

TIM barrel folding mechanism

The conservation of the TIM barrel fold is mirrored by the conservation of its equilibrium and kinetic folding mechanisms in bacterial paralogs with phylogenetically distinct lineages. Chemical denaturation of several natural^{1,2} and 2 designed TIM barrel variants², invariably involves a highly populated equilibrium intermediate. The kinetic intermediates that appear after dilution from highly denaturing solutions involve an early misfolded species that must at least partially unfold to access the productive folding pathway^{1,2}. The rate-limiting step in folding is the closure of the 8-stranded β -barrel, with the preceding, open barrel form corresponding to the equilibrium intermediate³ (Fig. 1). Native-centric molecular dynamics simulations recapitulate the experimental results and point the way to testable computational models for complex folding mechanisms⁴.

Fitness landscapes for the TIM barrel fold

TIM barrel proteins possess an unusually high sequence plasticity, forming large families of orthologous and paralogous enzymes in widely divergent organisms. This plasticity suggests a sequence landscape that allows for protein adaptation to a variety of environmental conditions, largely independent of phylogenetic history, while maintaining function. Chan et al⁵ used a deep mutational scanning approach and a competition assay⁶ to determine the fitness of all possible amino acid mutants across 80 positions in 3 hyperthermophilic indole-3-glycerolphosphate synthase (IGPS) TIM barrel enzymes in supporting the growth of a yeast host lacking IGPS. Although the 2 bacterial and 1 archaeal IGPS enzymes were only 30-40% identical in sequence, their fitness landscapes were strongly correlated: the same amino acids at the same positions in the three different proteins had very similar fitness. The correlation can be thought of as the conservation of the fitness landscape for a TIM barrel enzyme across evolutionary time (Fig. 2).

TIM barrel structures and stability

(To be added to the paragraph describing the stabilizing electrostatic network at the N-terminus of the β -barrel).

Another stabilizing element in TIM barrels is the $\beta\alpha$ -hairpin clamp. Side chain H-bond donors at the N-termini of even-numbered β -strands often form H-bonds with main chain amide hydrogens in preceding odd-numbered β -strands^{7,8}. These clamps (or hydrophobic side chain bridge analogs) are conserved in 3 indole-3-glycerolphosphate synthase TIM barrel orthologs from the bacterial and archaeal kingdoms⁵, implying they arose in their last common ancestor and have been preserved for over a billion years.

References

1. Topology and sequence in the folding of a TIM barrel protein: global analysis highlights partitioning between transient off-pathway and stable on-pathway folding intermediates in the complex folding mechanism of a $(\beta\alpha)_8$ barrel of unknown function from *B. subtilis*. W. R. Forsyth, O. Bilsel, Z. Gu and C. R. Matthews, *J. Mol. Biol.* (2007) 372, 236-253. DOI: 10.1016/j.jmb.2007.06.018
2. Conservation of the folding mechanism between designed primordial $(\beta\alpha)_8$ -Barrel proteins and their modern descendant. L. Carstensen, J. M. Sperl, M. Bocola, F. List, F. X. Schmid, and Reinhard Sterner. *J. Am. Chem. Soc.* *J. Am. Chem. Soc.* (2012), 134, 12786-12791. DOI: 10.1021/ja304951v
3. Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Gō-model simulation, Z. Gu, M. K. Rao, W. R. Forsyth, J. M. Finke, and C. R. Matthews, *J. Mol. Biol.* (2007) 374(2):528-46. DOI: 10.1016/j.jmb.2007.09.024
4. Frustration and folding in a TIM barrel protein. K. T. Halloran, Y. Wang, K. Arora, S. Chakravarthy, T. C. Irving, O. Bilsel, C. L. Brooks III, and C. R. Matthews, *Proc. Natl. Acad. Sci.* (2019) 116 (33) 16378-16383. DOI: 10.1073/pnas.1900880116
5. Correlation of fitness landscapes from three orthologous TIM barrels originates from sequence and structure constraints, Y. H. Chan, S. V. Venev, K. B. Zeldovich, and C. R. Matthews, *Nat. Commun.* 2017 Mar 6;8:14614. DOI: 10.1038/ncomms14614
6. Experimental illumination of a fitness landscape, R. T. Hietpas, J. D. Jensen, and D. N. A. Bolon, *Proc. Natl. Acad. Sci.* (2011) 108 (19) 7896-7901. DOI: 10.1073/pnas.1016024108
7. Long-range side-chain–main-chain interactions play crucial roles in stabilizing the $(\beta\alpha)_8$ barrel motif of the alpha subunit of tryptophan synthase, X. Yang, R. Vadrevu, Y. Wu and C. R. Matthews, *Protein Sci.* (2007) 16, 1398-1409. DOI: 10.1110/ps.062704507
8. $\beta\alpha$ -hairpin clamps brace $\beta\alpha\beta$ modules and can make substantive contributions to the stability of TIM barrel proteins. X. Yang, S. V. Kathuria, R. Vadrevu, and C. R. Matthews, *PLOS One* (2009). DOI: 10.1371/journal.pone.0007179.

