Published in final edited form as:

J Mol Biol. 2007 April 20; 368(1): 230–243.

Hydrophobic Surface Burial is the Major Stability Determinant of a Flat, Single-Layer β -Sheet

Shude Yan^{1,2}, Grzegorz Gawlak¹, Koki Makabe¹, Valentina Tereshko¹, Akiko Koide¹, and Shohei Koide^{1,2,*}

¹ Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, U.S.A.

² Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, U.S.A.

Abstract

Formation of a flat β -sheet is a fundamental event in β -sheet-mediated protein self-assembly. To investigate contributions of various factors to the stability of flat β -sheets, we performed extensive alanine-scanning mutagenesis experiments on the single-layer β -sheet segment of *Borrelia* outer surface protein A (OspA). This β -sheet segment consists of β -strands with highly regular geometries that can serve as a building block for self-assembly. Our Ala-scanning approach is distinct from the conventional host-guest method in that it introduces only conservative, truncation mutations that should minimize structural perturbation. Our results showed very weak correlation with experimental β -sheet propensity scales, statistical β -sheet propensity scales, or cross-strand pairwise correlations. In contrast, our data showed strong positive correlation with the change in buried nonpolar surface area. Polar interactions including prominent Glu-Lys cross-strand pairs marginally contribute to the β -sheet stability. These results were corroborated by results from additional non-Ala mutations. Taken together, these results demonstrate the dominant contribution of nonpolar surface burial to flat β -sheet stability even at solvent-exposed positions. The OspA single-layer β -sheet achieves efficient hydrophobic surface burial without forming a hydrophobic core by a strategic placement of a variety of side chains. These findings further suggest the importance of hydrophobic interactions within a β -sheet layer in peptide self-assembly.

Keywords

protein folding; energetics; hydrophobic effect; secondary structure propensity; X-ray crystallography

Introduction

Amyloid fibrils and other types of peptide self-assemblies are usually mediated by β -sheet formation. The prevalence of peptides and proteins that can be converted into fibrils indicates that filamentous β -rich self-assemblies are a major class of protein structure.¹ While globular proteins and β -rich self-assemblies both contain " β -sheets", the nature of β -sheets in these classes may be distinctly different. Recent high-resolution structural models of fibrils revealed that the β -sheets in fibrils are very flat.² In contrast, the backbone geometry of residues in

*Corresponding author: Shohei Koide Department of Biochemistry and Molecular Biology, The University of Chicago 920 E. 58th Street, Chicago, IL 60637, U.S.A. Fax: 1-773-702-0439 E-mail: skoide@uchicago.edu

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

natural β -sheet proteins is diffusely distributed in the Ramachandran plot.^{3,4} Furthermore, β -Sheets in water-soluble proteins and in designed β -sheet-peptides, particularly the edge β -strands, are highly distorted. Features such as strong twist and β -bulge are common in β -sheet in globular proteins. Such distortions create multiple negative-design elements against the formation of a continuous β -sheet.⁵ Thus, although there are a large number of studies on side chain contributions to β -sheet energetics in globular proteins, it is not clear how accurately we can extrapolate these results to the side chain energetics of flatter, contiguous β -sheets.

It is very difficult to perform quantitative mutation studies on peptide self-assembly, because the stoichiometry of the reaction is usually unknown and it is nearly impossible to confirm that the mutation of interest does not significantly alter the folded (or assembled) structure. While systematic mutagenesis experiments have yielded important insights into the energetics of β -rich peptide self-assembly,⁶⁻⁸ it is difficult to interpret such data in terms of molecular interactions at an atomic level. Therefore, a model system that permits quantitative thermodynamic measurements and high-resolution structural characterization of β -rich peptide self-assembly would significantly advance our understanding of this class of protein structure.

In this work, we used outer surface protein A (OspA) as a model for investigating side chain contributions to the energetics of flat β -sheets. OspA is a surface lipoprotein of the Lyme disease spirochete, *Borrelia burgdorferi*. The crystal structure of a 28-KDa soluble form of OspA revealed an unusual single-layer β -sheet (SLB) segment that connects two globular domains (Fig. 1A).⁹ Unlike amphipathic β -sheets commonly found in globular proteins, both faces of this single-layer β -sheet are exposed to the solvent, and thus residues in this sheet are not involved in extensive interactions with a hydrophobic core.^{9,10} Nevertheless, this single-layer β -sheet is rigid and highly stable in the context of OspA.¹⁰⁻¹²

The SLB region consists of two homologous β -hairpin units (β -strands 7–8 and β -strands 9–10). As the SLB is "capped" with the globular domains on both ends, the β -strand regions of the SLB are all internal and exhibit an unusually flat geometry. The repeat architecture and the flat geometry suggest that the β -hairpins in their native conformation are poised for self-assembly. Indeed, we proved this notion by demonstrating that the SLB segment can be stably extended by adding copies of the β -hairpin unit^{13,14} and that a peptide corresponding to the β -hairpin unit forms fibrils.¹⁵ Therefore, the SLB can be considered copies of a self-assembling peptide captured within a water-soluble protein.

The OspA SLB has good attributes as a platform for mutation analysis. The single-layer structure of this β -sheet affords a large number of solvent exposed sites for mutation analysis. Each β -strand residue forms two main-chain H-bonds (Fig. 1B), and thus their backbone conformation should undergo only small changes upon mutation than that of a residue on an edge β -strand, reducing errors in modeling the structures of mutants. We have established methods for characterizing mutation effects on thermodynamics and the folding mechanism of OspA.¹⁶ Recently, we have developed a set of surface mutations in the globular domains that greatly facilitates crystallization, which resulted in a 1.15-Å x-ray crystal structure.¹⁷ Taken together, OspA represents a unique model system for investigating the energetics of the formation of flat, nonglobular β -sheets.

