

β -Strand Flipping and Slipping Triggered by Turn Replacement Reveal the Opportunistic Nature of β -Strand Pairing

Koki Makabe, Shude Yan, Valentina Tereshko, Grzegorz Gawlak, and
Shohei Koide*

Contribution from the Department of Biochemistry and Molecular Biology, The University of
Chicago, Chicago, Illinois 60637

Received June 11, 2007; E-mail: skoide@uchicago.edu

Abstract: We investigated how the register between adjacent β -strands is specified using a series of mutants of the single-layer β -sheet (SLB) in *Borrelia* OspA. The single-layer architecture of this system eliminates structural restraints imposed by a hydrophobic core, enabling us to address this question. A critical turn (turn 9/10) in the SLB was replaced with a segment with an intentional structural mismatch. Its crystal structure revealed a one-residue insertion into the central β -strand (strand 9) of the SLB. This insertion triggered a surprisingly large-scale structural rearrangement: (i) the central strand (strand 9) was shifted by one residue, causing the strand to flip with respect to the adjacent β -strands and thus completely disrupting the native side-chain contacts; (ii) the three-residue turn located on the opposite end of the β -strand (turn 8/9) was pushed into its preceding β -strand (strand 8); (iii) the register between strands 8 and 9 was shifted by three residues. Replacing the original sequence for turn 8/9 with a stronger turn motif restored the original strand register but still with a flipped β -strand 9. The stability differences of these distinct structures were surprisingly small, consistent with an energy landscape where multiple low-energy states with different β -sheet configurations exist. The observed conformations can be rationalized in terms of maximizing the number of backbone H-bonds. These results suggest that adjacent β -strands “stick” through the use of factors that are not highly sequence specific and that β -strands could slide back and forth relatively easily in the absence of external elements such as turns and tertiary packing.

Introduction

A β -sheet is composed of specifically aligned β -strands. Specific alignment of β -strands is essential for highly reproducible folding of globular proteins and also the formation of β -rich self-assemblies in which short β -strand peptides are aligned in a very specific manner.^{1,2} A paucity of experimental systems has made it difficult to evaluate major determinants governing the alignment of adjacent β -strands.

Although β -sheets in small globular proteins are good model systems for characterizing different factors governing the stability of β -sheet structure,^{3–5} their structures are usually defined by extensive interactions with elements outside the β -sheet regions, typically those involving a hydrophobic core. Folding reactions of these proteins usually follow the two-state mechanism; i.e., they are either folded in one well-defined configuration or unfolded. Thus, different β -strand alignments are rarely observed in this type of β -sheet systems.

Short peptides that form a β -hairpin and other types of β -sheet-containing structures are other model systems for β -sheet formation, complementary to small globular proteins.^{6–8} Such

peptide systems are usually highly dynamic, sampling multiple conformations. While this conformational dynamics potentially offers an opportunity to evaluate factors that govern the relative stability among multiple conformational species, it is in practice extremely difficult to unambiguously determine the conformations of individual species and quantify their respective populations. Taken together, although a wealth of knowledge exists as to the factors stabilizing a β -sheet structure relative to the unfolded state, still little is known about why a specific β -sheet architecture is preferentially stabilized among many possible conformations.

Understanding factors governing β -strand alignment is also important for protein structure prediction and design. Predicting the structure of a predominantly β -sheet protein is still a major challenge.⁹ Accurate prediction of strands pairing in a β -sheet is clearly critical for successful prediction of the overall β -sheet architecture. The insufficient understanding of the physical basics of β -sheet formation also appears to contribute to the difficulty in computationally designing all- β proteins.¹⁰

(1) Dobson, C. M. *Nature* **2003**, 426, 884–890.

(2) Nelson, R.; Eisenberg, D. *Curr. Opin. Struct. Biol.* **2006**, 16, 260–265.

(3) Smith, C. K.; Regan, L. *Acc. Chem. Res.* **1997**, 30, 153–161.

(4) Minor, D. L. J.; Kim, P. S. *Nature* **1994**, 367, 660–663.

(5) Minor, D. L. J.; Kim, P. S. *Nature* **1996**, 380, 730–734.

(6) Serrano, L. *Adv. Protein Chem.* **2000**, 53, 49–85.

(7) Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, 2, 717–725.

(8) de Alba, E.; Jiménez, M. A.; Rico, M.; Nieto, J. L. *Fold. Design* **1996**, 1, 133–144.

(9) Steward, R. E.; Thornton, J. M. *Proteins* **2002**, 48, 178–191.

(10) Dantas, G.; Kuhlman, B.; Callender, D.; Wong, M.; Baker, D. *J. Mol. Biol.* **2003**, 332, 449–460.

Previous studies have indicated the importance of side-chain–side-chain pairing,^{11–14} backbone hydrogen-bonding,¹⁵ and β -turn structure^{16–18} for the determination of the strand register. However, important remaining questions in β -sheet formation include (i) whether or not interactions between two adjacent β -strands can specifically stabilize a particular pairing pattern and (ii) the contribution of such interstrand interactions relative to other factors that can also stabilize a β -sheet structure, such as hydrophobic core formation and effects of turns and loops linking β -strands.

We have investigated the molecular mechanism of β -sheet formation using the single-layer β -sheet (SLB) segment of outer surface protein A (OspA) from *Borrelia burgdorferi* as a model system. The solvent-exposed SLB segment, consisting of three β -strands, is located at the center of a large antiparallel β -sheet that spans the two globular domains (Figure 1A).¹⁹ Unlike most of β -sheets found in water-soluble proteins, the SLB segment is not associated with a hydrophobic core. Nevertheless, it is highly stable.^{20–22} We have determined the folding mechanism of OspA and stability contributions of SLB residues.^{22–24} We have also established a surface-engineered variant of OspA that consistently produces high-resolution crystal structures of SLB variants.^{24–26} Thus, the OspA SLB offers an excellent model system for studying the mechanism of β -sheet formation in that it minimizes structural context associated with a hydrophobic core, similar to peptide systems, and at the same time it allows precise characterization of conformation and energetics.

In this work, we examined how the β -strand pairing in the OspA SLB segment is determined. This investigation was made possible by our discovery that a perturbation of a critical turn resulted in a large-scale structural rearrangement of the SLB segment. We subsequently constructed a series of turn mutants and determined their X-ray crystal structures and stability changes. Results strongly suggest that backbone hydrogen-bonding is a major determinant of the pairing pattern of adjacent β -strands in the absence of external elements and that specific side-chain interactions play a minor role. This inherent low sequence-specificity of β -strand pairing has important implications in protein structure prediction, design, misfolding, and self-assembly.

