

Extension of a *de novo* TIM barrel with a rationally designed secondary structure element

Jonas Gregor Wiese¹  | Sooruban Shanmugaratnam^{1,2}  | Birte Höcker^{1,2} 

¹Max Planck Institute for Developmental Biology, Tübingen, Germany

²University of Bayreuth, Department for Biochemistry, Bayreuth, Germany

Correspondence

Birte Höcker, University of Bayreuth, Bayreuth, Germany.

Email: birte.hoecker@uni-bayreuth.de

Present address

Jonas Gregor Wiese, Technical University of Munich, Munich, Germany.

Abstract

The ability to construct novel enzymes is a major aim in *de novo* protein design. A popular enzyme fold for design attempts is the TIM barrel. This fold is a common topology for enzymes and can harbor many diverse reactions. The recent *de novo* design of a four-fold symmetric TIM barrel provides a well understood minimal scaffold for potential enzyme designs. Here we explore opportunities to extend and diversify this scaffold by adding a short *de novo* helix on top of the barrel. Due to the size of the protein, we developed a design pipeline based on computational *ab initio* folding that solves a less complex sub-problem focused around the helix and its vicinity and adapt it to the entire protein. We provide biochemical characterization and a high-resolution X-ray structure for one variant and compare it to our design model. The successful extension of this robust TIM-barrel scaffold opens opportunities to diversify it towards more pocket like arrangements and as such can be considered a building block for future design of binding or catalytic sites.

KEYWORDS

($\beta\alpha$)₈-barrel, *ab initio* folding, computational protein design, enzyme design, sTIM11, TIM barrel

1 | INTRODUCTION

The tight coupling of protein structure and function motivates the field of protein design to pursue the construction of modified or novel protein functions such as enzyme catalysis, signaling, binding and many more. The design of enzymes is of particular interest due to their applicability but also because it poses a thorough test for our understanding of what defines activity and selectivity. An intensively studied protein fold is the TIM or ($\beta\alpha$)₈-barrel. Enzymes with this fold are ubiquitously found among organisms and efficiently catalyze a wide variety of reactions.^{1,2} The fold is composed of an eight-fold repeat of $\beta\alpha$ -units, with eight parallel β -strands

forming a circular sheet in the core that is surrounded by the eight α -helices. While the “bottom side” of the barrel with its $\alpha\beta$ -loops provides stability to the barrel, the “upper” part including the $\beta\alpha$ -loops usually contains the catalytic function of TIM-barrel enzymes. The twisting central β -sheet provides a cavity at the opening of the barrel that is often used as substrate binding site in combination with extensions at the $\beta\alpha$ -loops. These extensions also often play an important role in positioning catalytic residues and shielding the catalytic site from solvent. Due to these exceptional characteristics of TIM barrels, they are highly attractive targets for enzyme design approaches. In fact, naturally occurring TIM barrel have already been modified to turn over non-natural

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

substrates, for example, carrying out a retro-aldol reaction.³ Computational approaches for the design of novel enzymes have been progressing in recent years and their combination with directed evolution has already proven very successful for some reactions.⁴

However, the ability to design an enzyme from scratch including its tailor-made scaffold is still a major aim that might be tackled with computational methods for structure prediction. Tools exist for structure prediction that can be used for its inverse, the protein design problem, to predict the structure of amino acid sequences. The widely used Rosetta molecular modeling suite provides tools for both problems. Its *ab initio* structure prediction algorithm generates fragment libraries for a target sequence from the protein structure database and searches conformational space using a Monte Carlo procedure.^{5,6} The protein design algorithm also uses Monte Carlo optimization to populate a given backbone with energetically optimal residues and uses structure evaluating energy functions for scoring.⁷ While the suite proved its utility in several CASP competitions,^{5,8} it was also the core design tool for the first *de novo* TIM barrel with four-fold symmetry.⁹ This design aimed at providing an idealized TIM barrel with minimal loops, formed from a four-fold repeated sequence. The design called sTIM11 is a monomeric and stable protein that with its X-ray structure provides a suitable starting point for further enzyme design approaches. With its compact structure it does however lack an extended surface or pocket as usually found in enzymes.

Functional sites in natural proteins are often built by loops that fit functional requirements, provide fine-tuned interactions or allow subtle dynamics. While rules have been defined for short loops that serve as connectors of secondary structure elements,¹⁰ the computational design of extended structured loops is still a major challenge.

