The TIM barrel fold

Nagarajan D. and Nanajkar N.

Comments and corrections:

Line 10: fix "*αhelices*" in "α-helices".

Lines 11-12: C-terminal loops are important for catalytic activity, while N-terminal loops are important for the stability of the TIM-barrels. This should be mentioned.

Line 14: The reference #7 is not related to the statement.

Line 14: There is a new EC classe (EC.7, translocases). Change "5 of 6" in "5 of 7".

Lines 26-27: It is not correct to state that the shear number of 8 for the TIM-barrels is due to "their staggered nature". Most of the β-barrels have a staggered nature, but their shear number is not 8.

Line 27: The reference #2 is imprecise. Wierenga did not defined himself the shear number of TIM-barrel proteins. Please check the 2 papers of Murzin AG, 1994, "Principle determining the structure of β-sheet barrels in proteins," I and II, and the paper of Liu W, 1998, "Shear numbers of protein β-barrels: definition refinements and statistics".

Line 29: Again, it is not correct to state that the 4-fold geometric symmetry depends on the stagger. Since the number of strands (n) is equal to the Shear number (S), side-chains point alternatively towards the pore and the core, giving a 4-fold symmetry.

Line 37: "historically" is a bit exaggerated for a reference dated 2015, especially if it comes from the author itself. Find a true historic reference, or just mention that you defined the regions "core" and "pore".

Line 43: "Consequently" is misleading. The fact that 11% of the core residues are polar does not depend by the fact that 95% of core residues are buried.

Lines 55-57: Reference #25 support the idea that the folding process is driven by hydrophobic interactions of branched aliphatic side-chains (leucine, isoleucine and valine). This theory is opposite to the one that you mention in lines 53-55 (polar residues stabilizing the foldons). Please make it clear that there are evidences for both theories.

Line 64: There is an open parenthesis that is never closed.

Lines 90-91: The fact that TIM-barrels evolved from a single ancestor, following gene duplication and fusion, is still a theory (the most supported, but still a theory). Please make it clear that it is a theory in this introductory sentence. Moreover pay attention to the sentence "forming an enzymatically active TIM-barrel", it suggest that the half barrels have no functions and that only TIM-barrels became enzymes. Evolutionary speaking it is quite unlikely that the half-barrels had no function, however there are no evidences to support (or deny) this theory. I will simply use "forming the actual structure of the TIM-barrels" or something similar, without references to the function.

Lines 118-124: You should re-organize this paragraph. Höcker et al. (reference #17) are the firsts that designed HisF-C*C in their paper of 2004, and should be mentioned at the beginning. Seitz et al. (reference #15) used the HisF-C*C designed by Höcker as basis to create HisF-C***C, which was then crystallized and its structure solved in 2009 (reference #16).

The TIM barrel fold

Deepesh Nagarajan[1*] **and Neha Nanajkar**[1]

[1]Department 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 1975[1]. TIM barrels contain 200-250 amino acid residues[2], which form 8 β-strands and 8 $\alpha helices$. 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 networks[3]. Loops at the C-terminal ends of the β-barrel are responsible for catalytic activity[4,5]. 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 fold[6,7]. Further, 5 of 6 enzyme commission (EC) enzyme classes include TIM barrel proteins[8,9]. The TIM barrel fold is evolutionarily ancient, with many of its members possessing little similarity today[10], instead falling within the 'twilight zone' of sequence similarity[11,12]. TIM barrels appear to have evolved through gene duplication and domain fusion events of half-barrel proteins[13], with a majority of TIM barrels originating from a common ancestor. This lead many TIM barrels to possess internal symmetries[14]. 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 designs[15–21], whereas *de novo* designs only yielded successes after 28 years of incremental development[22].