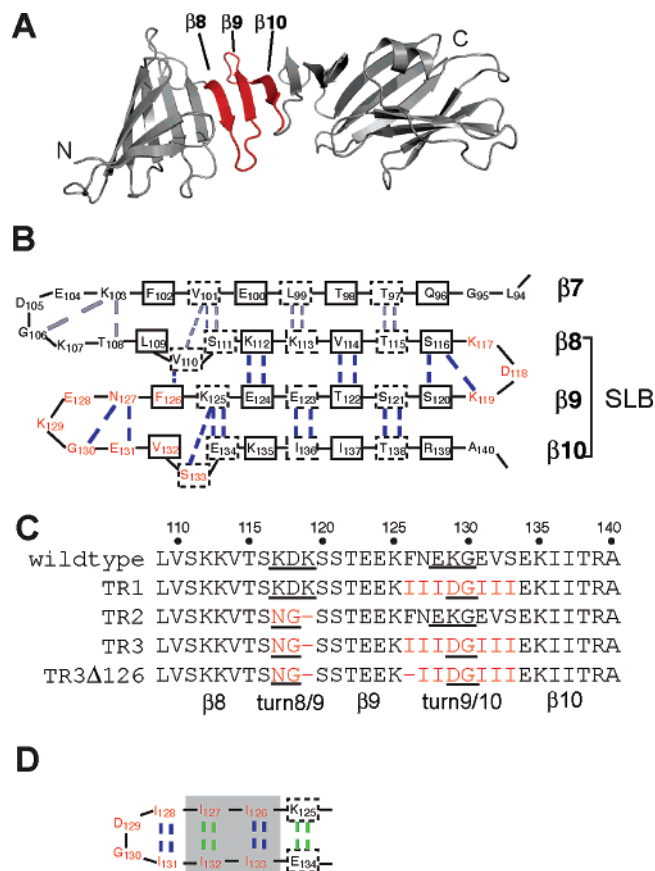


Figure 1. Architecture of OspA and the SLB segment. (A) Ribbon representation of the structure of wild-type OspA-sm1 (PDB ID 2G8C). The SLB segment (β -strands 8–10) is shown in red. (B) The amino acid sequence of the SLB segment, represented according to the secondary structure topology. Residues in a β -strand are enclosed in a box. The solid-line rectangles indicate residues whose side chain faces toward the reader. Main-chain hydrogen bonds are shown as dashed lines. Residues mutated in this study are shown in red. (C) Residues 109–140 of OspA mutants used in this study with mutated residues shown in red. (D) Schematic representation of H-bond mismatches introduced in the TR1 mutant. The backbone H-bonds expected for the hairpin structure of the IIIDGIII segment are shown in blue, and those expected from the H-bond pattern of β -strands 9 and 10 are in green. The region containing an incompatible H-bond pattern is shaded in gray.

Results

Turn Replacement Resulted in an Unexpected One-Residue Addition To β -Strand 9. In this work, we used the “sm1” version of OspA that contains 12 surface mutations in the two globular domains that facilitate crystallization.²⁶ These surface mutations are at distant locations from the SLB region and do not affect the structure of SLB.²⁶ The urea-induced unfolding of the sm1 OspA variant exhibited a two-state behavior (Supplementary Figure 1A, Supporting Information), as opposed to the three-state unfolding of the wild-type protein.^{21,23} This difference is presumably due to a collapse of two separate unfolding transitions that result from different degrees of destabilization of the two globular domains by the surface mutations. For brevity, we will refer to this “sm1” OspA that contains the wild-type SLB region as “wild type” hereafter.

In order to promote the formation of alternative β -sheet arrangements, we replaced a turn segment between β -strands 9 and 10 (residues 126–133; FNEKGEVS; Figure 1B) with a distinctly different sequence (IIIDGIII) (Figure 1C). According

- (11) Wouters, M. A.; Curmi, P. M. *Proteins* **1995**, *22*, 119–131.
- (12) Hutchinson, E. G.; Sessions, R. B.; Thornton, J. M.; Woolfson, D. N. *Protein Sci.* **1998**, *7*, 2287–2300.
- (13) Fooks, H. M.; Martin, A. C.; Woolfson, D. N.; Sessions, R. B.; Hutchinson, E. G. *J. Mol. Biol.* **2006**, *356*, 32–44.
- (14) Mandel-Gutfreund, Y.; Zaremba, S. M.; Gregoret, L. M. *J. Mol. Biol.* **2001**, *305*, 1145–1159.
- (15) Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. *Nature* **2004**, *430*, 101–105.
- (16) Searle, M. S.; Williams, D. H.; Packman, L. C. *Nat. Struct. Biol.* **1995**, *2*, 999–1006.
- (17) Blanco, F.; Ramirez-Alvarado, M.; Serrano, L. *Curr. Opin. Struct. Biol.* **1998**, *8*, 107–111.
- (18) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236–4237.
- (19) Li, H.; Dunn, J. J.; Luft, B. J.; Lawson, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3584–3589.
- (20) Pham, T.-N.; Koide, A.; Koide, S. *Nat. Struct. Biol.* **1998**, *5*, 115–119.
- (21) Koide, S.; Bu, Z.; Risal, D.; Pham, T.-N.; Nakagawa, T.; Tamura, A.; Engelman, D. M. *Biochemistry* **1999**, *38*, 4757–4767.
- (22) Yan, S.; Kennedy, S.; Koide, S. *J. Mol. Biol.* **2002**, *323*, 363–375.
- (23) Yan, S.; Gawlak, G.; Smith, J.; Silver, L.; Koide, A.; Koide, S. *J. Mol. Biol.* **2004**, *338*, 811–825.
- (24) Yan, S.; Gawlak, G.; Makabe, K.; Tereshko, V.; Koide, A.; Koide, S. *J. Mol. Biol.* **2007**, *368*, 230–243.
- (25) Makabe, K.; McElheny, D.; Tereshko, V.; Hilyard, A.; Gawlak, G.; Yan, S.; Koide, A.; Koide, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17753–17758.
- (26) Makabe, K.; Tereshko, V.; Gawlak, G.; Yan, S.; Koide, S. *Protein Sci.* **2006**, *15*, 1907–1914.