Advances in loop structure prediction and challenges in applying them to computational loop design are important for the design of functional sites as recently reviewed by Kundert and Kortemme.¹¹

Here we took steps towards the goal of diversifying *de novo* TIM barrel by extending sTIM11 by a simple structural element, namely an additional helix that we inserted in a loop at the upper part of the barrel. We show with a high-resolution X-ray structure that the core protein is properly folded and the extension is formed while the protein's biophysical properties are maintained. This *de novo* helical extension can be considered as a structural building block for further design approaches towards binding and catalytic sites by providing additional structural mass at the minimal sTIM11 barrel surface.

2 | RESULTS

2.1 | Scaffold and target site identification

The minimal sTIM11 barrel provides a fairly flat top surface with seven adjoining $\beta\alpha$ -loops pointing into the region that typically comprises the active site of TIM-barrel enzymes.⁹ These loops were considered most suitable to introduce additional functionalities and extensions. Since the introduction of a secondary structural element might interfere with the folding process we considered carefully where to insert the extension. Since the two halves of sTIM11 can assemble into a TIM-barrel like dimer, we believed the loop in the middle of the protein sequence to be maximal tolerant for the structural modification and chose it as the target site for the helical extension (Figure 1). We further introduced a small

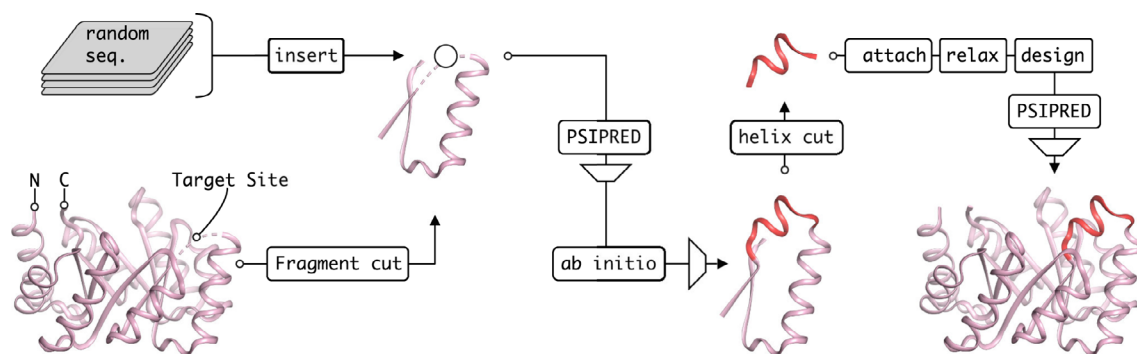


FIGURE 1 Design pipeline. A $\beta\alpha$ -fragment was extracted from sTIM11, whose structure could be predicted reliably and independently from its context. The target loop (circle) was substituted by random sequences of variable lengths that were predicted to fold into the desired helical structure by PSIPRED within the fragment context. Then the top scored sequences were folded using Rosetta and highly scored models that recover the barrel fragment with an additional helix on top were used further. The α -helix extension (red) was transferred back to the full barrel and adapted to the changed structural context by a Rosetta design step. Only those candidates were used for experimental investigation, that were still predicted to adapt the correct secondary structure according to another PSIPRED run

change to the sTIM11 scaffold by removing two cysteine residues that were originally introduced to form a disulfide bridge to improve barrel closure.⁹ However, this disulfide bond did not form, so we removed the two cysteines to generate a cysteine-free variant named sTIM11noCys whose characteristics and crystal structure are reported in Romero-Romero *et al.*¹²

2.2 | Design pipeline

Our aim was to extend sTIM11noCys by an arbitrary α -helix without any predefined backbone and thus the following design pipeline emerged (Figure 1). We wanted to design the model by structure prediction using the corresponding Rosetta *ab initio* method. However, the protein with its 184 amino acids is rather large and initial folding experiments on the entire barrel did not easily recover the structure. Further, since our design target was not a fully symmetric protein anymore, we decided to not use symmetry constraints such as in the design process of sTIM11. Instead, we simplified the problem by reducing the folding target to a smaller fragment of the barrel including the surroundings of the target site for insertion. A $\beta\alpha\beta$ -fragment of only 31 residues was found to fold into its tertiary structure independently from the full barrel context when using the Rosetta structure prediction method. We had considered the use of such a $\beta\alpha\beta$ -fragment as it is a super-secondary structural element that provides a hydrophobic core and as such has been regarded an evolutionary building block in the emergence of TIM barrel as well as other protein folds.^{13,14} In order to find insertions that were likely to fold into an α -helix, more than 150,000 random amino acid sequences of 8–19 residues were generated and inserted into the $\beta\alpha\beta$ -fragment substituting four amino acids (Asp91, Ala92, Thr93, Asp94) from the $\beta\alpha$ -loop. The resulting sequences of length 35–46 were used for secondary structure prediction using PSIPRED.¹⁵ The predictions were filtered for α -helical content at the insertion site and sorted by their prediction score to select the top 30 candidates. These sequences were inserted into the target barrel fragment, thus obtaining the final sequences for the actual Rosetta *ab initio* structure prediction experiments. 1,000 models were generated per candidate and multiple sequences were found to fold *in silico* into the desired tertiary structure, thereby recovering the barrel fragment with an additional α -helix at the target site. The structures were sorted by root-mean-square deviation (RMSD) when superimposing the $\beta\alpha\beta$ -fragment onto the parental barrel and for the top 12 candidates, the α -helix was manually transferred to the full barrel. An additional Rosetta relax step was