In this work, we performed Ala-scanning mutagenesis in the SLB segment in order to determine each residue's contribution to the β -sheet stability and to elucidate relative contributions of different factors. In our Ala-scanning approach, we aimed to minimize the influence of context associated with a particular host position by distributing mutation sites over many sites with similar backbone geometry. We also aimed to minimize structural perturbation by restricting the type of mutation to Ala-replacement and other conservative mutations so that we could model mutant structures with reasonable accuracy. We exploit the natural amino acid diversity

among the mutation sites to produce a sufficient variety of mutation types. The β -strand positions of the SLB are rich in amino acids that are generally ranked high in β -strand propensity scales (e.g. Thr, Ser, Val and Ile; Fig. 1B) and also in polar and charged residues and contain a large number of cross-strand Glu-Lys pairs that have been shown to have a high interaction energy.¹⁸ This β -sheet also buries hydrophobic surface areas to an extent comparable to that of small globular proteins.¹⁰ Taken together, this amino acid diversity allows us to address contributions of different factors.

Here, we analyzed the effects of 40 single-point mutants on stability. Unexpectedly, the stability changes showed no significant correlation with β -sheet propensity scales derived from other experimental systems or with statistical cross-strand pairwise correlations. In contrast, we found that the stability changes strongly correlate with hydrophobic surface burial of the mutated residues. We also present a crystal structure of the most destabilized mutant, which defines an extent of conformational perturbations by the mutations used in this work. We discuss the design features of the OspA single-layer β -sheet and implications of our findings for the formation of β -rich self-assembly, misfolding and design.

Results

Characterization of alanine-substitution mutants

We performed alanine-scanning mutagenesis at a total of 31 positions throughout the SLB region (Tables 1 and 2). All mutants were expressed as soluble proteins and monomeric. For each mutant, the free energy change for unfolding was determined from urea-induced unfolding reactions. All proteins studied here underwent the characteristic two-step (i.e. three-state) unfolding reactions (Supplementary Figure 1).¹⁶ A mutation in the SLB region can affect either of the two unfolding transitions (native to intermediate and intermediate to unfolded) and shift the relative placement of the intermediate state in the reaction coordinate due to the equilibrium between two unresolved species within the intermediate state that differ in the nature of the SLB region.¹⁶ Thus, the total effect of a mutation is manifested in the change in the free energy difference between the fully folded and fully unfolded states (ΔG_{NU} in Table 1). The effect of a mutation is expressed in terms of the change in ΔG_{NU} and designated as $\Delta\Delta G$ ($=\Delta G_{\text{NU}}^{\text{wild-type}} - \Delta G_{\text{NU}}^{\text{mutant}}$) (Tables 1 and 2).

The far-UV circular dichroism spectra of these proteins in the fully denatured state were essentially identical (Supplementary Figure 2). Also the ^1H , ^{15}N -HSQC spectra of the unfolded state of representative mutants showed no more than three peaks that significantly shifted compared with the wild-type spectra (Supplementary Figure 3). These shifted HSQC peaks are most likely to be for the mutated residues and its immediately preceding and following residues. These spectra strongly suggest that the mutations do not significantly affect the unfolded state of OspA and that there is no significant residual structure in the unfolded state. The far-UV CD spectra of the mutants in the native state were essentially identical (Supplementary Figure 2). The HSQC spectra for the native state had larger number of shifted peaks than the unfolded state spectra, but most shifts were very small, typically observed for subtle conformational changes (Supplementary Figure 3). These spectral data as well as their consistent unfolding curves indicate that the mutants are all fully folded in the absence of urea and that the point mutations caused only small, local conformational changes in the native state.

We first examined the residues in the β -strand regions of the SLB (Fig. 2). All hydrophobic residues (Phe, Leu, Val and Ile) have significant contributions, particularly Phe and Ile. Contributions of polar (Ser and Thr) and charged (Glu and Lys) residues are highly uneven, from negligible (e.g., T115) to almost 2 kcal/mol (e.g., K113). In the following sections, we examine contributions of different factors to the observed stability changes.

Contribution of β -sheet propensity

We first compared our data with β -sheet propensity scales. There are a number of experimental β -sheet propensity scales available in the literature, and we used the average of four scales that are derived from host-guest studies.¹⁹⁻²² No significant correlation was found ($r = 0.01$; Fig 3A). We also found little correlation ($r = 0.3$) between the data and a statistical β -sheet propensity scale by Chou and Fasman (data not shown).^{23,24} This little dependence of the $\Delta\Delta G$ values with β -sheet propensity is evident from the observation that residues of the same amino-acid type (i.e. four Ser, four Glu, four Lys, and three Thr) exhibit significantly different $\Delta\Delta G$ values (Fig. 3A).

Contributions of cross-strand pairwise interactions

We then examined correlation between our results with cross-strand statistical pair correlation data.²⁵ We used the statistical correlation data because experimental thermodynamic data is available only for a small number of cross-strand pair types.¹⁸ No significant overall correlation was found ($r = -0.22$; Fig 3B). However, the pair correlation energy for E100, K112 and E124 was significantly greater than those for the other residues and are close to the experimental $\Delta\Delta G$ values. These three residues are a part of a long cross-strand E-K ladder (E77-K87-E100-K112-E124-K135; Fig 1). These results suggest that cross-strand pairwise interactions may have significant contribution for residues in a long cross-strand E-K ladder but not for a single E-K cross-strand pair (K113-E123; Figs. 1A) or the other types of pairs tested here.