Structure

Topology

The X-ray crystallographic structure of triose phosphate isomerase (TIM) isolated from chicken muscles was first solved in 1975[1], lending its name to the TIM barrel fold (Figure 1A). TIM barrels contain 200-250 amino acid residues[2], 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 8[2], 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 $(\beta\alpha)_8$ 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

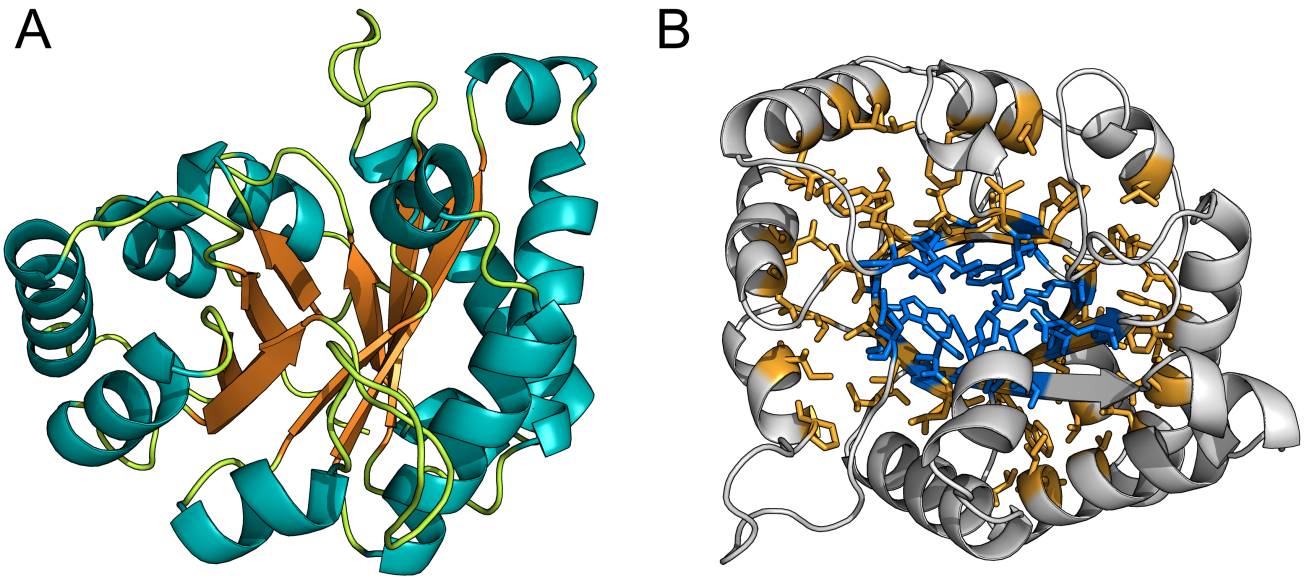

Figure 1. Triose phosphate isomerase (TIM) isolated from chicken muscles (PDB ID: 1TIM), the archetypal TIM barrel enzyme. **(A)** Cartoon representation of the TIM barrel structure. α-helices are colored teal, β-strands are colored orange, and loops are colored green. Note that the C-terminal ends of 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[24], 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'[23], 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%)[23]. 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[3].

The exact reason for the overrepresentation of polar residues and salt bridges within the pore remains unclear. One study[23]

