βα-Hairpin Clamps Brace βαβ Modules and Can Make Substantive Contributions to the Stability of TIM Barrel Proteins

Xiaoyan Yang*, Sagar V. Kathuria*, Ramakrishna Vadrevu**, C. Robert Matthews*

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts

Abstract
Non-local hydrogen bonding interactions between main chain amide hydrogen atoms and polar side chain acceptors that bracket consecutive βα or αβ elements of secondary structure in βTS from E. coli, a TIM barrel protein, have previously been found to contribute 4–6 kcal mol$^{-1}$ to the stability of the native conformation. Experimental analysis of similar βα-hairpin clamps in a homologous pair of TIM barrel proteins of low sequence identity, IGPS from S. solfataricus and E. coli, reveals that this dramatic enhancement of stability is not unique to βTS. A survey of 71 TIM barrel proteins demonstrates a 4-fold symmetry for the placement of βα-hairpin clamps, bracing the fundamental βαβ building block and defining its register in the (βαβ)$_8$ motif. The preferred sequences and locations of βα-hairpin clamps will enhance structure prediction algorithms and provide a strategy for engineering stability in TIM barrel proteins.

Introduction
The (βα)$_8$, TIM barrel is one of the most common folds in biology, supporting a myriad of catalytic functions essential to life [1]. Experimental [2] and bioinformatics [1,3,4] analyses of TIM barrel proteins have led to the conclusion that a pair of adjacent parallel β-strands and the intervening anti-parallel α-helix, i.e., the βαβ module, serve as the minimal unit of stability. Gene duplication of this elemental βαβ building block into higher-order structures has been suggested to result in several common βα-repeats, including the TIM barrel, Rossmann, flavodoxin and leucine-rich folds [3]. The interactions stabilizing this secondary structure include: (1) main chain-main chain (MC–MC) hydrogen bonds (H-bonds) between the β-strands, (2) intra-helical MC–MC H-bonds, (3) hydrophobic interactions between the side chains (SC) protruding from the β-strands and the α-helix, (4) side chain-side chain (SC–SC) H-bonds and salt bridges, (5) dipole-dipole interactions between the α-helix and the pair of β-strands on which it docks [5] and (6) main chain-side chain (MC–SC) H-bonds [6–8]. The surprising role of a subset of non-local MC–SC H-bond interactions in structure and stability is the subject of this communication.

A majority of MC–SC interactions in proteins are local in sequence, usually within six residues [7,8], and are often involved in capping either the N- or the C-termini of α-helices [9,10]. Mutational analysis has shown that these non-covalent interactions usually contribute modestly, typically in the range of 1–2 kcal mol$^{-1}$, to the stability of their resident proteins [11–14]. In contrast, the removal of three non-local MC–SC H-bond interactions each reduce the stability of the alpha subunit of tryptophan synthase (βTS), a TIM barrel protein, by 4–6 kcal mol$^{-1}$, and disrupt the complete formation of the TIM barrel motif [6]. These three interactions in βTS, between MC amide H-bond donors and SC H-bond acceptors, connect the N-terminus of one element of secondary structure, either β-strand or α-helix, to the C-terminus of the subsequent element of structure, either α-helix or β-strand, respectively. These non-local MC-SC interactions were designated as βα-hairpin clamps and βα-hairpin clamps, respectively [6]. The significant contribution to structure and stability by three such clamps in βTS [6] raises the possibility that potent βα- and αβ-hairpin clamps may be an important general feature of TIM barrel proteins.

A two-pronged approach was taken to probe the significance of βα-hairpin clamps in TIM barrel proteins. First, mutational analysis of two representative TIM barrel proteins, indole-3-glycerol phosphate synthase (IGPS) from S. solfataricus (sIGPS) and E. coli (cIGPS), shows that a subset of their βα-hairpin clamps make significant contributions to protein stability. Second, a survey of 71 TIM barrel proteins [15] explored the frequency, location and sequence preferences of all βα-hairpin clamps. The observed preferences for location and sequence for the βα-hairpin clamps and their contribution to the structure and stability of TIM barrel proteins suggest that the recognition of these interactions can enhance protein structure prediction algorithms and provide a strategy for engineering stability in TIM barrel proteins.
Results

Experimental analysis of βz-hairpin clamp interactions in two TIM barrel proteins

The generality of the potent hairpin clamps found in αTS was tested by mutational analysis of βz-hairpin clamps in two homologous TIM barrel proteins with low sequence identity (<30%) to αTS and to each other, sIGPS (Figure 1A) and eIGPS (Figure 1B), each contain three βz-hairpin clamps (Figure 1C and 1D), some of which are conserved in location with those in αTS and others between sIGPS and eIGPS. Figure 1 displays the distances between the donor and acceptor atoms of the βz-hairpin clamps interactions observed in sIGPS (Figure 1C) and eIGPS (Figure 1D).

The solvent-exposed surface area of the H-bond acceptor atoms ranges from 0.2 Å² (0%) to 11.8 Å² (25%), while the MC H-bond donor amide is typically completely excluded from solvent (Figure 1C and 1D). The βz1 clamp is observed in αTS and eIGPS (αTS-β1α1-F19NH→O62D46, eIGPS-β1α1-F50NH→O62S82), the βz2 clamp only appears in sIGPS (sIGPS-β2z2-S104NH→O62E74), the βz3 clamp is observed in all three proteins (αTS-β3α3-I97NH→O62D124, sIGPS-β3α3-I107NH→O62D126 and eIGPS-β3α3-I111NH→O62D132), and the βz7 clamp is observed in sIGPS and eIGPS (sIGPS-β7α7-K207NH→O62N228 and eIGPS-β7α7-V211NH→O62N231).