Contribution of surface burial

We first examined burial of nonpolar solvent accessible surface area ($\Delta \text{SASA}_{\text{np}}$), which confers the favorable hydrophobic effect. The minimal effects of mutations on the NMR and CD spectra described above strongly suggest that the structural effects of the mutations in the unfolded state are confined in the immediate vicinity of the mutated position and consequently most of the observed thermodynamic effects of the mutations are due to changes in the folded state. Therefore, in the following analysis we will rationalize the results in terms of structural properties of the folded state. The buried hydrophobic surface area shows significant correlation with the $\Delta\Delta G$ values ($r = 0.74$; Fig 3C). In addition to rationalizing the differences among nonpolar residues, the degrees of nonpolar surface burial partially explain $\Delta\Delta G$ values of polar and charged residues (e.g., the significantly larger contributions of K112 and K113 than that of K125). This strong correlation suggests that hydrophobic surface burial is the dominant stability determinant for residues in the single-layer β -sheet.

We next examined the effect of polar surface burial. A weak negative correlation was found ($r = -0.55$; Fig 3D). The negative correlation is consistent with the generally unfavorable nature of the desolvation of polar surface²⁶⁻²⁸. The desolvation penalty seems to explain why K135 contributes the least among the four Lys residues and why T115 contributes significantly less than T122, although they have similar amount of buried hydrophobic surfaces. E100 and E124 showed significantly greater contributions than expected from the overall correlation, and K112 and K113 also showed slightly higher-than-expected contributions. These results, together with the unusually high pairwise interaction energies of E100, K112 and E124 (Fig 3B) suggest that burial of complementary polar surface (as in a Glu-Lys pair) may contribute favorably to β -sheet stability (see below).

Contributions of turn and bulge residues

In addition to β -strand residues, we examined the side chain contributions of residues in turns (K117-K119 and N127-E131) and in β -bulge positions (V110 and S133) by alanine substitution (Tables 1 and 2). Most turn residues showed insignificant contributions, not unexpected from

the high degree of solvent exposure of their side chains and from small numbers of interactions involving them. N127 and D118 showed exceptionally large contributions (2.54 and 1.99 kcal/mol, respectively), probably because the side chains of these two residues form hydrogen-bonds with backbone atoms (D118 O^ε with K107 NH, and N127 O^ε with NH of residues 129–131). We found excellent correlation between our experimental values and the statistical propensities for turn and bulge ($r = 0.96$) of Thornton and coworkers,^{29,30} with the exception of E131 (Fig. 3E). O^ε of E131 forms a hydrogen-bond with NH of N127, which would not be accounted for by the propensity scale. This strong correlation suggests that the stability contributions of turn/bulge residues predominantly arise from intrinsic propensities.

Mutations for testing the contributions of cross-strand Glu-Lys pairs

The Glu-Lys pairs in the "Glu-Lys ladder" were outliers in the pair correlation and polar surface burial analyses above (Fig 3B and D). The Glu-Lys cross-strand pair has been shown to have high statistical correlation²⁵ and also high degrees of side-chain interaction energy as measured experimentally.¹⁸ It has been suggested that interaction between Glu and Lys may be cooperatively enhanced by forming a linear ladder or a network³¹. However, the energetic advantage of Glu-Lys pairs is not supported by theoretical studies^{26–28}, because of high desolvation penalty and a large loss of side-chain entropy.

To investigate the energetic effect of Glu-Lys pairs, we made non-alanine mutations (K112M, E124L, S111E and I136K). The K-to-M and E-to-L mutations each remove the charges but minimally alter the side chain size, and thus they should be better suited than the K-to-A and E-to-A mutations for specifically testing the role of electrostatic interactions. Interestingly, K112M and E124L had opposite effects, with K112M destabilizing and E124L stabilizing, although each of these mutations removed two cross-strand charge-charge interactions (K112 with E100 and E124, and E124 with K112 and K135). This difference can be rationalized by different degrees of desolvation of the Glu and Lys side chains. Because the Glu side chain is shorter than that of Lys, the carboxyl group of a Glu in a Glu-Lys ladder is much more desolvated than the amino group of a Lys (compare E100 and E124 with K112 and K135; Fig 3D and Table 2). Although the E124L mutation removes the favorable charge-charge interactions, it also removes the unfavorable desolvation penalty of the E124 carboxyl group and increases nonpolar burial, while it marginally affects desolvation of adjacent Lys amino groups (Fig. 1A and B). In contrast, because its terminal amine is highly exposed, the K112M mutation removes charge-charge interactions but only much smaller desolvation penalty of K112 and it does not significantly increase nonpolar burial, while leaving E100 and E124 still highly desolvated. Thus, the net effect of the K-to-M mutation in a Glu-Lys ladder is much more unfavorable than the E-to-L mutation.

Substitutions of a charged residue with an isosteric polar residue, E123Q and E124Q, were both destabilizing. This is consistent with the above analysis because these mutations remove favorable charge-charge interactions but still bury a similar amount of polar surface. Taken together, these results suggest that the Glu-Lys pairs in the single-layer β -sheet have significantly favorable electrostatic interactions that are however offset by desolvation penalty for polar surfaces. The data also demonstrates that replacing the buried polar surface in a Glu-Lys pair with a similar amount of hydrophobic surface is stabilizing.