performed to smooth introduced structural deviations from the manual intervention.

Since the promising tertiary structure had been found in the context of the fragment and not in the context of the entire barrel, the residues of the helix had to be optimized locally to fit with adjacent residues. Therefore, about 50% of the residues of the α -helix extension was allowed to mutate freely in a subsequent Rosetta design step, generating 750 designs per candidate. Once again, the secondary structure of the sequences was predicted by PSIPRED in the context of the barrel fragment and filtered to raise the chance of obtaining the desired secondary structure. Four low energy candidates were chosen manually based on hydrophobic packing and predicted polar contacts for experimental characterization. While all candidates expressed and indicated properly folded proteins (Figure S1), only one candidate, sTIM11_helix3, yielded diffracting crystals for which an X-ray structure could be solved.

2.3 | Biochemical evaluation of sTIM11_helix3

After cloning the designed sequence, the corresponding protein sTIM11_helix3 was expressed and purified. Subsequent analytical size exclusion chromatography of the purified protein revealed a single peak, corresponding to a homogeneous monomeric species (Figure 2(a)). Circular dichroism spectra were compatible with folded protein containing a mixed α -helical and β -sheet content comparable to sTIM11noCys (Figure 2(b)). Moreover, the design showed a similar thermal stability to sTIM11noCys with a melting temperature of 66°C (Figure 2(c)).

2.4 | Structural evaluation by X-ray crystallography

The protein was crystallized and diffraction data collected and processed to a resolution of 1.58 Å (Table 1). The X-ray structure was solved by molecular replacement with four copies of a quarter barrel of the four-fold symmetric sTIM11 (PDB: 5BVL). The B-factor distribution of the final structure revealed highly reliable areas for the core of the TIM barrel, which is clearly recovered in the structure. The introduced helix showed elevated B-factors, making it more difficult to model. Due to the increased flexibility in the region of interest we refined our structure using a feature enhanced map (FEM), reducing the level of noise and model bias.¹⁶ The obtained map allowed us to reliably build the entire structure, making

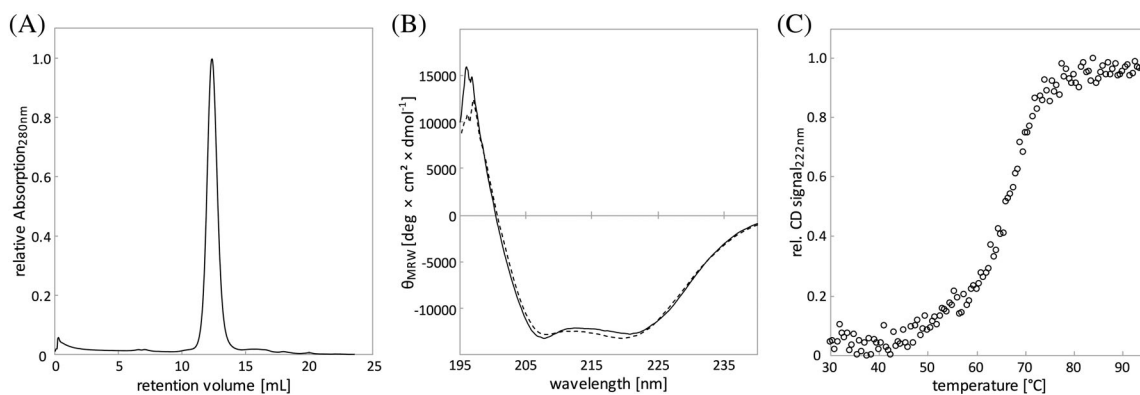


FIGURE 2 Biochemical Characterization of sTIM11_helix3. (A) Analytical size-exclusion chromatography, (B) circular dichroism measurements before heating (solid line) and after cooling down (dashed line) and (C) thermal melting upon heating

the introduced helix clearly observable so it can be compared to the theoretical design model (Figure 3).