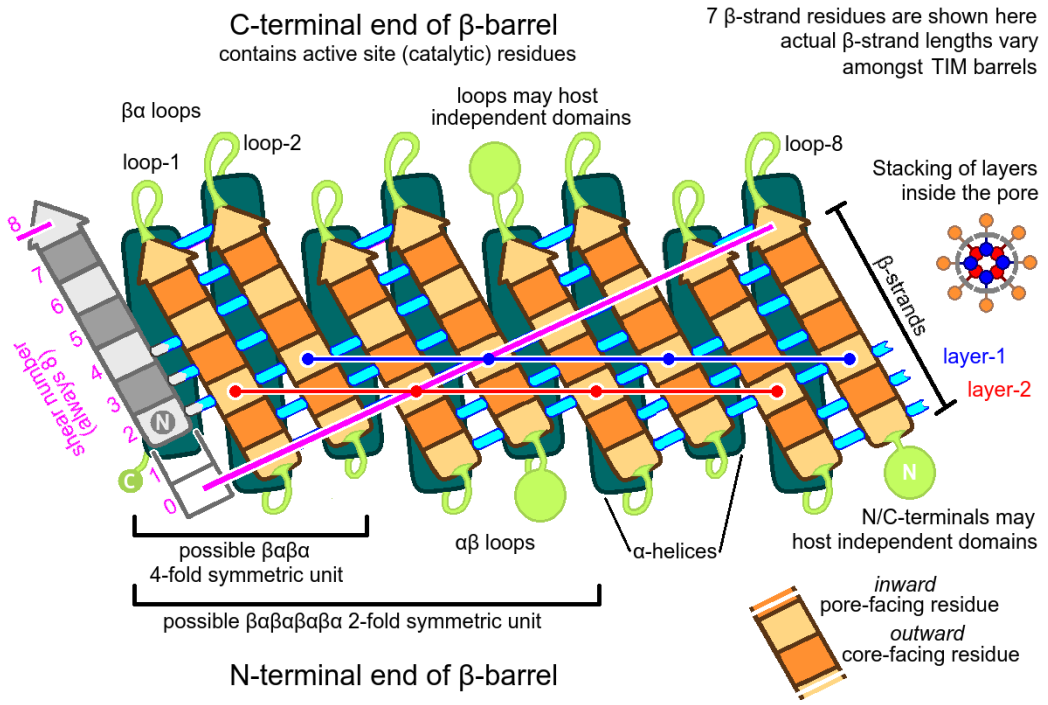

Figure 2. TIM barrel topology. α-helices are colored teal, loops are colored green, and β-strands are colored in two shades of orange. Lighter shades indicate residues pointing *inward*, towards the barrel pore. Darker shades indicate residues pointing *outward*, towards the barrel core. Cyan lines depict 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 $\beta\alpha\beta\alpha$ modules, called foldons, would be energetically stabilized by polar pore residues during this stage of folding. In another study[25] involving the *S. solfataricus* indole-3-glycerol phosphate synthase TIM barrel protein, a conserved $\beta\alpha\beta\alpha\beta$ 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.

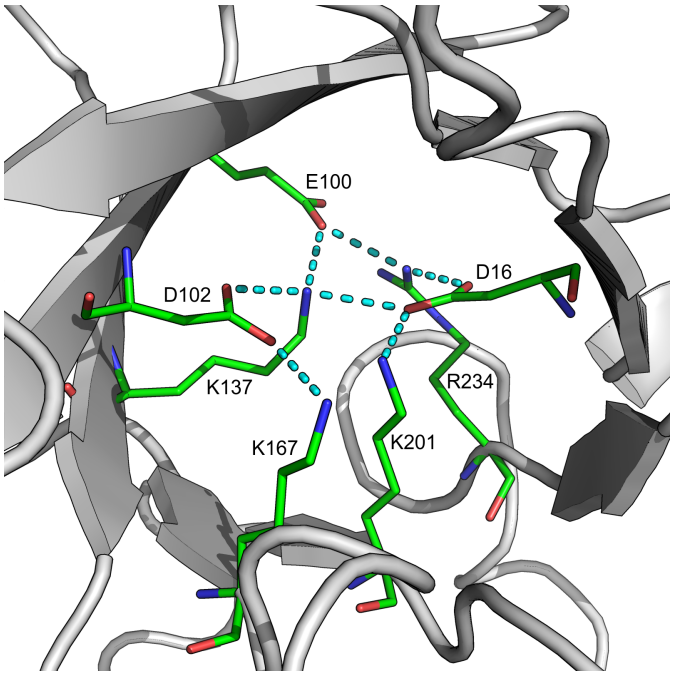

Figure 3. Example salt bridge network in 2-deoxyribose-5-phosphate aldolase (PDB ID: 1P1X). Interactions are shown as cyan dashed lines. Polar residues are colored green. Polar amino acids aspartate (D), glutamate (E), lysine (K), and arginine (R), are shown here.