We also tested mutations that introduce Glu-Lys pairs in place of nonpolar or neutral side chains. These are complementary to the disruption of Glu-Lys pairs described above. I136K, which could potentially create charge-charge interactions with E123, E134 and E146 (Fig 1), was strongly destabilizing, supporting a lack of energetic advantage of cross-strand Glu-Lys interactions compared with hydrophobic surface burial. S111E, which was intended to create two potential charge-charge interactions (with K113 and K125), was also destabilizing, further demonstrating the difficulty in creating favorable charge-charge interactions.

We further assessed the contributions of electrostatic interactions with the program *Delphi* that solves the Poisson-Boltzmann equation with a finite difference method (Table 1). The calculations predicted that all charged residues had large columbic energy and solvation energy of opposite signs, which were almost completely balanced and resulted in small net electrostatic energy ($\Delta\Delta E_{\text{delphi}}$). Except for N127, the $\Delta\Delta E_{\text{delphi}}$ values were generally unfavorable (i.e. negative) even for the three residues located in the Glu-Lys ladder (E100, K112 and E124). The calculations correctly predicted the destabilization effect of K122M, E123Q and E124Q, and the stabilization effect of E124L but not the stabilization effect of T122V (Table 2). Although these results may not be taken quantitatively because of the inaccuracy of the structural models for the mutants and other limitation of the method,³²⁻³⁴ they suggest that electrostatics interactions even on the solvent-exposed β -sheet surface may be generally unfavorable compared with hydrophobic interactions with equal amount of surface burial, consistent with conclusions by Tidor and coworkers.^{27,28,35}

Polar-to-nonpolar surface mutations stabilizes the β -sheet

To further test the dominance of hydrophobic burial at solvent-exposed positions in a β -sheet, we examined a series of polar-to-nonpolar substitutions. T122V and T122Y stabilized the β -sheet by 1.3 and 0.69 kcal/mol, respectively. A polar-to-polar mutation, S121T, had negligible effect, although the mutation adds a methyl group and Thr has greater β -sheet propensity than Ser (Table 2).

Crystal structure of the F126A mutant

To define the extent of structural changes by the mutations, we determined the X-ray crystal structure of F126A, the most destabilizing Ala-scanning mutant, at a resolution of 1.20 Å. The data collection and refinement statistics are summarized in Table 3. Crystallization was facilitated by a set of surface mutations (denoted "sm1") that we developed recently.¹⁷ All the surface mutations are located in the globular domains, distant from the SLB, and thus they should have little effect on the conformation of the SLB region. Also a comparison of a series of OspA variant crystal structures containing these surface mutations suggests that the surface mutations do not impose a single mode of crystal packing. The crystals of F126A mutant and wild-type-sm1 proteins are isomorphous and belong to monoclinic P2₁ space group with one protein molecule in asymmetric unit. The RMS deviation of the C α atoms common in F126-sm1 and wild-type-sm1 is 1.153 Å, indicating there are small differences between these structures (Fig. 4A, B). The F126A mutation triggered structural rearrangements around the mutational site (Fig. 4C). The neighboring residues of Ala126 repack to reduce the cavity caused by the mutation. These rearrangements accompany bending of the SLB, leading to the poor overall fit of the wild-type and mutant proteins. When each globular domain was separately compared, the RMSD values were smaller (the N-terminal domain (residues 28–94), 0.711 Å; the C-terminal domain (141–273), 0.685 Å). The structural rearrangements involved not only the residues in the close vicinity of the mutation site (i.e. the left side of the molecule as shown in Fig. 4A) but also residues on the other edge of the SLB (Fig. 4A and B; Supplementary Movie), indicating the presence of an extensive cooperative network throughout the SLB region.

Despite the rearrangements of residues, the F126A mutation reduced nonpolar surface burial by 98 Å². This actual change corresponds to 56% of the value predicted from modeling using the wild-type structure (Table 1). While conformational adjustments undoubtedly limits the accuracy of parameters derived from model structures, the strong correlation between nonpolar burial and stability change would hold even if we were to take effects of conformational changes into account, because of the following reason. We expect that the inaccuracy in our prediction of buried surface areas is approximately proportional to the size of the removed side chain. In the comparison of nonpolar surface burial and experimental stability changes, residues

with a large side chain are generally positioned above the regression line (Fig. 3C). Thus, a reduction of the nonpolar burial in the actual structure might bring these data points closer to the regression line and potentially improve the correlation.

Discussion

Side chain-dependent determinants of flat β -sheet stability

Our scanning mutagenesis study provides a large data set for systematically evaluating contributions of different factors to flat β -sheet stability. The poor correlation between β -sheet propensity and our experimental $\Delta\Delta G$ values suggest that β -sheet propensity is not a dominant factor (Fig 3A). This conclusion is consistent with the large discrepancies among different experimental and statistical β -sheet propensity scales.^{19-24,36,37} Our results also suggest that cross-strand side-chain interactions are not a dominant factor for β -sheet stability,²⁵ except for the Glu-Lys pairs in the extended E-K ladder. This is consistent with the fact that pairwise interaction energies exhibit large dependence on local environments of a cross-strand pair in a protein,³⁸ and that cross-strand pairs are not highly conserved in protein evolution³⁹.