2.5 | Comparison of design model and X-ray structure

Superposition of the design model and the X-ray structure show that most parts of the structures are identical. In the core TIM barrel, only the polypeptide chain subsequent to the target site of extension showed structural deviations (Figure 4). But also, the helical extension showed a small but significant deviation. While a helical structure was formed, it is wound slightly different than expected and forms a 3^{10} helix instead of a classical α -helix. In addition, its position is slightly twisted and it packs more closely and flat onto the barrel surface compared to the design model, which projected the helix to be more exposed. In fact, the design model showed Met94 and Ala95 to dock into a space at the barrel surface without any major structural changes to the main barrel. Interestingly, the experimental structure revealed a different hydrophobic packing and showed Met94 to rather enter the barrel by inserting its side chain in between the β -sheet core and the outer barrel α -helix (Figure 5). Due to this, several polar contacts that had been predicted between the first residues of the inserted helix and its surrounding could not form, but a polar contact between Ser90 and Asp71 is established instead. The following loop that connects the extension with the subsequent α -helix is initiated by Pro98 as designed. But the length of the loop turns out to be longer and more flexible than expected due to a slight unravelling of the barrel helix. In the X-ray structure the loop is solvent exposed and numerous water molecules make polar contacts with its residues. The slightly shortened TIM-barrel helix subsequent to the extension also deviates from its usual

position and orientation. It is shifted downwards. Thus, the *de novo* TIM barrel accommodates the newly inserted 3^{10} helix by a rearrangement of the packing of the barrel helix. The snug interaction of the inserted short helix provides a likely explanation for the observed structural deviations of the TIM-barrel core helix.

3 | DISCUSSION

In this work we provide a structural extension to the *de novo* TIM barrel sTIM11. Instead of inserting a full domain into the barrel as recently reported¹⁷ or attaching a natural occurring fragment to the barrel, we decided to introduce a short additional α -helix connected by short loops. This required careful integration to the residue environment of the barrel by comprising various computational tools for sequence screening and optimization. By not constraining the design target at the beginning to any fixed backbone, we could not use residue optimization methods initially but performed the first design step by using structure prediction methods instead. We found it to be sufficient to work with just a structural fragment of the barrel in order to find the desired structural extension in a randomly sampled sequence space. After adapting the variants to the adjacent residues in their barrel environment using Rosetta design, well-folded proteins were obtained and the successful extension of one design could be confirmed experimentally.

The experimentally derived X-ray structure revealed unexpected deviations from the design model. Methods from the Rosetta toolbox can approximate the structure and energies of proteins while it does not guarantee the experimental structure to have the exact shape. However, our design pipeline itself might have introduced structural deviations at various steps. We first worked with a barrel fragment in order to find our initial α -helical

TABLE 1 Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses

| | |
|---|-----------------------------------|
| Data collection | |
| Wavelength (Å) | 1.000030 |
| Resolution range (Å) | 47.23–1.58 (1.64–1.58) |
| Space group | P 4 ₁ 2 ₁ 2 |
| Unit cell | |
| a, b, c (Å) | 50.5 50.5133.3 |
| α, β, γ (°) | 90 90 90 |
| Total reflections | 197,971 (8063) |
| Unique reflections | 24,291 (1874) |
| Multiplicity | 8.1 (3.8) |
| Completeness (%) | 96.9 (78.4) |
| Mean I/sigma (I) | 12.75 (0.40) |
| Wilson B-factor (Å ²) | 25.4 |
| No. of molecules per a. u. | 1 |
| Matthews coefficient | 1.81 |
| R _{merge} | 0.081 (2.161) |
| R _{meas} | 0.086 (2.484) |
| R _{pim} | 0.029 (1.178) |
| CC _{1/2} | 0.999 (0.321) |
| CC* | 1.000 (0.697) |
| Refinement | |
| Reflections used in refinement | 23,865 (1873) |
| Reflections used for R _{free} | 1,191 (93) |
| R _{work} | 0.200 (0.469) |
| R _{free} | 0.240 (0.482) |
| CC _{work} | 0.966 (0.427) |
| CC _{free} | 0.948 (0.290) |
| Number of non-hydrogen atoms | |
| macromolecules | 1,660 |
| solvent | 104 |
| Protein residues | 190 |
| RMS bond lengths (Å) | 0.007 |
| RMS bond angles (°) | 1.100 |
| Ramachandran favored (%) | 95.7 |
| Ramachandran allowed (%) | 3.7 |
| Ramachandran outliers (%) | 0.5 |
| Rotamer outliers (%) | 1.2 |
| Clashscore | 2.42 |
| Average B-factor (Å²) | |
| macromolecules | 40.8 |
| solvent | 44.5 |
| Number of TLS groups | 1 |