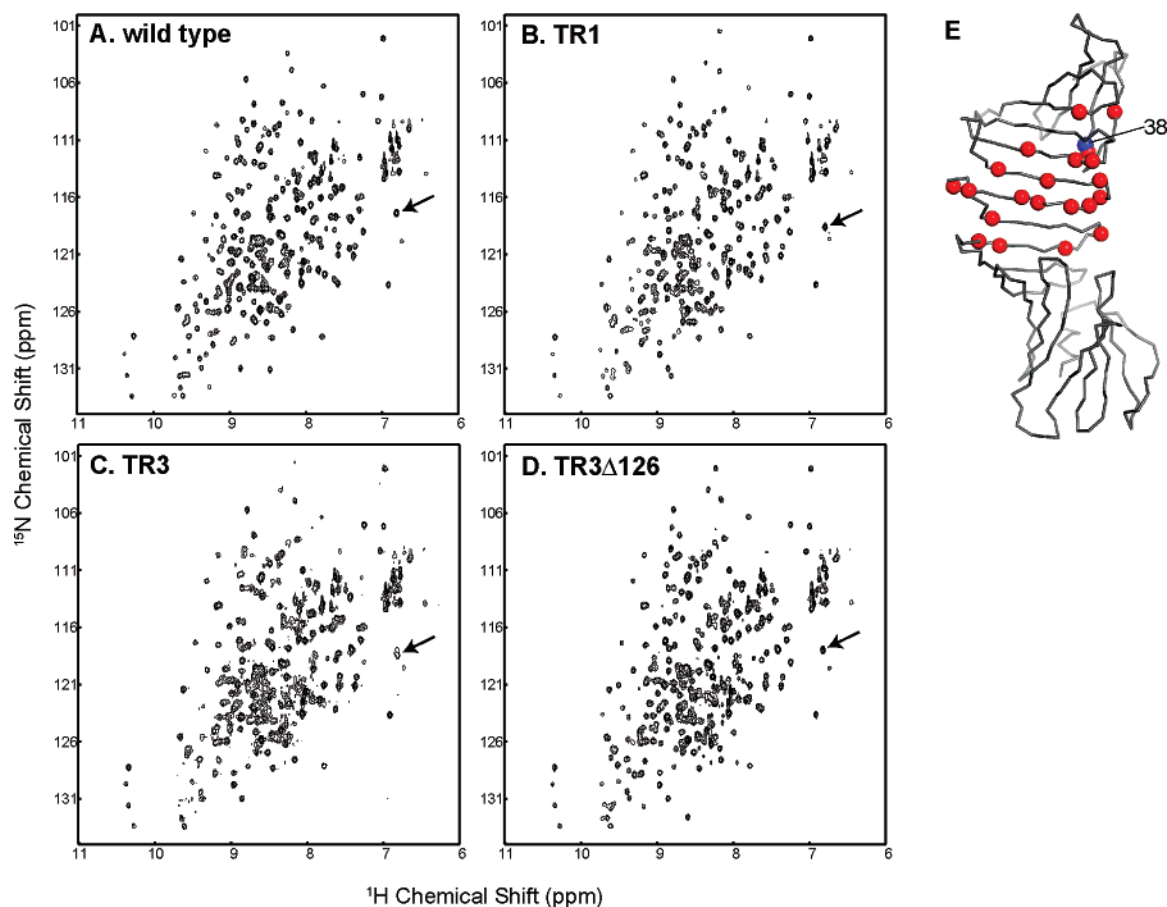


Figure 2. (A–D) ^1H , ^{15}N HSQC spectra of wild-type and mutant OspA proteins. The cross-peak for residue 38 is indicated with the arrow. The identity of samples is indicated in each panel. (E) The residues whose resonances are affected by the TR1 mutation mapped on the crystal structure of the wild-type protein. Residue 38 is also shown as a blue sphere. Note that only unambiguously identified residues are shown, and the absence of a sphere means either that the position is unaffected or that its peak has been excluded from analysis.

to the criteria of PROMOTIF,²⁷ F126 and V132 in the wild-type SLB are classified as β -strand residues (as parts of β -strands 9 and 10, respectively), residues 127–131 as a 3:5 turn, and S133 as a β -bulge (Figure 1B). F126 packs against V132, and it is the most important residue for stability among residues 126–133 as probed with Ala scanning mutagenesis.^{23,24} Because we anticipated a large loss of conformational stability upon removal of these critical residues, we chose the IIIDGIII sequence that might compensate the stability loss. It consists of residues with high β -sheet propensity (Ile)^{4,28} and a strong turn motif (Asp-Gly).²⁹ We envisioned that cross-strand interactions between the Ile residues might further stabilize the anticipated β -hairpin conformation. At the same time, the segment introduces a mismatch in the backbone H-bond pattern to the SLB. The Asp-Gly sequence has a high propensity to form a type I' turn, in which two backbone H-bonds are formed between the flanking residues (I128 and I131 in the II/DG/II sequence, the italicized residues). Consequently, if the two Ile₃ segments respectively form a β -strand, backbone H-bonds should also form between I126 and I133. Because K125 and E134, residues immediately adjacent to I126 and I133, respectively, are a H-bonded pair in the wild-type SLB (Figure 1B), simultaneous formation of H-bonds between I126 and I133 and

Table 1. Thermodynamic Parameters for the Turn Mutants^a

mutant	$C_{0.5}$ (M)	$\Delta G^{2.5M}$ (kcal mol ⁻¹)	$\Delta\Delta G^{2.5M}$ (kcal mol ⁻¹)	m (kcal mol ⁻¹ M ⁻¹)
wild type (OspA-sm1)	2.8	1.0 ± 0.1		-3.9 ± 0.1
TR1	2.1	-1.5 ± 0.2	2.5 ± 0.2	-4.0 ± 0.1
TR2	2.6	0.3 ± 0.2	0.7 ± 0.2	-3.5 ± 0.1
TR3	1.9	-2.0 ± 0.1	3.0 ± 0.1	-3.4 ± 0.1
TR3/ Δ 126	2.2	-1.2 ± 0.2	2.2 ± 0.2	-4.1 ± 0.1

^a Thermodynamic parameters were determined using global fitting of CD and fluorescence data based on a two-state model. $C_{0.5}$ is the urea concentration at the midpoint of transition. $\Delta G^{2.5M}$ is free energy difference at 2.5 M urea; $\Delta\Delta G^{2.5M}$ is the change in $\Delta G^{2.5M}$ caused by a mutation. A positive $\Delta\Delta G^{2.5M}$ value indicates destabilization. m is the dependence of ΔG on urea concentration. The errors indicated are the standard deviations determined from curve fitting.

those between K125 and E134 is not possible within a β -sheet structure. Thus, this turn replacement introduces severe structural strain that must be somehow accommodated through structural adjustments. We will refer to this mutant as TR1 ("turn replacement 1"; Figure 1C).

The TR1 mutant was expressed as a soluble protein and purified in the same manner as the wild type. Its NMR spectrum indicated that the protein was highly structured (Figure 2B). Although this protein was significantly less stable than the wild type, the level of destabilization ($\Delta\Delta G = 2.5$ kcal/mol; in this paper a positive $\Delta\Delta G$ value indicates destabilization; Table 1 and Supplementary Figure 1) was smaller than that induced by

(27) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1996**, *5*, 212–220.

(28) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510–5517.

(29) Ramirez-Alvarado, M.; Blanco, F. J.; Niemann, H.; Serrano, L. *J. Mol. Biol.* **1997**, *273*, 898–912.