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[27] (PDB ID: 1DBT, Figure 4A) is a TIM barrel protein displaying 4 α-helices in place of the $\beta\alpha$ loops (at the C-terminal of the β-barrel. One of these helices (R215$\rightarrow$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'-deoxyribulose-5'-phosphate (CdRP)[28]. PriA is a TIM barrel enzyme that accommodates both substrates using active site loops (loops 1, 5, and 6, extended $\beta\alpha$ loops at the C-terminal end of the β-barrel) that change conformation depending on the reactant present[29]. Loop 1 wraps over the active site only in the presence of ProFAR. Loop5 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

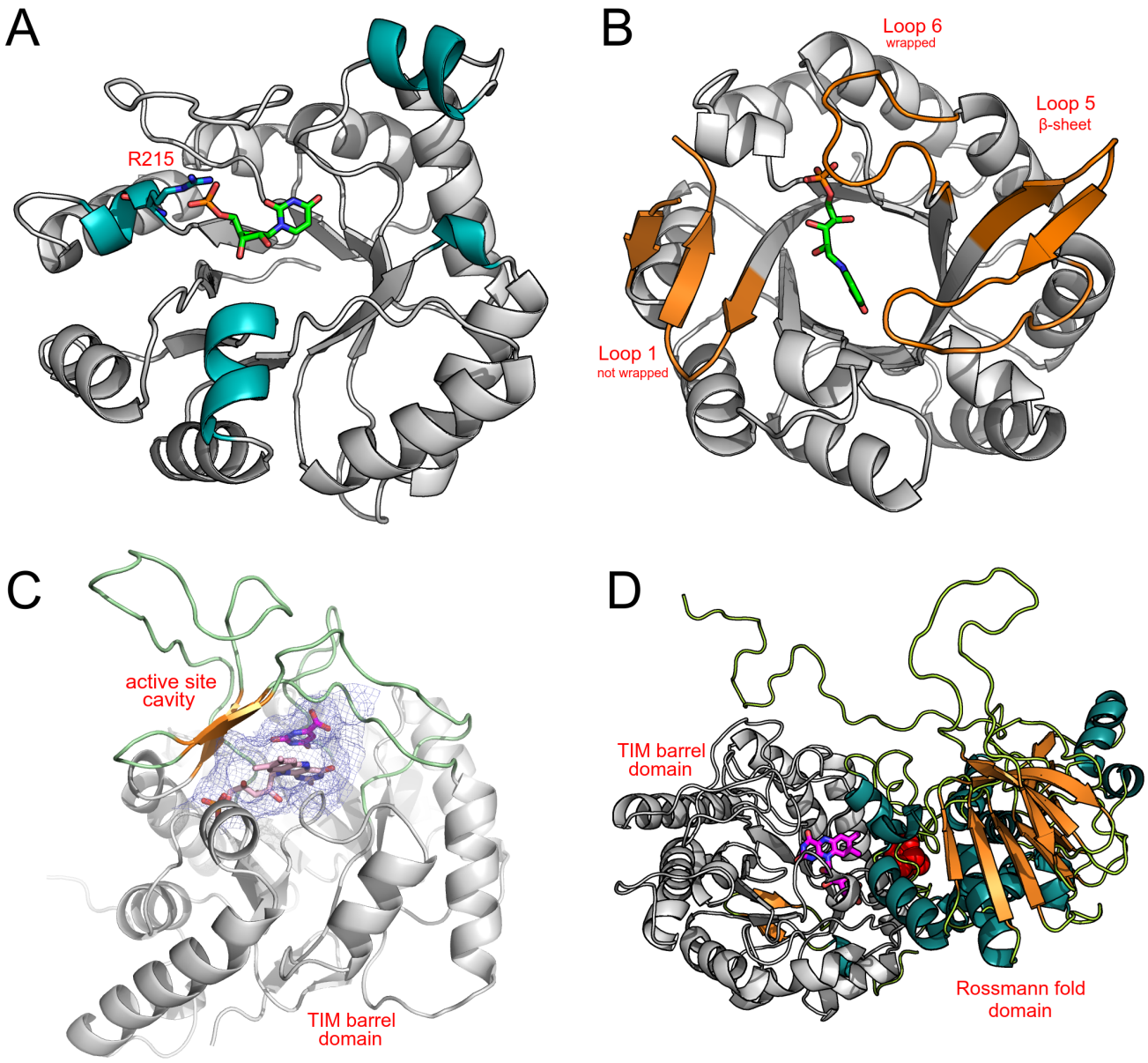

Figure 4. Examples of structural inserts at TIM barrel loop and N/C-terminal regions. **(A)** The *Bacillus subtilis* Orotidine 5-monophosphate decarboxylase (PDB ID: 1DBT). Orotidine 5-monophosphate is colored green. α-helical inserts are colored teal. The catalytic arginine residue (R215) is displayed 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[26] (PDB ID: 2TMD). The Rossmann fold domain is colored according to secondary structural elements. Cofactor FMN is colored magenta. The $[4Fe\text{-}4S]^{+}$ is colored red. Note that substrate/product were not crystallized.

to orotate[30,31], which is part of the *de novo* uridine 5'-monophosphate (UMP) synthesis pathway. This oxidation is mediated by flavin mononucleotide (FMN). Here, β-sheets and extended loops enclose the active site forming a cavity, while also hosting several catalytic residues.

The *Methylophilus methylotrophus* trimethylamine dehydrogenase[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[32]. This reaction requires both a reduced 6-S-cysteinyl Flavin mononucleotide (FMN) cofactor and a reduced iron-sulphur ($[4Fe\text{-}4S]^{+}$) center. FMN is covalently bound within the C-terminal region of the β-barrel. The $[4Fe\text{-}4S]^{+}$ 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.