Perturbation of the secondary and tertiary structure by clamp deletion in sIGPS and eIGPS. The contribution of each βz-clamp interaction to the structure of the TIM barrel proteins, sIGPS and eIGPS, was assessed by replacing the H-bond acceptor SC with alanine and monitoring the effects on the secondary and tertiary structure by far-UV and near-UV circular dichroism (CD) spectroscopy. The far-UV CD spectra for the wild-type (WT) and clamp-deletion variants of sIGPS (sIGPS-WT, sIGPS-Δβ2z2-E74A, sIGPS-Δβ3α3-D128A and sIGPS-Δβ7α7-N228A) and eIGPS (eIGPS-WT, eIGPS-Δβ1α1-S82A, eIGPS-Δβ3α3-D132A and eIGPS-Δβ7α7-N231A) are shown in Figure 2A and 2B, and the near-UV CD spectra are shown in Figure 2C and 2D. Relatively small changes in the far-UV and near-UV CD spectra are observed for sIGPS-Δβ2z2-D128A, eIGPS-Δβ1α1-S82A and eIGPS-Δβ3α3-D132A compared to their respective WT sequences. However, the significant changes in the near-UV CD spectra for the sIGPS-Δβ2z2-E74A, sIGPS-Δβ7α7-N228A and eIGPS-Δβ7α7-N231A variants imply that the deletion of the βz7 clamp in both proteins and the βz2 clamp in sIGPS result in altered aromatic side chain packing.

Figure 1. Ribbon diagrams of sIGPS (A) and eIGPS (B) highlighting the βz-hairpin clamps. (C) and (D) display the intervening elements of secondary structures between the residues forming the clamps for sIGPS: sIGPS-β2z2-S104NH→O62E74, sIGPS-β3α3-I107NH→O62D126, and sIGPS-β7α7-K207NH→O62N228 and for eIGPS: eIGPS-β1α1-F50NH→O62S82; eIGPS-β3α3-I111NH→O62D132; and eIGPS-β7α7-V211NH→O62N231. The SCs involved in the clamp interactions are highlighted with the H-bond donor and acceptor atoms shown in blue and red, respectively. The distances between the donor and acceptor atoms are indicated. The solvent exposed surface areas of the H-bond donor and acceptor atoms is shown in parenthesis. The H-bonds and their corresponding distances were determined by using the program HBPLUS [45]. The structures were generated using PyMOL v 0.99 [46], and the PDB codes are 2C3Z for sIGPS [21] and 1PII for eIGPS [22].

doi:10.1371/journal.pone.0007179.g001
Perturbation of stability by clamp deletion in sIGPS and eIGPS. The effect of β2-hairpin clamp deletion on the stability of sIGPS and eIGPS was determined by urea denaturation. As for αTS [16], both sIGPS [17] and eIGPS [18] unfold via a highly populated intermediate, and their unfolding titration curves are well described by a three-state model, $N \rightleftharpoons I \rightleftharpoons U$. With the exception of eIGPS-$\Delta \beta 7\alpha 7$-N231A, the urea-induced unfolding transition for each of the remaining five clamp-deletion variants is also well-described by this three-state model (Figure 3A and 3B).

Because a distinct transition between the native state (N) and the intermediate state (I) is not observed during the urea induced denaturation of eIGPS-$\Delta \beta 7\alpha 7$-N231A (Figure 3B), kinetic unfolding experiments were performed to verify the existence of I and measure the free energy difference between N and I [6]. The presence of I in eIGPS-$\Delta \beta 7\alpha 7$-N231A is verified by the observation of a slow kinetic unfolding phase, whose relaxation times decrease with increasing final denaturant concentration [19], when eIGPS is subjected to an unfolding jump from 0 to 3 M urea where I is not observed during the urea induced denaturation of eIGPS-$\Delta \beta 7\alpha 7$-N231A (Figure 3B), kinetic unfolding experiments were performed to verify the existence of I and measure the free energy difference between N and I [6].

The presence of I in eIGPS-$\Delta \beta 7\alpha 7$-N231A is verified by the observation of a slow kinetic unfolding phase, whose relaxation times decrease with increasing final denaturant concentration [19], when eIGPS is subjected to an unfolding jump from 0 to 3 M urea where I is expected to be highly populated. Because the amplitude of the unfolding phase is proportional to the population of N from which the reaction initiates, the decrease in the amplitude as a function of increasing initial urea concentrations (Figure S1) can be fit to a two-state model, $N \rightleftharpoons I$, to extract the stability, $\Delta G_{NI}$, and the urea dependence of the stability, $m_{NI}$ (Dataset S1). These parameters are used to fit the CD unfolding transition for eIGPS-$\Delta \beta 7\alpha 7$-N231A (Figure 3B) and to extract $\Delta G_{IU}$ and $m_{IU}$ (Methods S1).