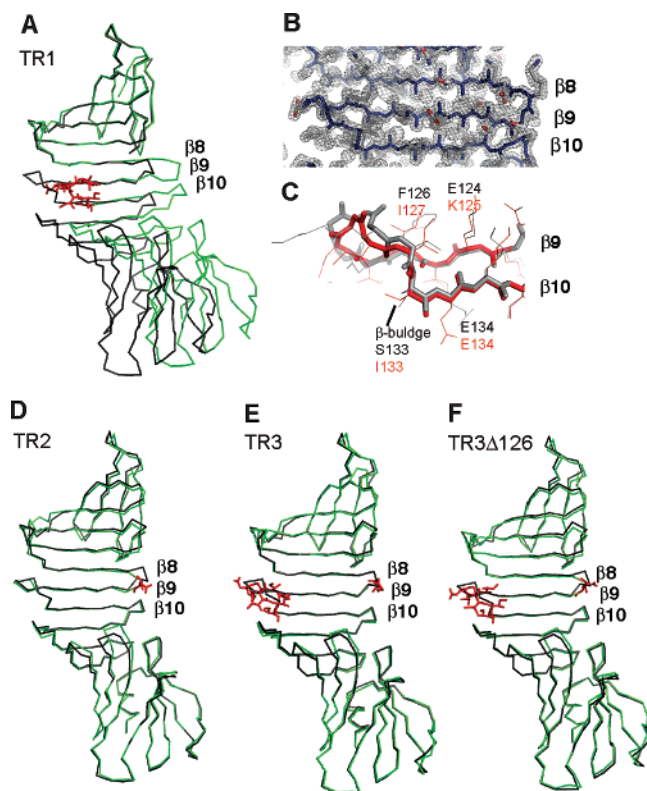


Figure 3. X-ray crystal structures of turn replacement mutants of OspA. (A) The C α trace of the X-ray crystal structure of TR1 (green) superposed on the wild-type structure (gray). Only the N-terminal domain was used for superposition. Mutated residues are shown as stick representation in red. (B) The σ_A -weighted $2F_o - F_c$ electron density map contoured at 1.0σ for the SLB region of TR1 structure. Solvent molecules are omitted for clarity. (C) Superposition of residues 123–135 of TR1 (red) and the wild type (gray), illustrating the conservation of the backbone conformation. Backbone atoms are shown as sticks, and side chains are shown as lines. The residues are labeled in cognate color. The β -bulge at position 133 is marked. (D–F) X-ray crystal structures of TR2 (D), TR3 (E), and TR3 Δ 126 (F) in blue superposed with the wild type in gray. Unlike in (A), the entire structures were superposed in (D–F).

a single-point mutation, F126A, in the same turn ($\Delta\Delta G = 3.3$ kcal/mol) that retained the overall fold.²³ These results indicate that the SLB segment managed to accommodate the new sequence without detrimental effects.

The X-ray crystal structure of TR1 at 1.35-Å resolution revealed a series of unanticipated structural rearrangements (Figure 3A,B). There is a large shift of the C-terminal domain relative to the N-terminal domain, suggesting a large magnitude of structural rearrangements (Figure 3A). Although all the rearrangements occurred simultaneously and cooperatively, for the sake of clarity we will dissect and describe them in a sequential manner.

The IDGI sequence in the mutant forms type I' turn as designed. These two residues replaced a three-residue turn in the wild-type structure, effectively shortening this turn segment by one residue. Although the introduced sequence is distinctly different from the original (Figure 1C), its backbone conformation is similar to that of the wild-type sequence except for the shorter turn. The β -bulge structure at position 133 (wild-type numbering) was retained in TR1 (Figure 3C). The bulge structure created a mismatch between the two flanking “III” sequences, which effectively resulted in a one-residue insertion between residues 125 and 126 of the wild-type structure, i.e., at the C-terminus of β -strand 9 (Figures 1B, 3A, and 3B).

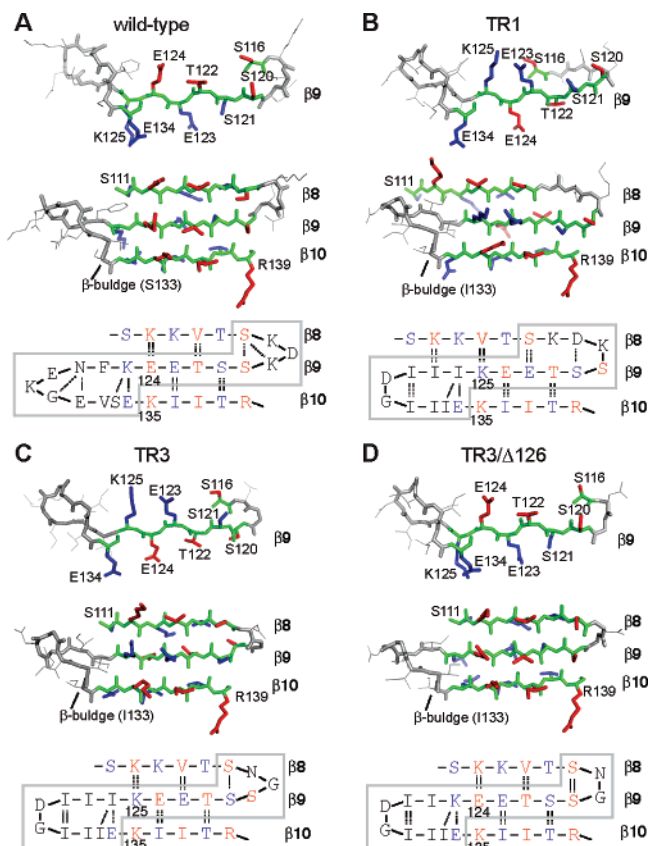


Figure 4. Conformation and β -strands register of the SLB region in the wild-type and mutant proteins. (A) The “top view” of residues 111–139 of the wild type is shown in the middle. The “side view” of residues 116–134 is shown at the top. The backbone atoms of residues 111–116, 120–125, and 134–139 are shown in green and those of the other residues in gray. The side chains that face toward the reader in the wild-type structure are shown as red sticks, and those facing backward in blue. The schematic drawing at the bottom is colored in the same manner as the above models, and H-bonds are depicted as dotted lines. The segment shown in the side view is enclosed in gray lines. (B–D) The structures of TR1 (B), TR3 (C), and TR3 Δ 126 (D) presented in the same manner as in (A). Residues are numbered according to the wild-type numbering as indicated in Figure 1C.

In the new turn structure, the side chains of the first, third, and last Ile residues in the IIDGIII motif (underlined residues) form a hydrophobic cluster on one face, and those of the second and fifth Ile residues are closely packed on the other face. Thus, we speculate that the presence of a strong turn motif and extensive hydrophobic contacts favors the observed, “curled” conformation, and this structure alleviates the introduced structural mismatch within the IIDGIII but excluded Ile126 from the turn region.

One-Residue Insertion Triggered β -Strand Flipping and Slipping. The extra residue (Ile126) in β -strand 9 was accommodated by a one-residue register shift of the rest of this β -strand toward the N-terminus (Figure 4B). Because of the two-residue periodicity of a β -strand, this register shift resulted in a complete flip of β -strand 9 with respect to the adjacent β -strands. That is, those side chains that face “up” in the wild-type structure now face “down” in the TR1 structure, and residues that form backbone H-bonds with those of strand 8 in the wild type now form H-bonds with strand 10 residue, and vice versa (compare Figure 4 panels A and B). This shift resulted in one extra residue in the turn region between β -strands 8 and 9 (referred to as turn 8/9). Residues 116–120 (SKDKS) in the wild type form a 3:5 turn (Figure 3A). In contrast, in the TR1 structure, residues

119–120 (SKDKS, the underlined residues) formed a type I (2:4) turn (Figure 3B), thus shifting the turn location. The new turn added two residues to the C-terminus of β -strand 8, which resulted in a three-residue register shift of the entire β -strand 8. Together, these rearrangements completely disrupted side-chain interactions between β -strands 8 and 9 and those between β -strands 9 and 10. Four cross-strand pairs of Glu and Lys were eliminated in TR1, and one Lys-Lys pair was introduced between residues 125 and 135 (Figure 4B). Therefore, we regarded the stability decrease of 2.5 kcal/mol as surprisingly small.