Evolution and origins

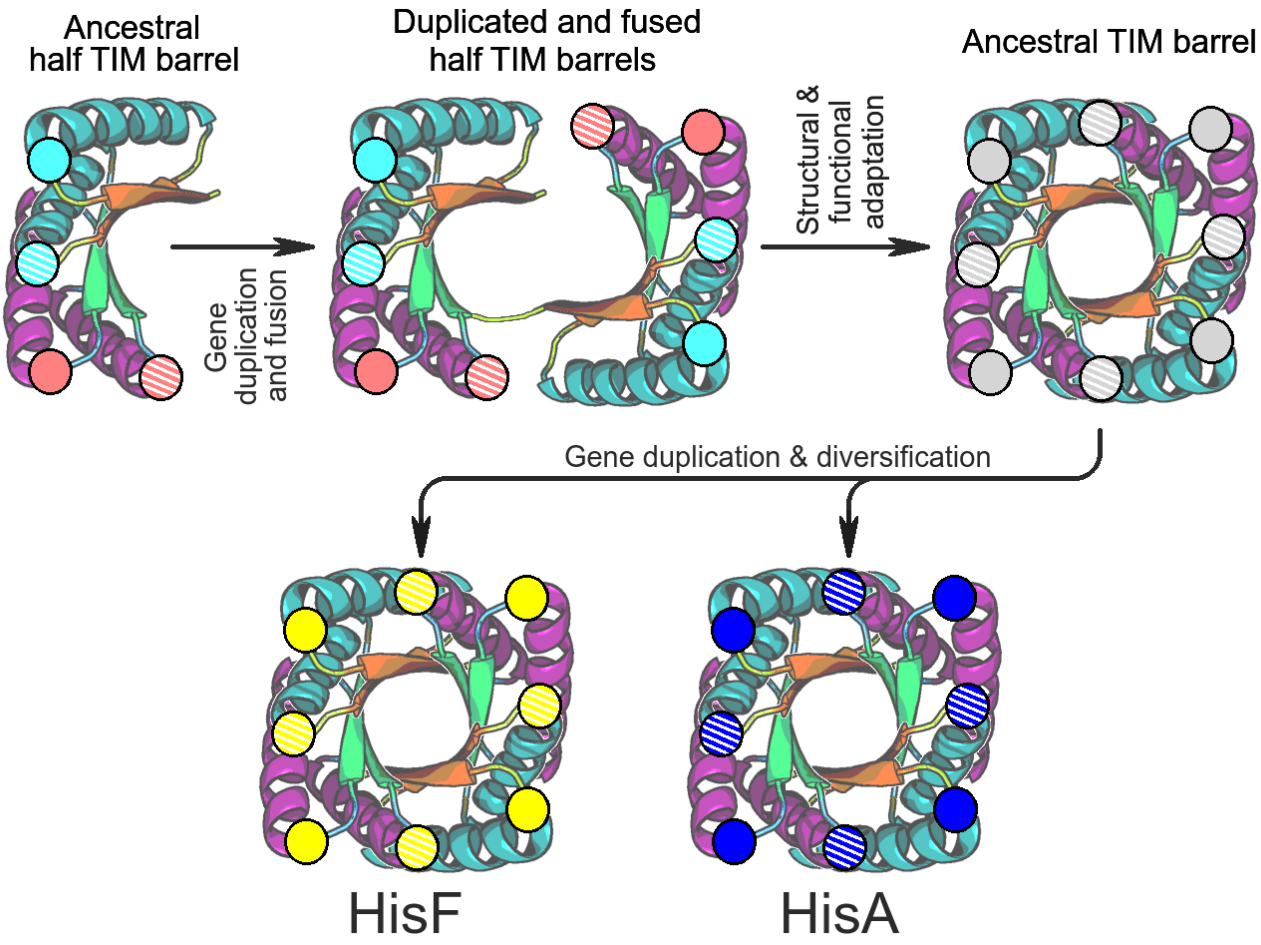

Figure 5. Model for the evolution of TIM barrels through gene duplication and domain fusion, as proposed by Lang *et al.*[13]. This model described the evolution of enzymes HisA and HisF of the histidine biosynthesis pathway. Two gene duplication steps are thought to have occurred. The first gene duplication 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.

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

below.

Evolution from a common ancestor

In the early 1990s, Farber *et al.*[4,5] noted that all TIM barrel structures solved at the time were enzymes, indicating divergence from a common ancestor. Further, all TIM barrels possessed active sites at the C-terminal end of β-barrels. Bränden *et al.*[24] 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.*[7] further studied these phosphate groups, concluding that 12 of 23 SCOP (structural classification of proteins)[33] TIM barrel families diverged from a common ancestor. Similarly, Nagano *et al.*[10] concluded that there were hints for common ancestry for 17 of the 21 CATH (classification of protein domain structures)[34] TIM barrel families. 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 $(\beta\alpha)_4$, $(\beta\alpha)_2$, or $\beta\alpha$ 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[13]. 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.*[13] 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[15] 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[16] 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[17]. These experiments lead Höcker *et al.* to propose a novel means of diversification and evolution of TIM-barrel enzymes through the exchange of $(\beta\alpha)_4$ 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[18], chimeric $(\beta\alpha)_5$-flavodoxin-like fold (CheY)/HisF TIM barrels[19,20], and a perfectly 2-fold symmetric HisF-based TIM barrel[21] have all been created.

The existence of 4/8-fold internal symmetry was suggested by Söding *et al.*[14] based on a computational analysis of TIM barrel sequences. For example, *Escherichia coli* KDPG aldolase[35] (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[22] 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.