The stabilities of N and I for the clamp-deletion variants and the WT parent sequences are illustrated graphically in Figure 3C and 3D for sIGPS and eIGPS, respectively. The free energy differences between N and I, $\Delta G_{NI}$, and between I and the unfolded state, U, $\Delta G_{IU}$, as well as the $m$-values, are tabulated in Table 1. The deletion of the $\beta 2\alpha 2$ clamp in sIGPS, sIGPS-$\Delta \beta 2\alpha 2$-E74A, only reduces the stability of N by 1.08 kcal mol$^{-1}$, and the deletion of the $\beta 3\alpha 3$ clamp, sIGPS-$\Delta \beta 3\alpha 3$-D128A, has no significant effect on its stability. By striking contrast, the elimination of the $\beta 7\alpha 7$ clamp, sIGPS-$\Delta \beta 7\alpha 7$-N228A, reduces the stability of N by 4.30 kcal mol$^{-1}$. Consistent with the absence of these clamps in I for all of these variants, the free energy differences between I and U for the clamp-deletion variants are comparable to the corresponding value for sIGPS-WT (Figure 3C and Table 1). Similar results are obtained for eIGPS. Only eIGPS-$\Delta \beta 7\alpha 7$-N231A decreases the stability of N significantly, $\Delta G_{NI} = 4.32$ kcal mol$^{-1}$. eIGPS-$\Delta \beta 1\alpha 1$-S82A and eIGPS-$\Delta \beta 3\alpha 3$-D132A have no significant effect on the stability of N, and none of the clamp-deletion variants perturb the stability of I relative to U (Figure 3D).

Figure 2. Ellipticity of wild-type and clamp-deletion variants of sIGPS and eIGPS. Far-UV (a, b) and near-UV (c, d) CD spectra of sIGPS (a) and (c): sIGPS-WT (--), sIGPS-$\Delta \beta 2\alpha 2$-E74A (••••), sIGPS-$\Delta \beta 3\alpha 3$-D128A (–), and sIGPS-$\Delta \beta 7\alpha 7$-N228A (•). and eIGPS (b) and (d): eIGPS-WT (--), eIGPS-$\Delta \beta 1\alpha 1$-S82A (••••), eIGPS-$\Delta \beta 3\alpha 3$-D132A (–), and eIGPS-$\Delta \beta 7\alpha 7$-N231A (•). Buffer conditions: 10 mM potassium phosphate, 0.2 mM K$_2$EDTA, 1 mM DME, pH 7.8 for sIGPS and pH 7.0 for eIGPS at 25°C. doi:10.1371/journal.pone.0007179.g002
and Table 1). Thus, while the elimination of either the β1z1, β2z2 or β3z3 clamps has only marginal effects on sIGPS and eIGPS, the β7z7 clamps in both proteins contribute significantly to both the structure and the stability of the native states for their resident TIM barrel protein.

**Survey of βα-hairpin clamps in the TIM barrel proteins**

The observation that βα-hairpin clamps can have a significant effect on structure and stability in three TIM barrel proteins motivated a survey of the prevalence of such non-local MC-SC H-bonds in the TIM barrel fold. This analysis was carried out for a structural database of 71 TIM barrel domains, previously reported bonds in the TIM barrel fold. This analysis was carried out for a survey of the prevalence of such non-local MC-SC H-bonds 

Forming H-bonds with the MC amide hydrogen of isoleucine, leucine and valine residues. Inspection of the location of the donor and acceptor residues in the β-strands reveals that every βα-hairpin clamp secures the N-terminus of one β-strand to the loop preceding the subsequent β-strand in the barrel. The locations of the entire group of 131 MC_{NH} → SC βα-hairpin clamps are displayed in Figure 4A, with each clamp interaction represented as a bridge across two adjacent β-strands. A very strong preference (77%) is seen for β1z1, β3z3, β5z5 and β7z7 clamps, where the SC acceptor is C-terminal to the MC H-bond donor. With the exception of 13 β8z8 clamps, the paucity of β2z2, β4z4 and β6z6 clamps is distinct from their odd β-strand counterparts. The relatively large number of clamps for the β8β1 interface may reflect the necessity for securing the N- and C-terminal β-strands. Far fewer βα-hairpin clamps, in which the SC acceptor is N-terminal to the MC H-bond donor, are observed. Highlighting the significance of this distribution pattern, the 55 Leu and Val [1/L/V] MC → SC Asp (D) sub-group of βα-hairpin clamps always have their MC H-bond donor I/L/V located in the odd-numbered stands, β1, β3, β5 or β7, and their SC acceptor, D, is always located before the succeeding even-numbered β-strands, β2, β4, β6 and β8. There is also a strong preference for the I/L/V residue to occupy the 2<sup>nd</sup> position in the odd-numbered β-strand and for the D residue to occupy the position immediately preceding the even-numbered β-strand (Figure 4B). This positional

**Figure 3. Stability perturbation of sIGPS and eIGPS by clamp deletion.** (A) and (B) display urea denaturation equilibrium unfolding curves of WT and clamp-deletion variants of IGPS, the lines represent fits of the data for each variant to a 3-state equilibrium folding model as described in the text. (a) sIGPS: sIGPS-WT (●, —), sIGPS-Δβ2z2-E74A (▲, —), sIGPS-Δβ3z3-D128A (●, —), and sIGPS-Δβ7z7-N228A (■, —). (b) eIGPS: eIGPS-WT (●, —), eIGPS-Δβ1z1-S82A (▲, —), eIGPS-Δβ3z3-D132A (●, —), and eIGPS-Δβ7z7-N231A (■, —). (C) and (D) are bar graphs representing the free energy differences for the N to I step in unfolding, ΔG_{N→I} (black bars) and the I to U step, ΔG_{I→U} (gray bars) for WT and the clamp-deletion variants of sIGPS (C) and eIGPS (D). The urea denaturation equilibrium unfolding curve of sIGPS-WT (A) and the corresponding folding free energy changes (C) are adapted from Forsyth et al. [17].