extension and then transferred the helix manually to the full barrel. This step might have introduced deviations that were not smoothed out entirely by a Rosetta relax step. Moreover, the helix was found in the context of the barrel fragment and was not optimized for the full barrel, which is why we added the additional design step of adapting the helix to the full barrel. However, in this step only small backbone deviations were allowed. The experimentally observed alternative hydrophobic packing, where the Met94 side-chain is inserted into the barrel core was probably out of the scope of allowed backbone deviations and thus was not predicted. Moreover, the length of the α-helical extension was optimized in the context of the barrel fragment but not allowed to change in the full barrel context.

Nonetheless, we were able to insert a new element into the TIM barrel without disrupting its structural integrity or stability. The newly inserted 3¹⁰ helix appears to fit even more snugly on top of the barrel, and the tight packing of the Met94 side chain with the hydrophobic core of the barrel might have caused the observed structural adjustments that even trigger the subsequent barrel helix to take an alternate conformation illustrating the plasticity of this *de novo* TIM barrel.

This work shows how idealized protein designs can be diversified and provides a first building block for further designs towards sTIM11-based enzymes. Due to the four-fold symmetry of sTIM11, we assume that this extension can even be transferred to any of the related βα-loops. It therefore provides a flexible extension that might as well be combined with other structural modifications helping to form binding or catalytic sites.

4 | METHODS

4.1 | Computational methods

In the first steps of the pipeline, we used only a 31-residue long βαβ-fragment from the barrel, and replaced the four-residue long βα-loop by randomly generated sequences that had lengths of 8–19 residues. So, the assembled sequences ranged from 35–46 residues and their secondary structure was predicted by PSIPRED (version 4.0) on local machines. The predicted secondary structures were filtered by motif filters, realized by regular expressions (Table S2). Sequences that were predicted to form helices of length 4–8 surrounded by 2–3 flexible residues were accepted.

The Rosetta molecular modeling suite (weekly release, first quarter 2016) was used on local machines

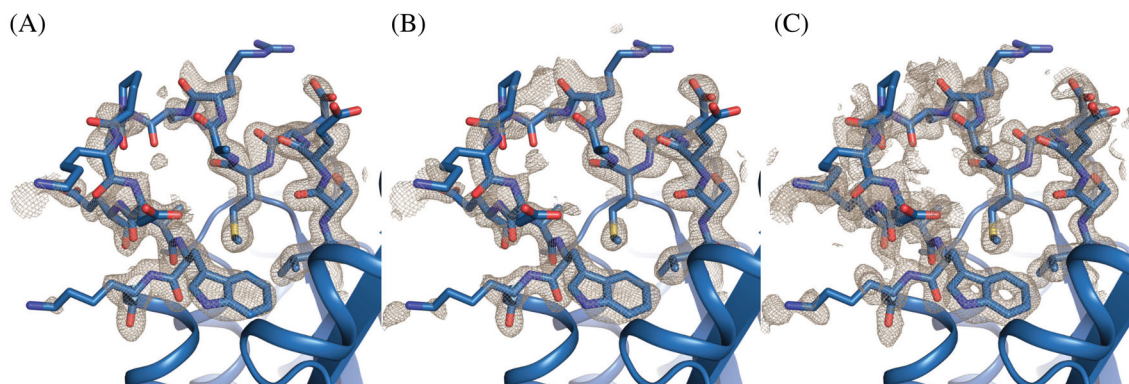


FIGURE 3 Comparison of density maps for the designed helix region. (A) 2mFo-DFc map after the last refinement. (B) 2mFo-DFc map after calculation of a composite omit map. (C) 2mFo-DFc map after calculation of a feature enhanced map. All maps are contoured at 1σ . Maps are cut off around 1.6 Å distance to the residues shown as sticks. The X-ray structure is shown as blue cartoon, while the sidechains in and flanking the helix region are depicted as sticks