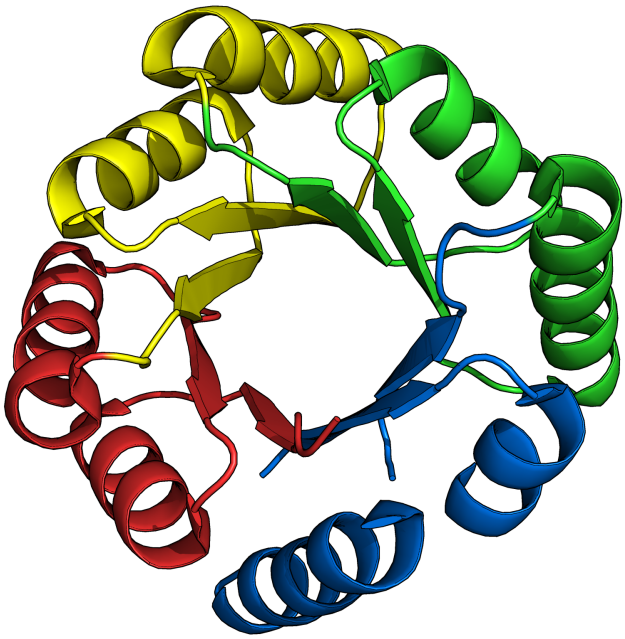

Figure 6. sTIM-11, the first successful *de novo* TIM barrel design. The asymmetric $(\alpha\beta)_2$ units are colored distinctly, highlighting the internal 4-fold symmetry.

De novo TIM barrel design

The TIM barrel fold has been a long-standing target for *de novo* protein designers. As previously described, numerous[15–21] 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[36].

The Octarellin series[37–41] 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[42] displayed a Rossmann-like fold under co-crystal conditions.

The Symmetrin series[23] 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[22] represented the first successful *de novo* TIM barrel design[36,43]. 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[44]. Previously-derived first principles[45] 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).