The β -strand register shift in the SLB segment is responsible for an offset between the two globular domains. This is clearly manifested in the decreased overall structural similarity between the wild type and TR1 (Figure 3A). The root-mean-square deviation (rmsd) for the equivalent C α atoms for the entire molecule is 2.33 Å, while those for the N- and C-terminal globular domains, calculated separately, are 0.66 and 0.54 Å, respectively.

We confirmed by NMR spectroscopy that the observed structural changes are not due to crystal packing. A comparison of ^1H , ^{15}N HSQC spectra of the wild-type and TR1 proteins revealed that the entire SLB region and parts of the globular domains are affected by the turn replacement, consistent with the extensive structural rearrangements seen in the TR1 crystal structure (Figure 2E). Also, we did not find clear evidence for the presence of multiple conformers in the TR1 NMR spectrum (Figure 2B). Taken together, the NMR data suggest that the TR1 crystal structure represents its dominant structure in solution.

A Stronger Turn Motif in Turn 8/9 Restores the Original Backbone Arrangement between β -Strands 8 and 9. Our previous study of a peptide fragment showed that the amino acid sequence for turn 8/9 (KDK, Figure 1B) in the wild type does not have strong turn-forming propensity.³⁰ We thus tested whether the replacement of this turn with a stronger turn-forming motif restored the β -strand register. We chose the Asn-Gly sequence for this purpose because it has a high propensity to form a type I' β -turn.³¹

First, we replaced the KDK sequence of the wild-type OspA with the NG sequence to check its compatibility. This mutant, termed TR2 (Figure 1C), was slightly destabilized compared with the wild type ($\Delta\Delta G = -0.7$ kcal/mol; Table 1 and Supplementary Figure 1). We determined its X-ray crystal structure at 1.7 Å resolution (Figure 3D). The mutated segment formed the type I' (2:2) turn as designed, with two backbone H-bonds formed between the (–B1) and (+B1) positions³² (S116 and S120, respectively; note that the numbering is according to the wild type, and residue 119 corresponds to one-residue deletion). The overall rmsd for the equivalent C α atoms between the wild type and the TR2 mutant was 1.03 Å, while those for the N- and C-terminal globular domains (residues 28–94 and 141–273, respectively) are 0.52 and 0.55 Å, respectively, suggesting a small change in the relative domain orientation. These results indicate that the new turn motif is well accommodated in the location. The small destabilization may be

attributed to the loss of interactions between this turn segment and adjacent regions.

Having confirmed the NG sequence's compatibility, we then introduced this turn motif into the TR1 mutant. The resulting mutant, termed TR3, thus contains the NG sequence in the 8/9 turn and the IIDGIII sequence in the 9/10 turn (Figure 1C). While this mutant was soluble and monomeric like the wild type and TR1, it was further destabilized ($\Delta\Delta G = 3.0$ kcal/mol; Table 1 and Supplementary Figure 1). The degree of destabilization was, however, no greater than the sum of the destabilizations observed for TR1 and TR2 (3.2 kcal/mol). The crystal structure of TR3, determined at 1.86 Å resolution (Figure 3E), revealed that the relative orientation of the N- and C-terminal domains was restored to that of the wild type, but strand 9 was still in the flipped configuration (Figure 4C). Despite the NG sequence's high propensity to form a two-residue turn, the segment formed a three-residue turn. A two-residue turn would have flipped strand 9 back to the wild-type configuration, which would have pushed strand 9 toward the C-terminus by one residue. The low stability of the TR3 mutant may be attributed to the presence of non-native strand register that includes unfavorable cross-strand interactions between positively charged residues (K112–K125–K135, Figure 4C) and a distorted turn between strands 8 and 9. These results strongly suggest marginal contributions of specific cross-strand amino acid pairings to specifying strand register.

The HSQC spectrum of TR3 retained overall similarity with those of the wild type and TR1, indicating that the global fold was retained in TR3 (Figure 2C). However, we found that the cross-peak for residue 38 was split into two overlapping peaks, indicative of the presence of multiple conformations. The small difference in the position of the two cross-peaks for residue 38 suggests that the difference between the two conformations is subtle. Residue 38 is located in a turn that interacts with β -strand 8 (Figure 2E). Thus, it is likely that the SLB segment in TR3 exists as an ensemble of multiple conformations.

Removal of the Additional Residue Eliminates the β -Strand Flip. To test whether the low stability of TR3 is due to a structural mismatch among the two new turns and the β -strand with one extra residue, we eliminated the extra residue (I126) from TR3. This mutant, termed TR3 Δ 126, contains an IIDGIII sequence in place of IIDGIII (Figure 1C). The X-ray crystal structure at 1.9-Å resolution of TR3 Δ 126 showed that β -strand 9 (residues 120–125) flipped back into the wild-type configuration and the turn between strand 8 and 9 formed the intended type I' turn as observed in TR2 (Figures 3F and 4D). Except for the presence of different turns, the overall structure of TR3 Δ 126 is very similar to the wild-type structure. The rmsd value for equivalent C α atoms between the wild type and TR3 Δ 126 is 0.87 Å. Its NMR structure is consistent with the presence of a single, dominant conformation in solution (Figure 2D).

The stability of TR3 Δ 126 was only marginally improved over that of TR3 (Table 1 and Supplementary Figure 1). This is interesting, considering that TR3 Δ 126 contains the fully restored strand 9 and the turn between strands 8 and 9 is identical to that of TR2. From the stability data of TR2 and TR3 Δ 126, we estimate the cost of replacing the original turn between strands 9 and 10 (FNEKGEVS) with IIDGIII to be 1.5 kcal/mol. From the difference between this value and the stability change of

(30) Ohnishi, S.; Koide, A.; Koide, S. *Protein Sci.* **2001**, *10*, 2083–2092.

(31) Ramirez-Alvarado, M.; Blanco, F. J.; Serrano, L. *Nat. Struct. Biol.* **1996**, *3*, 604–612.

(32) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170–174.