The strong correlation between our $\Delta\Delta G$ values and hydrophobic surface burial of mutated residues (Fig 3C) suggests that hydrophobic surface burial is the dominant factor for β -sheet stability. The conclusion is consistent with the finding by Minor and Kim²² that the discrepancies between two β -sheet propensity scales correlate well with amino-acid hydrophobicity. Our results also suggest that specific side-chain pattern (i.e., all surrounding residues for a given position) is responsible for burying hydrophobic surface. This is consistent with the finding that amino-acid pairs in the same strand and those diagonally cross-strand also showed significant statistical correlation⁴⁰ and they exhibited significant interaction energies in experimental studies.⁴¹ Interestingly, our results are in an apparent disagreement with results of Distefano et al. that revealed little effect of cross-strand interactions.⁴² However, a closer inspection suggests that this disagreement is due to a drastic difference in the structural contexts between their and our systems. The two positions (residues 44 and 53) that are systematically mutated in Distefano et al. are an H-bonded pair in which the side chains point away from each other. Residue 44 is located on the edge strand and residue 53's other cross-strand neighbor (Ile6) was mutated to Ala. Thus, their system should allow large degrees of structural relaxation of the mutated side chains. In contrast, we did not trim side chains around positions of interest. Taken together, our results suggest that the hydrophobic effect is not only the dominant driving force for the folding of globular proteins with a hydrophobic core^{43,44}, but also the driving force for the formation of protein structural elements such as SLBs.

Design features of the OspA single-layer β -sheet

Our mutation study revealed the dominant contribution of nonpolar surface burial to the SLB stability. We previously found that the OspA SLB buries hydrophobic surfaces to a degree similar to those found for common globular proteins.¹⁰ Together, these results indicate nonpolar burial to be a dominant factor for the positive design of the SLB.

Here, we examine how the SLB achieves efficient hydrophobic surface burial without forming a hydrophobic core. Fig 1B displays side-chain conformations of one face of the single-layer β -sheet. Residues in this face generally had larger stability contributions than those on the other face (Fig 2). F102 and F126, the two residues with the largest stability contributions, are almost completely buried in a hydrophobic cross-strand ladder consisting of F102, L109, F126, and V132.¹⁰ Their tight packing is facilitated by the β -bulges at positions 110 and 133 (Fig 1). The backbone of these two turns (residues K103 to T108 and N127 to E131) wraps around these hydrophobic side chains. Thus, this hydrophobic ladder is strategically placed next to the turns to cooperatively achieve efficient surface burial. The Glu-Lys ladder (E100-K112-E124-K135)

is located adjacent to the hydrophobic ladder, and the methylene groups of these residues closely interact with the hydrophobic array. A ladder of β -branched residues (T98, V114, T122, I137) is placed adjacent to the Glu-Lys array and all of their γ -methyl groups point toward the methylene groups of the adjacent Glu-Lys ladder, further enhancing the total hydrophobic surface burial. Thus the SLB structure achieves efficient hydrophobic surface burial by a strategic placement of amino-acid side chains.

The contributions of the side chains of the two faces are highly uneven, where the "front" face residues provide significantly more stabilization than the "back" face residues (Fig. 2). Consistent with the dominance of hydrophobic surface burial, the residues in the back face are less densely packed than the front face residues (Fig. 1C and D).

Although the dominant importance of nonpolar burial in the OspA SLB energetics implies that loading up the β -strand regions with nonpolar amino acids would stabilize the sheet, the OspA SLB is rich in polar and charged residues (Fig. 1B). Thus we initially expected that interactions between polar side chains might be critical for OspA stability, particularly the Glu-Lys pairs. However, the current results suggest that this is not the case. Both the experimental and computational results suggest Glu-Lys pairs in a β -sheet are generally unfavorable compared with hydrophobic interactions burying a similar amount of total surface^{26,27}. Then, what is the role of the abundant Glu and Lys residues in the single-layer β -sheet? We speculate that the Glu and Lys residues may be important for negative design.

First, the charged residues must be critical for defining the solvent exposed nature of the single-layer β -sheet. A β -strand consisting of predominantly nonpolar amino acids would also stabilize fibril-like structures and might drive the protein into fibrils or aggregates.^{45,46} As charged amino acids effectively inhibit fibril formation, the high-density of Glu and Lys in the SLB should ensure the formation of the single-layer architecture. Second, the charged amino acids may be important for specifying correct β -sheet register. For example, shifting strand 9 toward the N-terminus by one residue with respect to the other strands would create electrostatic repulsions among K112, K125 and K135, and shifting strand 9 toward the C-terminus by one residue would create repulsion between E124 and E134.^{15,47} Our results are consistent with conclusions from theoretical studies that salt bridges (ion pairs) in proteins are generally not for thermodynamic stability but for conformational specificity.^{26,27} Third, they may be important for preventing protein-protein interactions. In its native environment OspA is on the surface of the spirochete and under constant immune surveillance. An "entropy shield" by the long chains and of Lys and Glu and the desolvation penalties of their terminal charges collectively prevent binding of immunological molecules such as antibodies to OspA.^{48,49} Taken together, the abundance of Glu and Lys may be the consequence of negative design against aggregation and protein binding. Residues in the single-layer β -sheet are highly conserved,⁹ suggesting that the specific amino acid pattern in this sheet may have been selected in evolution to satisfy the positive and negative design elements described above.

Implications for β -sheet protein design and β -rich peptide self-assembly

Our result may have important implications for rationalizing factors determining the tendency of natural proteins to form amyloid-like aggregates.¹ Analysis of natural amyloidogenic sequences and systematic mutagenesis studies have established the importance of hydrophobicity, β -sheet propensity and charges in amyloidogenesis.^{6,7} DuBay et al. found hydrophobicity to exhibit the strongest correlation with the aggregation rates among parameters tested in regression analysis.⁴⁵ Hydrophobic amino acids are consistently ranked high for cross- β aggregation in systematic mutation studies by Lopez de la Paz et al.^{6,50} The importance of hydrophobicity is generally interpreted as the driving force for β -sheet lamination that creates a "dry interface" between adjacent β -sheets.⁵¹ Our finding of the dominant role of

nonpolar surface burial for *single-layer* β -sheet stability suggest that nonpolar surface burial is also critical for the formation of each layer of β -sheets.