Figures

Figure 1

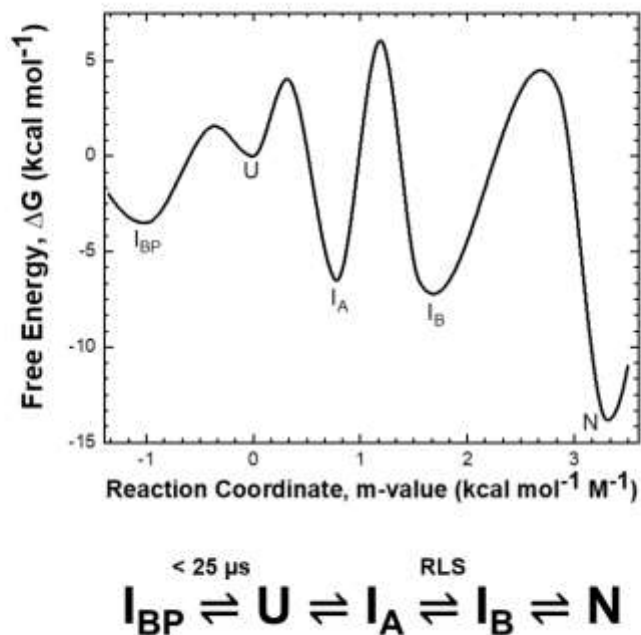


Fig. 1. The reaction coordinate diagram for SslGPS at pH 7.8 and 25 °C. The refolding reaction begins in the unfolded, U state, initially misfolds to the I_{BP} intermediate state, partially unfolds to reach the I_A intermediate state whose conversion to the subsequent I_B intermediate state is rate-limiting. The final step is the conversion of I_B to the native state, N. The I_A and I_B kinetic intermediates correspond to the intermediate observed in equilibrium unfolding studies. The ordinate represents the free energy of each state in the folding reaction mechanism in kcal mol^{-1} . The abscissa represents the dependence of the difference in free energy between 2 states on the denaturant concentration and is proportional to the change in buried surface, referenced to the U state. The kinetic folding mechanism, illustrating the flow of the unfolded protein to the native conformation is shown beneath the reaction coordinate diagram.

Fig. 2

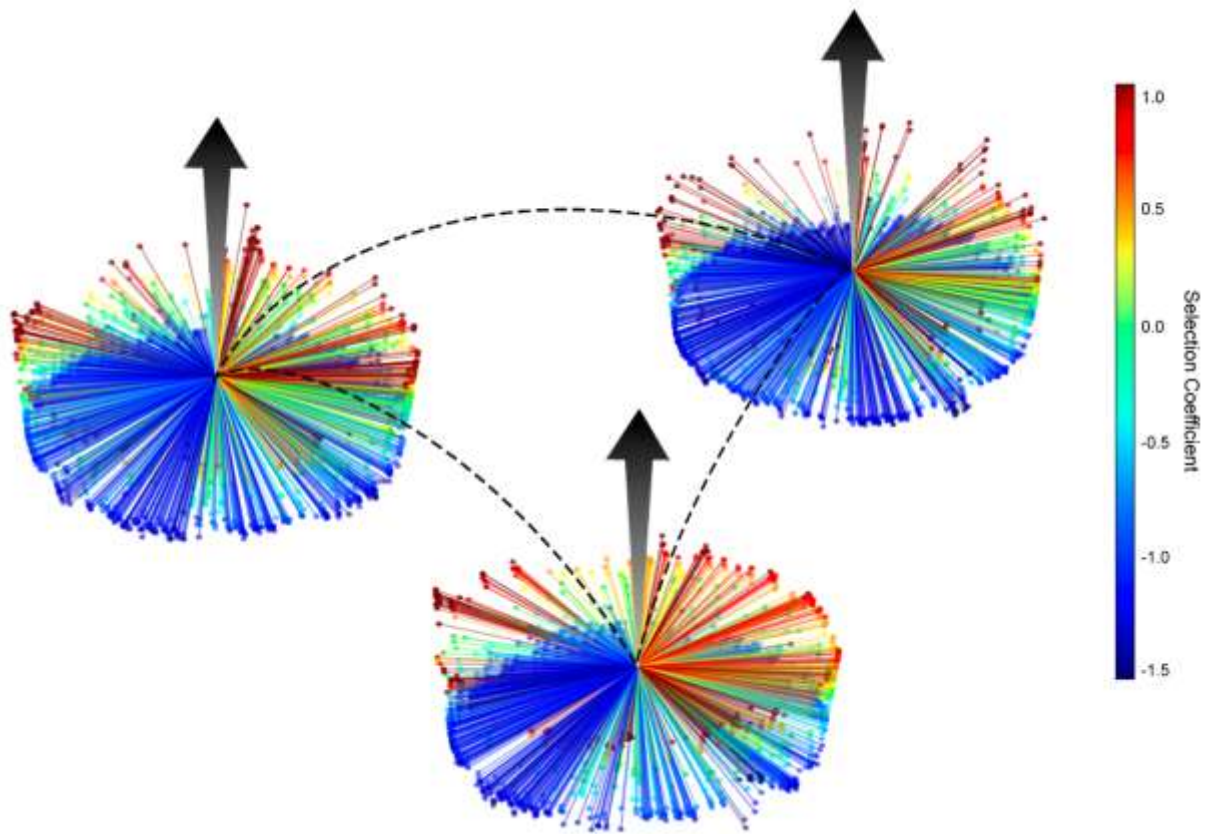


Fig. 2. Experimentally derived fitness landscapes mapped from point mutations represent single steps from WT sequence. Despite significant divergence of WT in sequence space, the fitness landscapes of IGPS orthologues remain correlated (dashed lines). Rather than traditional two-dimensional heatmaps, fitness values are displayed on a three-dimensional pinwheel, highlighting the wide range of possible fitness effects of a single sequence step. The profiles of the pinwheels are similar, indicating the correlation of fitness landscapes, even if WT sequences (centers of the wheels) are only ~40% identical and widely separated. Principal component analysis demonstrates a correlation between experimental fitness landscapes and amino-acid preferences in evolved sequences.