References

1. Banner, D. *et al.* Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 å resolution: using amino acid sequence data. *Nature* **255**, 609 (1975).
2. Wierenga, R. The tim-barrel fold: a versatile framework for efficient enzymes. *FEBS letters* **492**, 193–198 (2001).
3. Vijayabaskar, M. & Vishveshwara, S. Insights into the fold organization of tim barrel from interaction energy based structure networks. *PLoS computational biology* **8**, e1002505 (2012).
4. Farber, G. K. & Petsko, G. A. The evolution of α/β barrel enzymes. *Trends biochemical sciences* **15**, 228–234 (1990).
5. Farber, G. & Petsko, G. The structure and evolution of alpha/beta barrel proteins. *The FASEB J.* **9**, 497–503 (1995).
6. Jansen, R. & Gerstein, M. Analysis of the yeast transcriptome with structural and functional categories: characterizing highly expressed proteins. *Nucleic Acids Res.* **28**, 1481–1488 (2000).
7. Copley, R. R. & Bork, P. Homology among $(\beta\alpha)$ 8 barrels: implications for the evolution of metabolic pathways. *J. molecular biology* **303**, 627–641 (2000).
8. Nagano, N., Hutchinson, E. G. & Thornton, J. M. Barrel structures in proteins: automatic identification and classification including a sequence analysis of tim barrels. *Protein Sci.* **8**, 2072–2084 (1999).
9. Webb, E. C. *et al. Enzyme nomenclature 1992. Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes.* Ed. 6 (Academic Press, 1992).
10. Nagano, N., Orengo, C. A. & Thornton, J. M. One fold with many functions: the evolutionary relationships between tim barrel families based on their sequences, structures and functions. *J. molecular biology* **321**, 741–765 (2002).
11. Livesay, D. R. & La, D. The evolutionary origins and catalytic importance of conserved electrostatic networks within tim-barrel proteins. *Protein science* **14**, 1158–1170 (2005).
12. Chung, S. Y. & Subbiah, S. A structural explanation for the twilight zone of protein sequence homology. *Structure* **4**, 1123–1127 (1996).
13. Lang, D., Thoma, R., Henn-Sax, M., Sterner, R. & Wilmanns, M. Structural evidence for evolution of the β/α barrel scaffold by gene duplication and fusion. *Science* **289**, 1546–1550 (2000).
14. Söding, J., Remmert, M. & Biegert, A. Hhrep: de novo protein repeat detection and the origin of tim barrels. *Nucleic acids research* **34**, W137–W142 (2006).
15. Seitz, T., Bocola, M., Claren, J. & Sterner, R. Stabilisation of a $(\beta\alpha)$ 8-barrel protein designed from identical half barrels. *J. molecular biology* **372**, 114–129 (2007).
16. Hocker, B., Lochner, A., Seitz, T., Claren, J. & Sterner, R. High-resolution crystal structure of an artificial $(\beta\alpha)$ 8-barrel protein designed from identical half-barrels. *Biochemistry* **48**, 1145–1147 (2009).
17. Höcker, B., Claren, J. & Sterner, R. Mimicking enzyme evolution by generating new $(\beta\alpha)$ 8-barrels from $(\beta\alpha)$ 4-half-barrels. *Proc. Natl. Acad. Sci.* **101**, 16448–16453 (2004).
