had to be resolubilized from inclusion bodies for further characterization. Interestingly, Octarellin V.1 displayed a Rossmann-like fold under co-crystal conditions.

The Symmetrin series of proteins (Symmetrin-1 → 4) displayed more favorable biophysical characteristics. Symmetrin-1 was readily soluble, displayed circular dichroism spectra consistent with \(\alpha\beta\) proteins, and displayed excellent cooperative unfolding and refolding characteristics. Despite these advances, all proteins in this family displayed molten characteristics when analyzed using nuclear magnetic resonance (NMR), and further work to solve their structures could not be pursued.

Proteins of the sTIM series represented the first successful de novo TIM barrel design. sTIM-11 (PDB ID: 5BVL) was designed with an internal 4-fold symmetry, to reduce the complexity of computational design using the Rosetta software suite. Previously-derived first principles were used to delineate secondary structure topologies and lengths. sTIM-11 proved to be a highly thermostable, cooperatively folding design that adopted its intended structure (Figure 6).
Höcker, B., Claren, J. & Sterner, R. Mimicking enzyme evolution by generating new (17).

Livesay, D. R. & La, D. The evolutionary origins and catalytic importance of conserved electrostatic networks within (11).

Hocker, B., Lochner, A., Seitz, T., Claren, J. & Sterner, R. High-resolution crystal structure of an artificial (16).

Brändén, C.-I. The tim barrel—the most frequently occurring folding motif in proteins: Current opinion in structural biology 1991, 1: 978–983. Curr. Opin. Struct. Biol. 1, 978–983 (1991).
25. Gu, Z., Zitzewitz, J. A. & Matthews, C. R. Mapping the structure of folding cores in tim barrel proteins by hydrogen exchange mass spectrometry: the roles of motif and sequence for the indole-3-glycerol phosphate synthase from sulfolobus solfataricus. *J. molecular biology* 368, 582–594 (2007).

26. Barber, M. J., Neame, P. J., Lim, L. W., White, S. & Matthews, F. Correlation of x-ray deduced and experimental amino acid sequences of trimethylamine dehydrogenase. *J. Biol. Chem.* 267, 6611–6619 (1992).

27. Appleby, T. C., Kinsland, C., Begley, T. P. & Ealick, S. E. The crystal structure and mechanism of orotidine 5-monophosphate decarboxylase. *Proc. Natl. Acad. Sci.* 97, 2005–2010 (2000).

28. Juárez-Vázquez, A. L. et al. Evolution of substrate specificity in a retained enzyme driven by gene loss. *Elife* 6, e22679 (2017).

29. Due, A. V., Kuper, J., Geerlof, A., von Kries, J. P. & Wilmanns, M. Bisubstrate specificity in histidine/tryptophan biosynthesis isomerase from mycobacterium tuberculosis by active site metamorphosis. *Proc. Natl. Acad. Sci.* 108, 3554–3559 (2011).

30. Jang, M.-H., Basran, J., Scrutton, N. S. & Hille, R. The reaction of trimethylamine dehydrogenase with trimethylamine. *J. Biol. Chem.* 274, 13147–13154 (1999).

31. Borman, S. Protein designers roll out a barrel. *Chem. & Eng. News* 93, 6–6 (2015).

32. Goraj, K., Renard, A. & Martial, J. A. Synthesis, purification and initial structural characterization of octarellin, a de novo polypeptide modelled on the α/β-barrel proteins. *Protein Eng. Des. Sel.* 3, 259–266 (1990).

33. Figueroa, M. et al. Spectroscopic investigation of structure in octarellin vi: Using rosetta to design a putative artificial (β/α) 8 protein. *PLoS one* 8, e71858 (2013).

34. Offredi, F. et al. De novo backbone and sequence design of an idealized αβ-barrel protein: evidence of stable tertiary structure. *J. molecular biology* 325, 163–174 (2003).

35. Figueroa, M. et al. The unexpected structure of the designed protein octarellin v. 1 forms a challenge for protein structure prediction tools. *J. structural biology* 195, 19–30 (2016).

36. Nanda, V. Protein design: Getting to the bottom of the tim barrel. *Nat. chemical biology* 12, 2 (2016).

37. Kaufmann, K. W., Lemmon, G. H., DeLuca, S. L., Sheehan, J. H. & Meiler, J. Practically useful: what the rosetta protein modeling suite can do for you. *Biochemistry* 49, 2987–2998 (2010).

38. Koga, N. et al. Principles for designing ideal protein structures. *Nature* 491, 222 (2012).