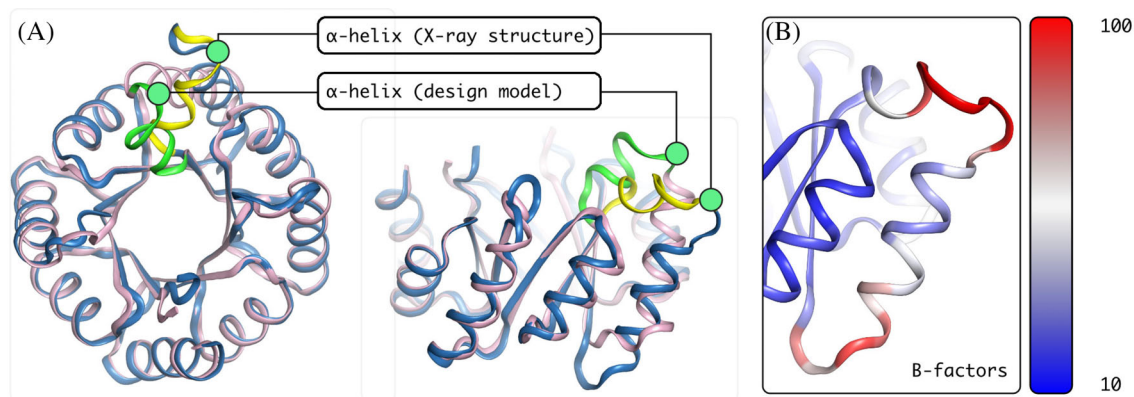


FIGURE 4 Characterization of the X-ray structure in comparison with the design model. The superposition of the X-ray structure (blue) with the design model (purple) in a top-down view (A, left) reveals a close to identical fit. The region of extension shows some structural difference, the inserted α -helix in the X-ray structure (yellow) is twisted in comparison to the prediction (green). The side view (A, right) illustrates that the α -helix is in a different position than predicted and that the following helix within the barrel spatially deviates as well. The B-factor coloring (B) points at the apparent high flexibility in the region of interest

for *ab initio* folding of the extension containing barrel fragment by using the *ab initio* relax method (Table S3-S5). The Rosetta relax method was used to smooth structural deviations from manual intervention, when transferring the folded helix extension to the entire barrel, which was performed using PyMOL (version 1.8). Finally, the Rosetta enzyme design method was used to adjust the helix residues to adjacent barrel residues.

4.2 | Cloning methods

In order to insert the gene fragment encoding the extension, we generated two overlapping fragments in a first PCR reaction. Both samples were loaded on a 1% agarose gel and after electrophoresis, appropriate bands were excised and purified

with the PCR clean-up kit (Qiagen). These two purified DNA fragments served as templates for the second PCR reaction, where both fragments were combined at their overlapping region. The PCR sample was loaded on a 1% agarose gel and after electrophoresis, the DNA band with the correct size was excised and purified as before.

Purified DNA fragment and vector pET21b(+) were then double-digested using *NdeI* and *XhoI* (Fermentas) at 37°C for 2 h. After digestion, the cut vector was purified using the PCR clean-up kit (Qiagen), while the DNA fragment was purified and in parallel concentrated using the DNA Clean & Concentrator Kit (Zymo Research). Vector and fragment were ligated in an overnight reaction at 4°C using T4 DNA Ligase (New England Biolabs). After transformation, successful clones were verified by sequencing.