doi:10.1371/journal.pone.0007179.g003
Table 1. Thermodynamic parameters for the urea-induced unfolding of sIGS, eIGPS, zTS and eight βz-hairpin clamp-deletion variants.a

| Donor and acceptor pairs | Donor and acceptor distance (Å) | Variants | ΔG_unf (H2O) (kcal mol⁻¹) | Δm_H2O (kcal mol⁻¹ M⁻¹) | ΔG_unf (H2O, variant) (kcal mol⁻¹) | Δm_H2O (kcal mol⁻¹ M⁻¹) | ΔΔG_unf (H2O, WT) (kcal mol⁻¹)b |
|--------------------------|-------------------------------|----------|--------------------------|--------------------------|----------------------------------|--------------------------|-----------------------------|
| sIGPS                    |                               |          |                         |                          |                                  |                          |                             |
|                          | WT                            | 5.60±0.99| 2.46±0.42                | 12.39±0.60               | 2.60±0.13                       | -                        |                             |
|                          | F50–582                       | 3.0      | 4.84±0.19                | 2.05±0.08                | 13.24±0.20                      | 2.68±0.04                | -0.76±1.01                |
|                          | I111-D132                     | 2.9      | 6.69±1.57                | 3.34±0.77                | 13.36±1.22                      | 2.99±0.26                | 1.09±1.86                 |
|                          | V211-N231                     | 2.8      | 1.28±0.15                | 0.89±0.11                | 11.84±1.45                      | 2.62±0.31                | -4.32±1.00                |
| eIGPS                    |                               |          |                         |                          |                                  |                          |                             |
|                          | WT                            | 7.19±0.58| 2.85±0.24                | 3.04±0.85                | 0.81±0.17                       | -                        |                             |
|                          | F19–D46                       | 2.8      | 1.98±0.45                | 0.78±0.17                | 4.97±1.96                       | 1.07±0.39                | -5.21±0.73                |
|                          | I97-D124                      | 2.6      | 2.53±0.40                | 1.12±0.19                | 3.81±0.64                       | 0.79±0.16                | -4.66±0.70                |
| zTS                      |                               |          |                         |                          |                                  |                          |                             |
|                          | WT                            | 1.12     | 0.46                     | 0.73                     | 0.13                            | -                        |                             |
|                          | F19–D46                       | 2.8      | 1.98±0.45                | 0.78±0.17                | 4.97±1.96                       | 1.07±0.39                | -5.21±0.73                |
|                          | I97-D124                      | 2.6      | 2.53±0.40                | 1.12±0.19                | 3.81±0.64                       | 0.79±0.16                | -4.66±0.70                |

a Buffer conditions: 10 mM potassium phosphate, 0.2 mM K2EDTA, 1 mM jIME, pH 7.8 for sIGPS and pH 7.0 for eIGPS at 25°C.
b Perturbation in stability for the N to I reaction, calculated by ΔΔG_unf = ΔG_unf (H2O, variant) - ΔG_unf (H2O, WT).

Discussion

Experimental analysis of βz-hairpin clamps between MC H-bond donors and SC H-bond acceptors in three TIM barrel proteins, zTS [6], sIGPS and eIGPS, has shown that a subset of these non-covalent interactions make substantive contributions to stability. Comparisons of the potency of the βz-hairpin clamps in these three proteins shows no correlation between the contributions of these clamps to stability and either the location of the clamps in the structure, their contributing residues or their relative exposure (0–25%) to the solvent. The observation of potent clamps formed by the neutral N228 in sIGPS and N231 in eIGPS, the βzα7 clamps, also shows that the formal negative charge on the aspartic acid H-bond acceptor in the remaining two potent βz-hairpin clamps is not deterministic of the strength of the clamp interaction. An examination of the crystal structures of the three proteins, however, suggests that the length of the H-bond in each structure differentiates between the clamps that make major or minor contributions to stability (Table 1). Although the nominal resolutions of the crystal structures of these proteins, 2.0 to 2.8 Å [21–23], dictate that the correlation between H-bond length and the clamp contribution to protein stability be viewed as tentative, it appears βz-hairpin clamps whose H-bonds are less than 2.8 Å in length are those, which when replaced with alanine, reduce the stability of the native state by 4–6 kcal mol⁻¹. The apparent correlation provides a logical and testable hypothesis for future experiments on βz-hairpin clamps in other TIM barrel proteins.

The assay for the contribution of the βz-hairpin clamps to the stability of three TIM barrel proteins involves the replacement of the polar side chain H-bond acceptors, asparagine, aspartic acid, glutamic acid and serine, with alanine. The absence of the H-bond acceptor moiety is accompanied by the introduction of a potential void for these buried side chains, reflecting the absence of chemical mass as the side chain is truncated to the β-carbon. The perturbations in the secondary and/or tertiary structures induced by the mutations (Figure 2C and 2D) show that the loss of the clamp is propagated to numerous other non-covalent interactions via the global cooperativity of the native conformation.