18. Claren, J., Malisi, C., Höcker, B. & Sterner, R. Establishing wild-type levels of catalytic activity on natural and artificial $(\beta\alpha)$ 8-barrel protein scaffolds. *Proc. Natl. Acad. Sci.* **106**, 3704–3709 (2009).
19. Bharat, T. A., Eisenbeis, S., Zeth, K. & Höcker, B. A $\beta\alpha$-barrel built by the combination of fragments from different folds. *Proc. Natl. Acad. Sci.* **105**, 9942–9947 (2008).
20. Eisenbeis, S. *et al.* Potential of fragment recombination for rational design of proteins. *J. Am. Chem. Soc.* **134**, 4019–4022 (2012).
21. Fortenberry, C. *et al.* Exploring symmetry as an avenue to the computational design of large protein domains. *J. Am. Chem. Soc.* **133**, 18026–18029 (2011).
22. Huang, P.-S. *et al.* De novo design of a four-fold symmetric tim-barrel protein with atomic-level accuracy. *Nat. chemical biology* **12**, 29 (2016).
23. Nagarajan, D., Deka, G. & Rao, M. Design of symmetric tim barrel proteins from first principles. *BMC biochemistry* **16**, 18 (2015).
24. 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. Rowland, P., Larsen, S., Björnberg, O., Nielsen, F. S. & Jensen, K. F. The crystal structure of lactococcus lactis dihydroorotate dehydrogenase a complexed with the enzyme reaction product throws light on its enzymatic function. *Protein Sci.* **7**, 1269–1279 (1998).

31. Björnberg, O., Rowland, P., Larsen, S. & Jensen, K. F. Active site of dihydroorotate dehydrogenase a from lactococcus lactis investigated by chemical modification and mutagenesis. *Biochemistry* **36**, 16197–16205 (1997).

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

33. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. Scop: a structural classification of proteins database for the investigation of sequences and structures. *J. molecular biology* **247**, 536–540 (1995).

34. Orengo, C. A. *et al.* Cath–a hierarchic classification of protein domain structures. *Structure* **5**, 1093–1109 (1997).

35. Wymer, N. *et al.* Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from escherichia coli. *Structure* **9**, 1–9 (2001).

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

37. 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).

38. Beauregard, M. *et al.* Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the α/β-barred packing). *Protein Eng. Des. Sel.* **4**, 745–749 (1991).

39. Houbrechts, A. *et al.* Second-generation octarellins: two new de novo (β/α) 8 polypeptides designed for investigating the influence of β-residue packing on the α/β-barrel structure stability. *Protein Eng. Des. Sel.* **8**, 249–259 (1995).

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

41. Figueroa, M. *et al.* Octarellin vi: Using rosetta to design a putative artificial (β/α) 8 protein. *PloS one* **8**, e71858 (2013).

42. 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).

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

44. 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).

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