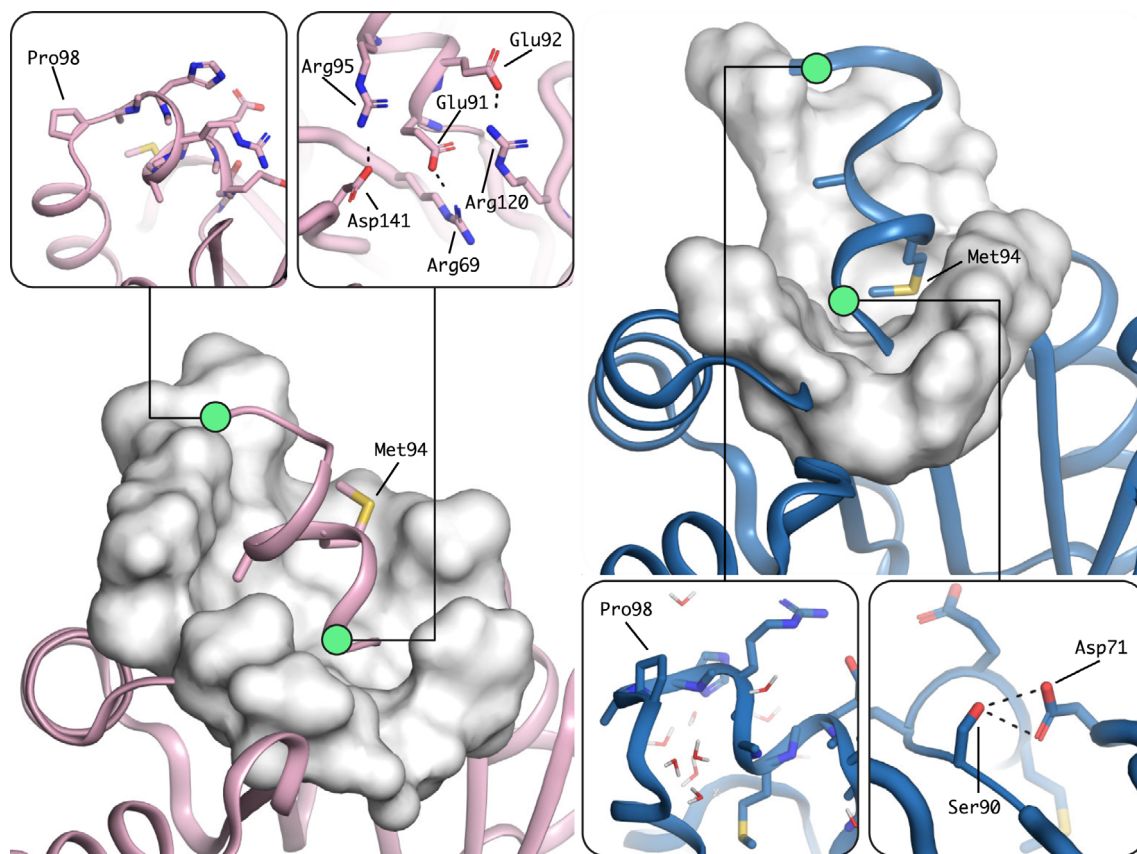


FIGURE 5 Features of the helix in the design model (left) and in the X-ray structure (right). Hydrophobic residue interactions differ (large scale images). In the design hydrophobic residues are packing flat onto the top of the barrel, while in the X-ray structure Met94 rather pierces the top of the barrel in between β -sheet and α -helices. The helix breaking Pro98 (small scale images, left) resulted in an enlarged loop compared to the design which is surrounded by water in the X-ray structure and polar contacts (small scale images, right) are less prominent and differently formed in the X-ray structure

4.3 | Protein expression and purification

Transformants of *E. coli* BL21 were grown in LB media at 37°C in a shaker until the culture reached an OD₆₀₀ of 0.6. Protein expression was induced by addition of IPTG (1 mM). The culture was incubated at 37°C for another 4 h and then harvested by centrifugation. After washing and resuspension in buffer A (50 mM potassium phosphate buffer pH 8.0, 150 mM NaCl, 20 mM Imidazole), cells were lysed by pulsed sonication. Soluble and insoluble components were separated by centrifugation (20,000 g).

4.4 | Biochemical characterization

The biochemical characterization was done in potassium phosphate buffer (50 mM pH 8.0, 150 mM NaCl) using a sample concentration of 0.2 mg ml⁻¹. For circular dichroism (CD), the sample was placed in a 1 mm cuvette and the spectrum was recorded from 240–195 nm on a Spectropolarimeter (Jasco J-810). Measured values were normalized to the molar ellipticity per amino acid.

The melting temperature was determined using the same sample. CD signal was tracked at 222 nm, while the instrument heated up or cooled down respectively from 30°C to 95°C with a rate of 1 K min⁻¹. Measured values were transformed to display the fraction of folded and unfolded protein.

Analytical size-exclusion chromatography was done with a Superdex 75 10/300 GL (GE Healthcare Life Sciences) on an ÄKTApurifier system (GE Healthcare Life Sciences). The sample (0.5 ml) was loaded on the buffer equilibrated column and the absorption was tracked at 280 nm.

4.5 | Structure determination

Crystallization conditions were screened using commercially available sparse-matrix screens (Qiagen) and a protein concentration of 4.7 mg ml⁻¹. Sitting drops (1 μ l) in a ratio of 1:1 (protein:screening solution) were pipetted into a 3-well Intelli-Plate (Art Robbins Instruments) using a Honeybee 961 (Genomic solutions). Plates were sealed with clear tape and incubated at 20°C. Crystals

were obtained in a condition containing 0.2 M Lithium sulfate, 0.1 M Tris pH 8.5, and 30% PEG 4000. Crystals were mounted and flash-cooled in liquid nitrogen. Diffraction images were collected at 100 K on a Pilatus 2 M-F at the beamline X06DA (PX III, Swiss Light Source, PSI). 1,400 images were collected using an oscillation of 0.1° per image.