The absence of the βz-hairpin clamps in the I states of all three TIM barrel proteins demonstrates that the potent effects of these clamps only appear as the N state appears [6]. Kinetic folding studies on zTS revealed further that each clamp is crucial for accessing the transition state ensemble required to reach the properly-folded structure [6]. Although the local connectivity of the ββ modules might have been expected to enable the clamp to have a role in the early stages of the folding reaction, the primary role of the potent set of clamps is to drive the final stage of the reaction to completion and fully develop global cooperativity. The 4-fold symmetry of the preferred βz-hairpin clamps is mirrored, not only in the symmetry of the ββ modules, but also in the packing of the side chains in the interior of the β-barrel. A residue oriented towards the inside of the β-barrel from an odd-numbered β-strand is at the same level as corresponding residues from the three remaining odd-numbered β-strands. The next layer is comprised of the four side chains from the even-numbered β-strands (Figure 5); the third, and usually final layer, is comprised again of side chains from the odd-numbered β-strands [20]. The layering of side chains inside the barrel has its origin in the tilt of the β-strands (35°) with respect to the central axis of the β-barrel [24]. The resulting S = 8 sheaf [20,24] provides a favorable orientation for the H-bonding network between adjacent parallel β-strands and provides opportunities for MC-SC ββ-hairpin clamp interactions. Together, these non-covalent interactions and others stabilize the ββ repeats.
The chemical origin for the asymmetry between odd- and even-numbered \( \beta \)-strands is apparent from an inspection of the residue preference (\( \sim 15\% \)) at positions preceding the N-terminus of each \( \beta \)-strand (Figure 5). The conserved proline just before odd-numbered \( \beta \)-strands provides a kink in the backbone that marks the beginning of a \( \beta \)-strand [5]. The preferred sequence pattern of the tight turn connecting the \( \alpha \)-helix and the subsequent even-numbered \( \beta \)-strand (Figure 5), GAD, has been reported previously [1]. The positive Q angle allowed by glycine and the hydrophobic nature of alanine immediately following the \( \alpha \)-helix enables a Schellman motif for the C-terminal capping of the helix [10] and a tight turn to the next \( \beta \)-strand. The aspartic acid just prior to the beginning of even-numbered \( \beta \)-strands forms the \( \beta \beta \)-hairpin clamp and braces the \( \beta \beta \beta \) module. This N-terminal cap for the odd-numbered \( \beta \)-strand is very often complemented by a MC–MC H-bond, with the amide group of the aspartic acid acting as the donor to the MC carbonyl oxygen of the partner residue. While other SC acceptors are observed (Table S1), the length of the aspartic acid side chain appears to be optimal for the reinforcement of the MC–SC H-bond with the MC–MC H-bond, providing a plausible explanation for its higher frequency in \( \beta \beta \)-hairpin clamps.

The preference for I/L/V residues at the MC NH\_H-bond donor position may reflect, in part, the \( \sim 40\% \) occurrence of these residues in parallel \( \beta \)-strands of TIM barrel proteins [20]. Further, along with alanine and glycine, I/L/V are the only amino acids that do not partition favorably from the vapor phase to water [25]. As such, these large aliphatic side chains are especially effective at excluding water from MC–SC H-bonds in the \( \beta \beta \)-hairpin clamps. The exclusion of water, that is apparent from the limited access to solvent for the H-bond donor and acceptor atoms of potent clamp interactions in \( \alpha \)TS, sIGPS and eIGPS (Figure 1C and 1D), is expected to strengthen these H-bonds and make them more resistant to exchange with solvent, as observed previously for \( \alpha \)TS [26–28]. This presumption is supported by the conclusions of Gao et al. [29], who recently reported that the strength of a MC–MC H-bond is inversely related to the polarity of its local environment. Valine more effectively screened an underlying \( \beta \)-sheet MC–MC H-bond from solvent than alanine in a Pin WW domain, increasing the strength of the H-bond by up to 1.2 kcal mol\(^{-1}\).

The occurrence of the \( \beta \beta \)-motif in a large number of protein families [3,20] suggests that the N-terminal capping of \( \beta \)-strands by \( \beta \beta \)-hairpin clamps, akin to the analogous N-capping of \( \alpha \)-helices [9,10], may be a useful property for the refinement of protein fold prediction and for engineering stability in \( \beta \beta \)-repeat proteins. \( \beta \beta \)-repeat proteins are readily recognized from their sequences and the predicted alternating patterns of \( \alpha \)-helices and \( \beta \)-strands [30]. The refinement of the 3D structures predicted from

**Figure 4. Positional preference of \( \beta \beta \)-hairpin clamps in 71 TIM barrel proteins.** (A) The TIM barrel architecture is represented by a cross-sectional view of the 8 \( \beta \)-strands, represented as rectangles and the strand number is indicated. The number of MC\_NH \( \rightarrow \) SC \( \beta \beta \)-hairpin clamp interactions connecting adjacent \( \beta \)-strands with SC H-bond acceptors C-terminal to the MC\_NH donors (\( \cdots \)), and with SC H-bond acceptors N-terminal to the MC\_NH donors (\( \cdots \)) are indicated. The number of \( \beta \beta \)-hairpin clamps with (I/L/V) MC \( \rightarrow \) SC (D) is represented in parenthesis. (B) The positional preference of (I/L/V) MC \( \rightarrow \) SC (D) relative to the \( \beta \)-strands. The MC donor prefers either the first or second position of the \( \beta \)-strand and the SC acceptor prefers to be in the loop immediately preceding the subsequent \( \beta \)-strand. The number of times each pair of interactions occurs in the 55 I/L/V MC \( \rightarrow \) SC D sub-set is indicated. doi:10.1371/journal.pone.0007179.g004

N-terminus of \( \beta \)-strand MC\_NH \( \rightarrow \) SC C-terminus of \( \alpha \)-helix

\[ \text{ILV} \rightarrow \text{D} \]
knowledge-based potentials [31], threading [32] and homology modeling [33] of these protein sequences, could be enhanced by screening for βα-hairpin clamps between the MC amide hydrogens at favored positions near the N-terminus of a β-strand and H-bond acceptor SC in the loop before the subsequent β-strand (25 residues apart in sequence). These clamps would establish the register of the pair of β-strands, and, with the very short loop linking the intervening α-helix to the second β-strand, it might be possible to establish the register of the α-helix on the β-strand pair in the βαβ module. Although TIM barrel proteins typically contain only a few βα-hairpin clamps, defining the spatial relationships of the components of a subset of βαβ modules might increase the probability of predicting the packing of adjacent βα repeats in the structures. The effect of accurately predicting the structure of one βαβ module might, therefore, propagate throughout the TIM barrel protein.