Although the final, stable fibrils contain multiple β -sheet layers, interactions within a single-layer sheet are still present in such multi-layer structures. Flat, monolayer β -sheets have been implicated as intermediates in fibril formation. Using computational simulations, Nussinov et al. demonstrated that a SLB conformation of peptide self-assembly is stably maintained, and they propose the possibility of SLB structures being intermediates ("protofilaments") in fibril formation.^{52,53} The initial step of β -2-microglobulin self-assembly is the formation of an anti-parallel SLB between two monomers.⁵⁴ Thus, we suggest that interactions of peptide monomers within a single-layer β -sheet are important during the nucleation and also during lateral propagation of a β -sheet layer, in which a peptide unit is added to an existing β -sheet structure. The dominant importance of nonpolar burial in the single-layer β -sheet stability found in this work strongly suggests two complementary roles of hydrophobicity in fibril formation: stabilization of single-layer β -sheet formation and stabilization of β -sheet lamination.

Our results also have implications in protein design. Design of predominantly β -sheet proteins remains a great challenge,⁵⁵⁻⁵⁷ probably due to inadequate understanding of β -sheet energetics and the strong tendency of designed proteins to aggregate. Our results suggest that finding a right balance between stabilizing the target β -sheet structure and destabilizing flat, contiguous β -sheet(s) through efficient nonpolar surface burial may be the key to successful design of β -sheet proteins that are water-soluble and stable. Our results also suggest that proper utilization of charged residues is critically important for negative design against aggregation.^{5,58} The recent atomic models of fibrils and a series of extended OspA SLB structures¹⁴ may serve as useful templates for negative design.

Methods

Protein preparation and stability measurement

Methods for site-directed mutagenesis and protein preparation have been described previously.¹⁶ Mutations at positions 100, 102 and 109–116, which we predicted to reduce the separation between the two unfolding transitions, were made in addition to the E160L mutation in order to obtain a better separation of the two unfolding transitions and thus greater accuracy in $\Delta\Delta G$ values.¹⁶ Urea-induced unfolding reactions were monitored simultaneously with far-UV CD and Trp fluorescence, and unfolding curves were analyzed in terms of a three-state model ($N - I - U$) using a global fitting method as described previously.¹⁶ The m -values for the $N - I$ and $I - U$ transitions (m_{NI} and m_{IU}) were allowed to vary, but the total m -value ($m_{NU} = m_{NI} + m_{IU}$) was fixed at the average value (4.4).¹⁶ The use of the constant m_{NU} value is justified because the m -value is proportional to the total change in surface burial upon folding⁵⁹ and the spectral data (see Results) indicates minimal effects of the point mutations on the conformations of the native and unfolded states of OspA, suggesting the conservation of the amount of surface burial upon folding. This method is equivalent to the commonly used one for comparing mutants of protein that undergoes a two-state unfolding in which a single, average m -value is used for obtaining $\Delta\Delta G$ values (e.g. refs ⁶⁰ and ⁶¹). Stability data for the wild type and 19 mutants were taken from Yan et al,¹⁶ which had been determined in the same manner as described above.

Conversions of statistical correlations

We converted statistical correlations into pseudo-energy terms by applying the Boltzman equation. A residue in the β -sheet has two cross-strand neighbors (e.g. E124 with K112 (H-

bonded pair) and K135 (non-H-bonded pair); Fig. 1A), and thus a mutation alters two cross-strand pair correlations. Accordingly, we define the cross-strand pair correlation energy as

$$\Delta\Delta G_{\text{pair}} = RT\ln(g_{ij-H}^{\text{WT}}/g_{ij-H}^{\text{mut}}) + RT\ln(g_{jk-\text{nonH}}^{\text{WT}}/g_{jk-\text{nonH}}^{\text{mut}}),$$

where g_{ij-H}^{WT} and $g_{jk-\text{nonH}}^{\text{WT}}$ are statistical correlations for hydrogen-bonded and non-hydrogen-bonded pairs of the wild type, respectively, and g_{ij-H}^{mut} and $g_{jk-\text{nonH}}^{\text{mut}}$ are the equivalent terms for the mutant.

Likewise, we converted the statistical probability for the turn and bulge by the Boltzman equation, $\Delta\Delta G_{\text{turn/bulge}} = RT\ln(P^{\text{WT}}/P^{\text{mut}})$, where P^{WT} is the probability of the wild type residue and P^{mut} is that of the mutant.

Structural models of mutant proteins

The atomic coordinates of alanine mutants were generated from the wild-type OspA structure (1OSP chain O) by truncating the side chain of the mutated residue beyond C_{β} . Those of non-alanine mutants were made by replacing the mutated residue and choosing the side-chain rotamer that is the closest to that of the wild-type residue from a backbone-dependent rotamer library⁶² using the Swiss-PDBViewer program.

Surface area calculation

Solvent accessible surface areas (SASA) of the folded state were calculated using the CNS program with a probe radius of 1.4Å.^{63,64} SASA for residues in a random coil conformation was from Lesser and Rose.⁶⁵ The atomic coordinates for the mutants were generated as described above. Buried surface areas for each atom of a residue were calculated by subtracting its SASA value in the folded structure from its respective the random coil value. Buried hydrophobic and polar surface areas, $\Delta\text{SASA}_{\text{hp}}$ and $\Delta\text{SASA}_{\text{pol}}$, were the total buried surface areas of aliphatic and aromatic carbons and those of the other types of atoms, respectively.