Collected data were processed using XDSAPP 2.0.¹⁸ Phases were solved by molecular replacement using PhaserMR with a quarter of sTIM11 (PDB: 5BVL) as search model.¹⁹ Refinements were done using Phenix and manual model building was performed with Coot.^{20,21}

ACKNOWLEDGMENT

We acknowledge financial support and allocation of beamtime by HZB and thank the beamline staff at BESSY for assistance. We further thank Kaspar Feldmeier for detailed explanations about the design and structure of sTIM11, and Horst Lechner, André C. Stiel and Sergio Romero-Romero for helpful comments on the manuscript.

AUTHOR CONTRIBUTIONS

Jonas Gregor Wiese: Conceptualization; data curation; investigation; writing-original draft; writing-review & editing. **Sooruban Shanmugaratnam:** Data curation; investigation; writing-original draft; writing-review & editing. **Birte Höcker:** Conceptualization; funding acquisition; project administration; writing-original draft; writing-review & editing.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY

Coordinate and structure files have been deposited to the Protein Data Bank (PDB) with accession code 7A8S.

ORCID

Jonas Gregor Wiese  <https://orcid.org/0000-0002-8925-2823>

Sooruban Shanmugaratnam  <https://orcid.org/0000-0002-2614-6046>

Birte Höcker  <https://orcid.org/0000-0002-8250-9462>

REFERENCES

1. Wierenga RK. The TIM-barrel fold: A versatile framework for efficient enzymes. *FEBS Lett.* 2001;492:193–198.
2. Sterner R, Höcker B. Catalytic versatility, stability, and evolution of the ($\beta\alpha$)8-barrel enzyme fold. *Chem Rev.* 2005;105:4038–4055.
3. Jiang L, Althoff EA, Clemente FR, et al. De novo computational design of retro-aldol enzymes. *Science.* 2008;319:1387–1391.
4. Lechner H, Ferruz N, Höcker B. Strategies for designing non-natural enzymes and binders. *Curr Opin Chem Biol.* 2018;47:67–76.
5. Bonneau R, Tsai J, Ruczinski I, et al. Rosetta in CASP4: Progress in ab initio protein structure prediction. *Proteins Struct Funct Genet.* 2001;45:119–126.
6. Simons KT, Kooperberg C, Huang E, Baker D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J Mol Biol.* 1997;268:209–225.
7. Kuhlman B, Baker D. Native protein sequences are close to optimal for their structures. *Proc Natl Acad Sci U S A.* 2000;97:10383–10388.
8. Raman S, Vernon R, Thompson J, et al. Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins Struct Funct Bioinforma.* 2009;77:89–99.
9. Huang PS, Feldmeier K, Parmeggiani F, Velasco DF, Hocker B, Baker D. De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat Chem Biol.* 2016;12:29–34.
10. Koga N, Tatsumi-Koga R, Liu G, et al. Principles for designing ideal protein structures. *Nature.* 2012;491:222–227.
11. Kundert K, Kortemme T. Computational design of structured loops for new protein functions. *Biol Chem.* 2019;400:275–288.
12. Romero-Romero S, Costas M, Manzano D-AS, et al. Epistasis on the stability landscape of de novo TIM barrels explored by a modular design approach. *bioRxiv.* 2020;319103.
13. Söding J, Lupas AN. More than the sum of their parts: On the evolution of proteins from peptides. *Bioessays.* 2003;25:837–846.
14. Romero-Romero S, Kordes S, Michel F, Höcker B. Evolution, folding, and design of TIM barrels and related proteins. *Curr Opin Struct Biol.* 2021;68:94–104.
15. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol.* 1999;292:195–202.
16. Afonine PV, Moriarty NW, Mustyakimov M, et al. FEM: Feature-enhanced map. *Acta Cryst D.* 2015;71:646–666.
17. Caldwell SJ, Haydon IC, Piperidou N, et al. Tight and specific lanthanide binding in a de novo TIM barrel with a large internal cavity designed by symmetric domain fusion. *Proc Natl Acad Sci U S A.* 2020;117:30362–30369.
18. Sparta KM, Krug M, Heinemann U, Mueller U, Weiss MS. XDSAPP2.0. *J Appl Cryst.* 2016;49:1085–1092.
19. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Cryst.* 2007;40:658–674.
20. Adams PD, Afonine PV, Bunkóczi G, et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D.* 2010;66:213–221.
21. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Cryst D.* 2010;66:486–501.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Wiese JG, Shanmugaratnam S, Höcker B. Extension of a *de novo* TIM barrel with a rationally designed secondary structure element. *Protein Science.* 2021; 30:982–989. <https://doi.org/10.1002/pro.4064>