The TIM barrel architecture provides a scaffold that is capable of a very diverse set of enzymatic functions [1], and this property has enabled TIM barrel enzymes to be re-engineered in order to accommodate alternative substrates [34–37] and even to catalyze non-biological reactions [38]. Oligonucleotides for mutagenesis were purchased from Eurofins MWG Operon (Huntsville, AL), and the Quickchange™ site-directed mutagenesis kit was obtained from Stratagene (La Jolla, CA). The site-directed mutations were confirmed by DNA sequence analysis (Genewiz Inc, NJ).

### Materials and Methods

#### Clamp-deletion variants

The plasmid encoding a truncated version of sIGPS, in which the non-canonical additional α-helix (α00) at the N-terminus was deleted to eliminate aggregation during folding, pTNI4 [17], was obtained from Dr. K. Kirschner (University of Basel, Switzerland). The plasmid coding for eIGPS, pJB122 [39], was obtained from Dr. J. M. Blackburn (University of the Western Cape, South Africa). The eIGPS, with an additional Ala residue after the start codon and a C-terminal FLAG peptide sequence (GSDYKDDDDK), is fully folded and catalytically active [39]. The plasmid coding for eIGPS, pJB122 [39], was obtained from Dr. J. M. Blackburn (University of the Western Cape, South Africa). The eIGPS, with an additional Ala residue after the start codon and a C-terminal FLAG peptide sequence (GSDYKDDDDK), is fully folded and catalytically active [39]. Oligonucleotides for mutagenesis were purchased from Eurofins MWG Operon (Huntsville, AL), and the Quickchange™ site-directed mutagenesis kit was obtained from Stratagene (La Jolla, CA). The site-directed mutations were confirmed by DNA sequence analysis (Genewiz Inc, NJ).

#### Protein expression and purification

The sIGPS protein and its variants were expressed in BL21/DE3 cells and purified as described previously [17]. The expression and purification of eIGPS and its variants followed the same protocol, with the exception that the procedures were conducted at pH 7.0. The purity (>95%) was demonstrated by the appearance of a single band Coomassie blue stained PAGE and confirmed using electrospray mass spectrometry at the

---

**Figure 5. Architectural principles of the TIM barrel fold.** The strand number of the 8 β-strands of the TIM barrel architecture (•) is indicated at the C-terminus of each β-strand. To convey the closed barrel architecture, β8 is shown adjacent to β1 (•) as well as adjacent to β7. The position of each residue on the β-strands, with SC pointing into the β-barrel (⋯) and SC pointing towards the α-helices (○), is indicated. The one letter code for the most common amino acids (>15%) in the loop preceding the β-strand in the 71 TIM barrel proteins database is shown. The H-bond network for the β-barrel (→), the βα-hairpin clamp interactions between the second residue of an odd-numbered β-strand and the side chain of the residue immediately preceding the subsequent even-numbered β-strand (→), and the MC-MC interactions between the same two residues (→→) are indicated.

doi:10.1371/journal.pone.0007179.g005
Proteomics Facility at the University of Massachusetts Medical School (Worcester, MA).

Circular dichroism
Far- and near-UV CD spectroscopy was employed to monitor the secondary and the tertiary structure near aromatic side chains, respectively. Spectra were obtained on a Jasco Model J-810 spectropolarimeter equipped with a thermostatted cell holder. Far-UV CD data were collected from 280 nm to 185 nm at a scan rate of 50 nm/min and at 1 nm intervals using a 0.1 cm path length cell, with an averaging time of 8 s. Three replicate spectra were collected and averaged. The protein concentration was 50–150 μM. Near-UV CD data were collected from 350 nm to 250 nm at 5 nm/min using a 0.5 cm path length cell, and the protein concentration was 50–150 μM. The temperature was maintained at 25°C with a computer-controlled Peltier system.

Thermodynamic measurements
The stability of the IGS clamp-deletion variants was measured by urea denaturation as described previously [17] in a buffer containing 10 mM potassium phosphate, pH 7.8 for sIGPS and pH 7.0 for eIGPS. 0.2 mM K2EDTA, and 1 mM βME. A Hamilton 540B automatic titrator was used to prepare the samples containing 0 to 8 M urea at concentration increments of 0.2 M urea to enhance the precision of the measurements. The samples were incubated overnight at 25°C for eIGPS, 0.2 mM K2EDTA, and 1 mM βME. A Hamilton 540B automatic titrator was used to prepare the samples containing 0 to 8 M urea at concentration increments of 0.2 M urea to enhance the precision of the measurements. The samples were incubated overnight at 25°C to ensure equilibration.

Data analysis
Equilibrium CD data at 222 nm were fit to a three-state model, N ↔ I ↔ U, as described previously [40]. All thermodynamic folding data were fit using Savuka version 6.2, an in-house, non-linear, least-squares program [40].