Electrostatics calculation with Delphi

Linearized Poisson-Boltzmann equation was solved with the *Delphi* program (version 4).⁶⁶ The unfolded state is modeled as an extended Gly-X-Gly peptide, assuming minimal side-chain interactions in the unfolded state.²⁷ ΔE_{delphi} and $E_{\text{solvation}}$ are defined as $\Delta E_{\text{delphi}} = E_{\text{coulombic}} + E_{\text{solvation}}$, and $E_{\text{solvation}} = E_{\text{crf}} + E_{\text{ion}}$. $E_{\text{coulombic}}$ and E_{crf} (corrected reaction field energy) are obtained from the *Delphi* run with ionic strength (**I**) of 0 and E_{ion} is the difference between the total grid energy from a *Delphi* run with **I** = 50 mM and that with **I** = 0. The final electrostatics contribution, $\Delta\Delta E_{\text{delphi}}$, was calculated with $\Delta\Delta E_{\text{delphi}} = \Delta E_{\text{delphi}}$ (mutant, folded) - ΔE_{delphi} (mutant, unfolded) - ΔE_{delphi} (WT, folded) + ΔE_{delphi} (WT, unfolded). The parameter settings were as follows: grid size = 165, grid scale = 1.25, internal dielectric constant = 4, external dielectric constant = 80, probe radius = 1.4, and the Debye-Hückel boundary condition. The results showed little dependence on the grid size and scale setting (data not shown).

X-ray crystallography—The gene for the F126A mutant was subcloned in the equivalent region of the OspA surface mutant for crystallization ("OspAsm1")¹⁷ using *SpeI* and *PstI* restriction enzyme sites, resulting twelve surface mutations in total. The mutant was expressed and purified as described before.¹⁷ The crystals were grown in 34% PEG400, 0.1M Tris-HCl pH9.0 using the hanging drop vapor diffusion method. The X-ray diffraction data were collected at the 19-BM beamline (the Advanced Photon Source at the Argonne National Laboratory). Crystal data, data collection statistics and PDB IDs are summarized in Table 2. X-ray diffraction data were processed with HKL2000.⁶⁷ The structure was determined by

molecular replacement with the program MOLREP in CCP4.⁶⁸ The "wild-type" OspA-sm1 structure (PDBID: 2G8C) was used as the search models. CNS1.1⁶³ and Refmac5⁶⁹ were used for the structural refinement, and final positional and anisotropic temperature factor refinement was performed using Refmac5. Model building was carried out using the Coot program.⁷⁰ Molecular graphics were generated using Pymol (<http://www.pymol.org>). Crystal data, data collection and refinement statistics are summarized in Table 3. The coordinates and structure factors for the F126A-sm1 mutant have been deposited in the Protein Databank (ID 2I5Z).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are thankful to Lin Silver for technical assistance. This work was supported in part by NIH grants R01-GM57215 and U54-GM074946. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Biological and Environmental Research, under Contract No. W-31-109-ENG-38.

References

1. Dobson CM. Protein folding and misfolding. *Nature* 2003;426:884–90. [PubMed: 14685248]
2. Nelson R, Eisenberg D. Recent atomic models of amyloid fibril structure. *Curr Opin Struct Biol* 2006;16:260–5. [PubMed: 16563741]
3. Serrano L. The relationship between sequence and structure in elementary folding units. *Adv Protein Chem* 2000;53:49–85. [PubMed: 10751943]
4. Jha AK, Colubri A, Zaman MH, Koide S, Sosnick TR, Freed KF. Helix, sheet, and polyproline II frequencies and strong nearest neighbor effects in a restricted coil library. *Biochemistry* 2005;44:9691–702. [PubMed: 16008354]
5. Richardson JS, Richardson DC. Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc Natl Acad Sci U S A* 2002;99:2754–9. [PubMed: 11880627]
6. Lopez de la Paz M, Serrano L. Sequence determinants of amyloid fibril formation. *Proc Natl Acad Sci U S A* 2004;101:87–92. [PubMed: 14691246]
7. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 2003;424:805–8. [PubMed: 12917692]
8. Williams AD, Shivaprasad S, Wetzel R. Alanine scanning mutagenesis of Abeta(1–40) amyloid fibril stability. *J Mol Biol* 2006;357:1283–94. [PubMed: 16476445]
9. Li H, Dunn JJ, Luft BJ, Lawson CL. Crystal structure of Lyme disease antigen outer surface protein A complexed with an Fab. *Proc Natl Acad Sci USA* 1997;94:3584–3589. [PubMed: 9108020]
10. Pham T-N, Koide A, Koide S. A stable single-layer β -sheet without a hydrophobic core. *Nat Struct Biol* 1998;5:115–119. [PubMed: 9461076]
11. Koide S, Bu Z, Risal D, Pham TN, Nakagawa T, Tamura A, Engelman DM. Multi-step denaturation of *Borrelia burgdorferi* OspA, a protein containing a single-layer β -sheet. *Biochemistry* 1999;38:4757–4767. [PubMed: 10200164]
12. Yan S, Kennedy S, Koide S. Thermodynamic and Kinetic Exploration of the Energy Landscape of *Borrelia burgdorferi* OspA by Native-state Hydrogen Exchange. *J Mol Biol* 2002;323:363. [PubMed: 12381326]
13. Koide S, Huang X, Link K, Koide A, Bu Z, Engelman DM. Design of single-layer beta-sheets without a hydrophobic core. *Nature* 2000;403:456–460. [PubMed: 10667801]
14. Makabe K, McElheny D, Tereshko V, Hilyard A, Gawlak G, Yan S, Koide A, Koide S. Atomic structures of peptide self-assembly mimics. *Proc Natl Acad Sci U S A* 2006;103:17753–17758. [PubMed: 17093048]