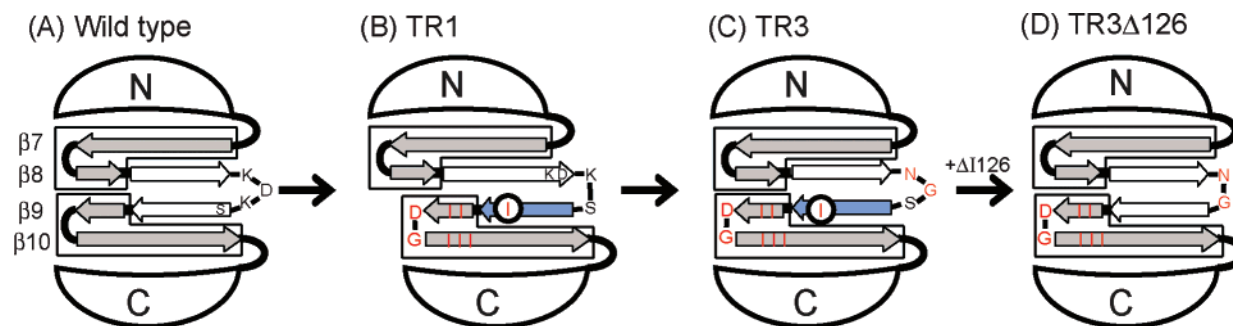


Figure 5. Schematic representation of structural rearrangements triggered by the mutations. The N- and C-terminal globular domains are depicted as half circles. Stable segments within the β -strands 7–10 region are shown as gray arrows and enclosed in solid lines. Mutations are shown in red. In (B) and (C), the flipped β -strand is shown as the blue arrow. The replacement of turn 8/9 with the NG sequence restored the relative orientation between the N- and C-terminal domains, but β -strand 9 remained flipped, as depicted in (C). Finally, the deletion at residue 126 restored the original orientation of β -strand 9 (D).

TR1 (2.5 kcal/mol), we estimated the total cost of inserting an Ile residue at position 126, flipping the entire strand 9, restructuring the turn 8–9, and changing the register between strand 8 and 9 to be only 1.0 kcal/mol.

Residues 117–125 Are a Plastic Segment Responsible for β -Strand Sliding. The conformational rearrangements of the SLB were limited to the close vicinity of strand 9, i.e., turn 8/9, strand 9, and their relative locations with adjacent strands. The conformations of the two β -strands (8 and 10) that directly interact with the globular domains were largely unaffected, indicating that the extensive interactions with the globular domains provide specific stabilization of their respective native conformations. Therefore, one can view the observed conformational rearrangements as movements of a plastic segment (residues 117–133) between two rigid frameworks (Figure 5).

In both TR1 and TR3, the replacement of turn 9/10 with the IIIDGIII sequence resulted in a one-residue shift and flipping of strand 9. This structure is energetically more favorable than others, e.g., one in which strand 9 and turn 8/9 maintain their respective wild-type configurations and the extra Ile residue is somehow accommodated in the 9–10 turn. This is quite surprising, particularly in the case of TR3, where the wild-type conformation of strand 9 would have restored the side-chain interactions involving strand 9 residues as well as the type I' turn in the 8/9 turn. These results indicate that the bulged turn conformation of the IIIDGIII with an Ile at position 126 is highly stable, effectively freezing the conformation of residues 126–133. This view is consistent with the small stability loss (~ 1.5 kcal/mol) caused by this turn replacement. Thus, the presence of this stable structural motif further limited the plastic segment to residues 117–125.

Main-Chain H-Bonds Are a Major Determinant of β -Strand Register. To evaluate factors governing the conformational preference of this plastic segment in TR1, we constructed topological models containing an alternative mode of β -strand pairing (Figure 6) and analyzed their properties. As described in the previous section, only residues 117–125 were regarded plastic; i.e., β -strand sliding occurs only between strands 8 and 9.

We first examined the role of the side-chain pair correlation using the statistical cross-strand pair correlation data.¹¹ Because different sheet pairing patterns in these models resulted in different numbers of cross-strand pairs, we calculated the average of the correlation values for each model, referred to as the “correlation index”, with a higher value indicating a more

frequent pair. Intriguingly, we found that the TR1 crystal structure topology had the smallest correlation index value (1.07) among the models (Table 2). We also evaluated the turn propensity of turn 8/9 of the models. The TR1 crystal structure, Models “–4” and “+4” had higher values (1.51, 1.59, and 1.37, respectively) than the others. Taken together, we found no significant correlation between our results and prediction from statistical data.

We found that the β -strand arrangement of TR1 structure maximized the number of main-chain H-bonds among these models (Figure 6 and Table 2). In the TR1 crystal structure, residues 112–127 form a total of seven cross-strand main-chain H-bonds, while the same segment forms fewer H-bonds in models “–2” (six H-bonds) and “–4” (five H-bonds). In addition to having fewer H-bonds in the interface between strands 8 and 9 (six and five H-bonds for models “+2” and “+4”, respectively), models “+2” and “+4” also have two fewer main-chain H-bonds between β -strands 9 and 10 (six versus eight in the other models, Table 2). Additionally, model “+4” has Thr at the (L2) position (residue 122), where a β -branched is disfavored. Thus, models “+2” and “+4” are likely to represent high-energy structures.

Discussion

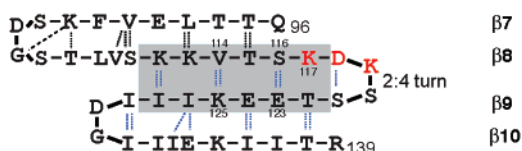
The high-resolution X-ray crystal structures of the series of SLB mutants provided a rare opportunity to observe how a β -sheet adjusts its structure to accommodate mutations. Strikingly, the mutations in turn 9/10 did not simply change the equilibrium between the native state and the unfolded state but rather caused a large structural rearrangement, including a drastic change in the backbone H-bond pattern and incorporation of a turn segment into a β -strand. These results indicate that the mutations stabilized alternative conformations to such a degree that they are more stable than the wild-type native conformation and the unfolded state. A similar case of β -strand inversion upon turn replacement has been reported in the context of a highly flexible β -hairpin peptide derived from ubiquitin,¹⁶ but the dynamic nature of short peptides made it difficult to estimate energy difference between the wild type and alternative conformations.

Mutation-induced conformational changes have been well documented (e.g., domain swapping,³³ strand/helix conversion,³⁴

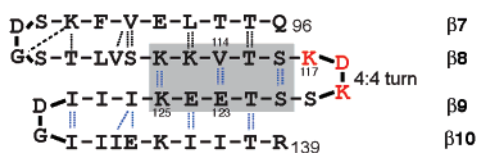
(33) Liu, Y.; Eisenberg, D. *Protein Sci.* **2002**, *11*, 1285–1299.

(34) Cordes, M. H.; Walsh, N. P.; McKnight, C. J.; Sauer, R. T. *Science* **1999**, *284*, 325–328.