Survey of TIM barrel proteins
A database of 71 TIM barrel proteins has been previously developed [http://www.cbcr.ipsr-gromiha/tim/proteinlist.html](http://www.cbcr.ipsr-gromiha/tim/proteinlist.html) from the SCOP [41] and HOMSTRAD [42] databases, with a pair-wise sequence homology of <25%. The highest resolution structure for each domain was chosen from the Protein Data Bank [43]. The secondary structure was calculated using the DSSP program [44] and the H-bond interaction parameters were calculated using default settings of the HBPLUS program [45].

Definitions of βα-hairpin clamp interactions
For each protein, the 8 canonical β-strands and α-helices in the context of the TIM barrel architecture were identified and labeled accordingly. H-bonding partners identified using the HBPLUS program [45], were subjected to the following filters: i) the H-bonds must be between a MC amide donor and a SC acceptor, ii) the amino acid chain length between the donor and acceptor must be ≥15 residues thereby eliminating shorter-range helix-capping interactions [9,10] and 3) the chain must include exactly one β-strand and one α-helix identified in the context of the TIM barrel architecture. For the case of the βα-hairpin clamps, MC_H → SC H-bonds between the residues prior to β1 and β8 were included. The H-bonds that passed each stage of the filtering process were exported to a PyMOL [46] script in color-coded fashion for manual confirmation.

Statistical significance of residue preference for βα-hairpin clamps.
The frequency of MC_H → SC H-bonds in the 71 TIM barrel proteins, where the donor and acceptor residues were at least 15 amino acids apart and were not involved in βα-hairpin clamp interactions, was determined. This frequency was used to calculate the expected frequency of H-bonding between any two types of residues and compared to that observed in βα-hairpin clamps. Four categories, Ile MC → SC D, Leu MC → SC D, Val MC → SC D, Other MC → SC Other, were used to determine the χ² distribution probabilities, with Yates correction [47], of observed βα-hairpin clamps.

Supporting Information

**Dataset S1** Thermodynamic analysis of the eIGPS Δβα7 N231A clamp-deletion variant. Found at: doi:10.1371/journal.pone.0007179.s001 (0.04 MB DOC)

**Methods S1** Kinetic experiments. Found at: doi:10.1371/journal.pone.0007179.s002 (0.04 MB DOC)

**Figure S1** Unfolding amplitude of eIGPS Δβα7 N231A as a function of initial urea concentration. Found at: doi:10.1371/journal.pone.0007179.s003 (0.26 MB DOC)

**Table S1** Sequence preferences for βα-hairpin clamps in 71 TIM barrel proteins. Found at: doi:10.1371/journal.pone.0007179.s004 (0.07 MB DOC)

Acknowledgments
The authors would like to thank Dr. K. Kirschner (University of Basel, Switzerland) and Dr. J. M. Blackburn (University of the Western Cape, South Africa) for sharing the plasmids encoding the sIGPS and eIGPS proteins, respectively. We also thank Dr. Jill A. Zitewitz, Dr. Osman Bäkel, Dr. Zhenyu Gu, and Can Kayatekin for helpful discussions and assistance in the preparation of this manuscript.

Author Contributions
Conceived and designed the experiments: XY RV CRM. Performed the experiments: XY. Analyzed the data: XY SVK RV. Contributed reagents/materials/analysis tools: SVK. Wrote the paper: XY SVK RV CRM.