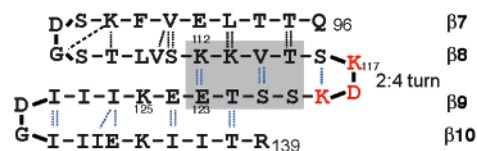
A. TR1 crystal structure



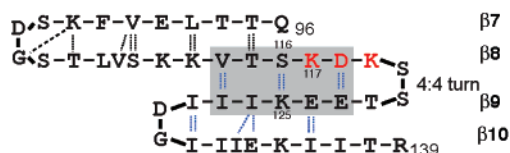
B. Model (-2)



C. Model (-4)



D. Model (+2)



E. Model (+4)

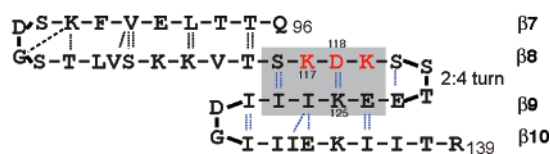


Figure 6. Schematic representations of the SLB segment of the TR1 and four models with an alternative mode of strand register. These models differ in the residue pairing between β -strands 8 and 9. The H-bonding pattern between strands 8 and 9 restricts the turn 8/9 to be either a two-residue β -turn (2:4 turn) or a four-residue β -turn (4:4 turn), but not a three-residue turn that would break all H-bonded pairs between strands 8 and 9. Under these restrictions, we considered four alternative modes of β -strand pairing in addition to that of the TR1 crystal structure. In models “-2” and “-4”, β -strand 9 shifts relative to β -strand 8 in such a way that a given residue forms H-bonds with $i - 2$ and $i - 4$ residues with respect to the i th residue in the TR1 structure, respectively (e.g., E123 forms H-bonds with S116 in TR1, with V114 in model “-2”, and with K112 in model “-4”). Likewise, in models “+2” and “+4”, new H-bonds are formed with $i + 2$ and $i + 4$ residues, respectively. Model “-2” corresponds to the relative position of the N- and C-terminal domains seen in the wild type and also in TR3. Residues 123 and 125 are labeled and serve as reference points. Their H-bonded partners are also labeled.

and helix/loop conversion³⁵). Of particular relevance to the present work, Sagermann et al. showed that duplication of a β -strand-turn segment resulted in a displacement of an adjacent β -strand in a T4 lysozyme β -sheet.³⁵ This somewhat irregular β -sheet consists of three β -strands, each containing only three residues, much smaller than the OspA SLB. Together, these

Table 2. Structural Analysis of TR1 and Models with an Alternative β -Strand Register Shown in Figure 6

	pairwise index ^a	turn index ^b	H-bonds ^c
TR1 crystal structure	1.07	1.51	24 (7, 8)
model “-2”	1.94	0.983	23 (6, 8)
model “-4”	1.80	1.59	22 (5, 8)
model “+2”	1.34	1.15	21 (6, 6)
model “+4”	1.75	1.37	20 (5, 6)
wild type	1.63	— ^d	25 (7, 9)

^a Side-chain pairwise index calculated for residue pairs between β -strands 8 and 9 (enclosed in a gray box in Figure 5). ^b Turn propensity index.

^c Number of cross-strand main-chain hydrogen bonds among β -strands 7–10. The numbers in the parentheses indicate the number of H-bonds between β -strands 8 and 9 and between β -strands 9 and 10, respectively.

^d No statistical propensity is available for the wild type, 3:5 turn.

examples of mutation-induced conformational changes are consistent with the concept of a rugged free energy landscape of proteins with multiple local minima that are energetically close to the native state.

The analysis based on the alternative sliding models of TR1 (Figure 6) strongly suggest that the backbone H-bonds are the dominant factors for the formation of β -strand/ β -strand interactions. This assessment is consistent with a previous backbone perturbation study on the PIN WW domain, demonstrating significant contribution of main-chain H-bonds buried in hydrophobic clusters to β -sheet stability.¹⁵ In addition, hydrophobic burial, which should be correlated with the number of H-bonded pairs, is likely to be important, as Ala scanning mutagenesis studies of the OspA SLB demonstrated the dominant contribution of hydrophobic burial to the stability of the SLB.²⁴

An intriguing feature found in the TR1 structure is the integration into β -strand 8 of the residues that were a part of turn 8/9 in the wild type (Figure 4). This structural transition was partly enabled by the weak turn-forming propensity of the sequence (KDK) of the original turn 8/9. This segment appears “ambivalent” enough to be incorporated in either a turn or a β -strand, reminiscent of the “chameleon” sequence.⁵ When K117-D118-K119 was replaced with an Asn-Gly in TR3, β -strand 8 was terminated at this position and the β -strand register was restored. These results suggest that a strong “ β -strand stop” signal, such as Gly and Pro residues, is important for preventing register sliding of β -strands. In the wild-type OspA, β -strand sliding is prevented by the stable structure of turn 9/10, and the KDK sequence forms a turn, probably passively as a consequence of specific and stable zipping of β -strands 9 and 10 and interactions between this β -hairpin and β -strand 8 that interacts tightly with the N-terminal domain. Consistent with this view, the wild-type structure forms the largest number of backbone H-bonds among the models analyzed in Figure 6 (Table 2). This notion suggests an interesting possibility that the SLB could accommodate a large variety of sequences as long as a strong turn motifs are placed along one edge of the SLB and the number of backbone H-bonds is maintained at the level found for the wild type.

The findings in this work have important implications in protein structure prediction, design, protein fold evolution, and misfolding. The inherent low sequence specificity of β -strand pairing rationalizes the difficulties in predicting and designing predominantly β -sheet structures.¹⁰ External elements such as turns and hydrophobic core packing, rather than amino acid

(35) Sagermann, M.; Gay, L.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9191–9195.

Table 3. X-ray Data Collection and Refinement Statistics for the Crystal Structures of OspA Mutants

protein (PDB code)	TR1 (2OL7)	TR2 (2OL6)	TR3 (2OY1)	TR3/Δ126 (2OL8)
Data Collection Statistics				
space group	$P1$	$P2_1$	$P2_1$	$P2_1$
cell parameters	$a = 37.13$ $b = 55.33$ $c = 64.79$ $\alpha = 85.68$ $\beta = 74.05$ $\gamma = 86.45$	$a = 32.88$ $b = 51.51$ $c = 65.64$ $\beta = 98.56$	$a = 36.05$ $b = 56.01$ $c = 66.20$ $\beta = 97.09$	$a = 33.36$ $b = 53.23$ $c = 65.20$ $\beta = 98.83$
beamline	APS 17-ID	APS 22-ID	APS 23-ID	APS 23-ID
wavelength	1.00000 Å	1.00000 Å	1.0332 Å	0.97934 Å
resolution (Å) (highest resolution shell) ^a	50–1.35 (1.40–1.35)	50–1.60 (1.66–1.60)	50–1.85 (1.92–1.85)	50–1.80 (1.86–1.80)
completeness (%)	95.6 (93.1)	97.0 (95.6)	99.8 (100.0)	99.8 (99.5)
$I/\sigma(I)$	18.87 (4.34)	21.51 (2.96)	13.90 (3.61)	27.43 (4.17)
R_{merge}^b	0.069 (0.324)	0.048 (0.475)	0.084 (0.374)	0.064 (0.335)
average redundancy	3.4 (3.4)	3.7 (3.7)	3.6 (3.4)	4.0 (3.6)
Refinement Statistics				
resolution range (Å)	20.0–1.35	20.0–1.70	20.0–1.86	20.0–1.80
reflections used (free)	97 870 (5139)	22 172 (1191)	20 929 (1398)	19 753 (1071)
R factor ^c	0.184	0.214	0.206	0.240
R_{free}^d	0.218	0.258	0.236	0.290
rms deviations				
bonds (Å)	0.013	0.015	0.023	0.014
angles (°)	1.479	1.611	1.971	1.471
no. protein residues	497	245	247	244
no. waters	585	175	88	51
average B factor (Å ²)	15.29	26.60	39.33	32.95
Ramachandran Plot Statistics				
most favored (%)	91.6	92.8	90.1	93.2
additionally allowed (%)	7.6	6.8	7.2	5.9
generally allowed (%)	0.9	0.5	1.3	0.5

^a Highest resolution shell is shown in parentheses. ^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I(hkl)_i - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I(hkl)_i$ over i observations of a reflection hkl . ^c R -factor = $\sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$. ^d R_{free} is R with 5% of reflections sequestered before refinement.

sequences within β -strands, probably play dominant roles in specifying β -strand register. Though the amphipathic nature of β -strands commonly found in globular proteins may well prevent β -strand flipping due to the high penalty of burying polar side chains, a strand shift by an even number of residues (e.g., +2 and −2) that maintains the amphipathic binary pattern is still possible. Thus, we suggest that successful design of predominantly β -sheet proteins requires implementation of negative design elements against nonspecific β -strand pairing.³⁶ Our results show that register shifts between adjacent β -strands can be triggered by mutations located at a distant site. This type of structural arrangement can create a distinct protein structure with a distinct distribution of functional groups, and thus it could serve as a mechanism to generate a new function. Finally, strand register shifts may occur within a β -rich conformation of a fibril-forming peptide if it lacks a strong β -strand-stop motif. Such shifts could generate distinct conformers and potentially contribute to the so-called prion strain phenomenon.³⁷ The importance of maximizing the number of backbone H-bonds in SLB energetics is consistent with recent models of amyloid core structures where peptide units are aligned in the exact register that maximizes the backbone H-bonds between adjacent units.²

Experimental Section

Mutagenesis and Protein Production. Mutagenesis was performed on the expression vector encoding residues 27–273 of OspA as described previously.²³ The genes for the mutants were subcloned in

the equivalent region of the OspA surface mutant for crystallization (“OspAsm1”)²⁶ using *SpeI* and *PstI* restriction enzyme sites, resulting in 12 surface mutations in total on N- and C-terminal domains. Mutant proteins were expressed in *Escherichia coli* BL21 (DE3) as described before.²⁶ Protein purification using an immobilized metal affinity chromatography column (HisTrap HP, Amersham Biosciences) was performed as described previously,²⁶ except that cation-exchange chromatography step was omitted.

Stability Measurements. Urea-induced unfolding of OspA variants were monitored simultaneously with far-UV circular dichroism and Trp fluorescence as described previously.²³ Mutants studies in this work exhibited a single-step transition that is consistent with the two-state transition. Therefore, we analyzed the unfolding data in terms of the standard two-state model.

Crystallization. The optimal concentration of OspA mutants for crystallization trials was determined using PCT kit (Hampton Research). Crystal screening was performed by the sitting-drop vapor-diffusion method using Hampton Research Crystal Screens I and II and Emerald Biostructures Wizard 1 and 2. Optimization was performed using the hanging drop vapor-diffusion method, where a drop initially contained 1 μ L each of the reservoir solution and a purified OspA mutant (15.8–29.0 mg/mL in 10 mM Tris-HCl pH 8). Crystallization conditions were as follows: 20% PEG3350, 6% isopropanol, 200 mM NaCl, and 100 mM HEPES pH 7.5 for TR1; 32% PEG400, 4% MPD, and 100 mM imidazole pH 7.1 for TR2; 26% PEG3350, and 100 mM imidazole pH 8.0 for TR3; 38% PEG400 and 100 mM imidazole pH 6.45 for TR3Δ126.

X-ray Data Collection and Structure Determination. X-ray diffraction data were collected at the 17-ID, 22-ID, and 23-ID Sectors (Advanced Photon Source at the Argonne National Laboratory). Data collection and processing as well as structural determination were performed in the same manner as for OspA-sm1,²⁶ except that the

(36) Richardson, J. S.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2754–2759.

(37) Krishnan, R.; Lindquist, S. L. *Nature* **2005**, *435*, 765–772.

molecular replacement was performed with OspA-sm1 structure (PDB ID 2G8C) as the search model and final refinement was done using isotropic temperature factors. The asymmetric unit of the TR1 crystal contains two molecules, and the rmsd value between them is 1.020 Å. Molecule A of the TR1 crystal has clearer electron density than that of molecule B; thus, we used the molecule A coordinates for further analysis. The TR2 and TR3 Δ 126 crystals are isomorphous to the OspA-sm1 crystal. Structural superposition was performed with the program LSQKAB.³⁸ PROMOTIF was used to identify secondary structure elements.²⁷ Molecular graphics were generated using PyMol (www.pymol.org). The PDB IDs and crystal statistics for the OspA mutants are summarized in Table 3.

NMR Spectroscopy. ¹⁵N-enriched samples were prepared as described previously.²⁴ Sample concentrations were ~0.2 mM. ¹H, ¹⁵N heteronuclear single-quantum correlation (HSQC) spectra were acquired on a Varian Inova 600 spectrometer using a pulse sequence supplied by the manufacturer. Data were processed and analyzed as described previously.²² In addition, HSQC spectra were collected on the wild type and TR1 that did not contain the surface mutations present in the sm1 protein, so that a greater number of HSQC peaks of the wild-type spectrum could be unambiguously assigned using previously established resonance assignments.³⁹ These spectra were used to map chemical shift perturbations caused by the TR1 mutation (Figure 2E) using a previously described method.⁴⁰

Modeling and Analysis of Alternative Conformations. Side-chain pairwise indexes were calculated by taking the average over the pair

correlation values of Wouters and Curmi¹¹ for cross-strand pairs between β -strands 8 and 9. Turn propensity values were calculated from position-specific statistics of Hutchinson et al.⁴¹

Acknowledgment. We are thankful to Dr. A. Koide for helpful discussions. This work was supported in part by NIH grants R01-GM57215 and U54-GM074946, NSF grant CMMI-0709079, and the University of Chicago Cancer Research Center. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. Use of the IMCA-CAT Sector 17-ID was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with the Center for Advanced Radiation Sources at the University of Chicago. Use of SER-CAT Sector 22-ID was supported by institutions listed at www.ser-cat.org/members.html. GM/CA CAT Sector 23-ID has been funded in whole or in part with federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104).

Supporting Information Available: Supplementary Figure 1, showing urea-induced unfolding curves of OspA variants. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA074252C

(38) Kabsch, W. *Acta Crystallogr A* **1976**, 32, 922–923.

(39) Pham, T.-N.; Koide, S. *J. Biomol. NMR* **1998**, 11, 407–414.

(40) Huang, X.; Yang, X.; Luft, B.; Koide, S. *J. Mol. Biol.* **1998**, 281, 61–67.

(41) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, 3, 2207–2216.