References
1. Nagano N, Orengo CA, Thornton JM (2002) One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. J Mol Biol 321: 741–763.
2. Zitewitz JA, Gualfetti PJ, Perkons IA, Wasta SA, Matthews CR. (1999) Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a β/α barrel protein. Protein Sci 8: 1200–1209.
3. Gerstein M (1997) A structural census of genomes: comparing bacterial, eukaryotic, and archaeal genomes in terms of protein structure. J Mol Biol 274: 1143–1159.
4. Frenkel ZM, Trifonov EN (2005) Closed loops of TIM barrel protein fold. J Biomol Struct Dyn 22: 643–656.
5. FarzadFard F, Gharaei N, Pesezhik H, Marashi S-A (2008) [beta]-Sheet capping: Signals that initiate and terminate [beta]-sheet formation. Journal of Structural Biology 161: 101–110.
6. Yang X, Vadrevu R, Wu Y, Matthews CR (2007) Long-range side-chain-main-chain interactions play crucial roles in stabilizing the (betaalpha)8 barrel motif of the alpha subunit of tryptophan synthase. Protein Sci 16: 1398–1409.
7. Baker EN, Hubbard RE (1984) Hydrogen bonding in globular proteins. Prog Biophys Mol Biol 44: 97–179.
8. Stickle DF, Presta LG, Dill KA, Rose GD (1992) Hydrogen bonding in globular proteins. J Mol Biol 226: 1143–1159.
9. Presta LG, Rose GD (1988) Helix signals in proteins. Science 240: 1632–1641.
10. Aurora R, Rose GD (1998) Helix capping. Protein Sci 7: 21–38.
11. Horwitz A, Serrano I, Avron Y, Bycroft M, Fersht AR (1990) Strength and cooperativity of contributions of surface salt bridges to protein stability. J Mol Biol 216: 1031–1044.
12. Serrano L, Kellis JT, Jr, Cann P, Matouschek A, Fersht AR (1992) The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. J Mol Biol 224: 783–804.
13. Myers JK, Face CN (1996) Hydrogen bonding stabilizes globular proteins. Biophys J 71: 2033–2039.
14. Ibarr-Molero B, Zitewitz JA, Matthews CR (2004) Salt-bridges can stabilize but do not accelerate the folding of the homodimeric coiled-coil peptide GGNH-p1. J Mol Biol 336: 969–990.
15. Gromiha MM, Pujadas G, Magyar C, Selvaraj S, Simon I (2004) Locating the stabilizing residues in (alpha/beta)8 barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation. Proteins 55: 316–329.
16. Matthews CR, Crisanti MM (1981) Urea-induced unfolding of the alpha subunit of tryptophan synthase: evidence for a multistate process. Biochemistry 20: 784–792.
17. Forsyth WR, Matthews CR (2002) Folding mechanism of indole-3-glycerol phosphate synthase from Sulfolobus solfataricus: a test of the conservation of folding mechanisms hypothesis in (beta/alpha)/8 barrels. J Mol Biol 320: 1119–1133.
18. Sanchez del Pino MM, Fersht AR (1997) Nonsequential unfolding of the alpha/ beta barrel protein indole-3-glycerol-phosphate synthase. Biochemistry 36: 5360–5365.
19. Matthews CR (1987) Effect of point mutations on the folding of globular proteins. Methods Enzymol 134: 498–511.
20. Branden C, Tooze J (1999) Introduction to Protein Structure. New York: Garland Science Publishing. 302 p.
21. Schneider B, Knochel T, Darimont B, Hennig M, Dietrich S, et al. (2005) Role of the N-terminal extension of the (betaalpha)8-barrel enzyme indole-3-glycerol phosphate synthase for its fold, stability, and catalytic activity. Biochemistry 44: 16403–16412.
22. Wilman M, Priestle JP, Niermann T, Jansonius JN (1992) Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase: indolglycerol phosphate synthase from Escherichia coli refined at 2.0 A resolution. J Mol Biol 225: 477–507.
23. Hyde CC, Ahmed SA, Padlan EA, Miles EW, Davies DR (1988) Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. J Biol Chem 263: 17057–17071.
24. Wierenga RK (2001) The TIM-barrel fold: a versatile framework for efficient coupling between hydrogen bonding and microenvironment polarity substructures. J Mol Biol 307: 479–502.
25. Radzicka A, Wolfenden R (1988) Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry 27: 1664–1670.
26. Vadhrevu R, Wu Y, Matthews CR (2008) NMR analysis of partially folded states and persistent structure in the alpha subunit of tryptophan synthase: implications for the equilibrium folding mechanism of a 29-kDa TIM barrel protein. J Mol Biol 377: 294–306.
27. Wu Y, Vadhrevu R, Kathuria S, Yang X, Matthews CR (2007) A tightly packed hydrophobic cluster directs the formation of an off-pathway sub-millisecond folding intermediate in the alpha subunit of tryptophan synthase, a TIM barrel protein. J Mol Biol 366: 1624–1638.
28. Gu Z, Zitewitz JA, Matthews CR (2007) 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 Mol Biol 368: 582–594.
29. Gao J, Bosco DA, Powers ET, Kelly JW (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. Nat Struct Mol Biol 16: 684–690.
30. Kunin V, Chan B, Sitton E, Lihvick G, Pietrokovski S (2001) Consistency analysis of similarity between multiple alignments: prediction of protein function and fold structure from analysis of local sequence motifs. J Mol Biol 307: 939–949.
31. Rohl GA, Strass CE, Chivian D, Baker D (2004) Modeling structurally variable regions in homologous proteins with rosetta. Proteins 55: 656–677.
32. Brylinki M, Skolnick J (2009) FINDSITE: a threading-based approach to ligand homology modeling. PLoS Comput Biol 5: e1000405. doi:10.1371journal.pcbi.1000405.
33. Delano WL (2002) The PyMOL Molecular Graphics System: DeLano Scientific.
34. Kunin V, Chan B, Sitton E, Lihvick G, Pietrokovski S (2001) Consistency analysis of similarity between multiple alignments: prediction of protein function and fold structure from analysis of local sequence motifs. J Mol Biol 307: 939–949.
35. Jurgens C, Stroon A, Wegener D, Hettwer S, Wilmanns M, et al. (2006) Directed evolution of a (beta/alpha)8-barrel enzyme to catalyze related reactions in two different metabolic pathways. Proc Natl Acad Sci U S A 97: 9925–9930.
36. Golz JA, Babliet FC (2009) Enzyme design: lessons from natural evolution and computation. Curr Opin Chem Biol 13: 10–18.
37. Sanchez del Pino MM, Fersht AR (1997) Nonsequential unfolding of the alpha/ beta barrel protein indole-3-glycerol-phosphate synthase. Biochemistry 36: 5360–5365.
38. Rothlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, et al. (2008) Kemp elimination catalysts by computational enzyme design. Nature 453: 190–195.
39. Alaniunmaa MM, Blackburn JM, Aguayo C, Fersht AR (2000) Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold. Nature 403: 617–622.
40. Bixel O, Zitewitz JA, Bowers KE, Matthews CR (1999) Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. Biochemistry 38: 1018–1029.
41. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 246: 403–412.
42. Mizuguchi K, Deane CM, Blundell TL, Overington JP (1998) HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci 7: 2469–2474.
43. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235–242.
44. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22: 2573–2637.
45. Rohl GA, Strass CE, Chivian D, Baker D (2004) Modeling structurally variable regions in homologous proteins with rosetta. Proteins 55: 656–677.
46. Yang X, Matthews CR (2007) 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 Mol Biol 368: 582–594.
47. Golz JA, Babliet FC (2009) Enzyme design: lessons from natural evolution and computation. Curr Opin Chem Biol 13: 10–18.