15. Ohnishi S, Koide A, Koide S. Solution Conformation and Amyloid-like Fibril Formation of a Polar Peptide Derived from a β -Hairpin in the OspA Single-Layer β -Sheet. *J Mol Biol* 2000;301:477–489. [PubMed: 10926522]
16. Yan S, Gawlak G, Smith J, Silver L, Koide A, Koide S. Conformational heterogeneity of an equilibrium folding intermediate quantified and mapped by scanning mutagenesis. *J Mol Biol* 2004;338:811–25. [PubMed: 15099747]
17. Makabe K, Tereshko V, Gawlak G, Yan S, Koide S. Atomic-resolution crystal structure of *Borrelia burgdorferi* outer surface protein A via surface engineering. *Protein Sci* 2006;15:1907–1914. [PubMed: 16823038]
18. Smith CK, Regan L. Guidelines for protein design: The energetics of β -sheet side chain interactions. *Science* 1995;270:980–982. [PubMed: 7481801]
19. Kim CA, Berg JM. Thermodynamic beta-sheet propensities measured using a zinc-finger host peptide. *Nature* 1993;362:267–270. [PubMed: 8459852]
20. Smith CK, Withka JM, Regan L. A thermodynamic scale for the β -sheet forming tendencies of the amino acids. *Biochemistry* 1994;33:5510–5517. [PubMed: 8180173]
21. Minor DLJ, Kim PS. Measurement of the β -sheet-forming propensities of amino acids. *Nature* 1994;367:660–663. [PubMed: 8107853]
22. Minor DLJ, Kim PS. Context is a major determinant of β -sheet propensity. *Nature* 1994;371:264–267. [PubMed: 8078589]
23. Chou PY, Fasman GD. Prediction of protein conformation. *Biochemistry* 1974;13:222–45. [PubMed: 4358940]
24. Chou PY, Fasman GD. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* 1978;47:45–148. [PubMed: 364941]
25. Wouters MA, Curmi PM. An analysis of side chain interactions and pair correlations within antiparallel beta-sheets: the differences between backbone hydrogen-bonded and non-hydrogen-bonded residue pairs. *Proteins* 1995;22:119–131. [PubMed: 7567960]
26. Honig B, Nicholls A. Classical electrostatics in biology and chemistry. *Science* 1995;268:1144–1149. [PubMed: 7761829]
27. Hendsch ZS, Tidor B. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 1994;3:211–226. [PubMed: 8003958]
28. Sindelar CV, Hendsch ZS, Tidor B. Effects of salt bridges on protein structure and design. *Protein Sci* 1998;7:1898–914. [PubMed: 9761471]
29. Hutchinson EG, Thornton JM. A revised set of potentials for beta-turn formation in proteins. *Protein Sci* 1994;3:2207–16. [PubMed: 7756980]
30. Chan AW, Hutchinson EG, Harris D, Thornton JM. Identification, classification, and analysis of beta-bulges in proteins. *Protein Sci* 1993;2:1574–90. [PubMed: 8251933]
31. Kumar S, Nussinov R. Close-range electrostatic interactions in proteins. *Chembiochem* 2002;3:604–17. [PubMed: 12324994]
32. Fogolari F, Brigo A, Molinari H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *J Mol Recognit* 2002;15:377–92. [PubMed: 12501158]
33. Schutz CN, Warshel A. What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins* 2001;44:400–17. [PubMed: 11484218]
34. Simonson T. Macromolecular electrostatics: continuum models and their growing pains. *Curr Opin Struct Biol* 2001;11:243–52. [PubMed: 11297935]
35. Hendsch ZS, Tidor B. Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Sci* 1999;8:1381–92. [PubMed: 10422826]
36. Otzen DE, Fersht AR. Side-chain determinants of β -sheet stability. *Biochemistry* 1995;34:5718–5724. [PubMed: 7727432]
37. Munoz V, Serrano L. Intrinsic secondary structure propensities of the amino acids, using statistical ϕ - ψ matrices: comparison with experimental scales. *Proteins* 1994;20:301–311. [PubMed: 7731949]

38. Zaremba SM, Gregoret LM. Context-dependence of amino acid residue pairing in antiparallel beta-sheets. *J Mol Biol* 1999;291:463–79. [PubMed: 10438632]
39. Mandel-Gutfreund Y, Zaremba SM, Gregoret LM. Contributions of residue pairing to beta-sheet formation: conservation and covariation of amino acid residue pairs on antiparallel beta-strands. *J Mol Biol* 2001;305:1145–59. [PubMed: 11162120]
40. Cootes AP, Curmi PM, Cunningham R, Donnelly C, Torda AE. The dependence of amino acid pair correlations on structural environment. *Proteins* 1998;32:175–89. [PubMed: 9714157]
41. Syud FA, Stanger HE, Gellman SH. Interstrand side chain–side chain interactions in a designed beta-hairpin: significance of both lateral and diagonal pairings. *J Am Chem Soc* 2001;123:8667–77. [PubMed: 11535071]
42. Distefano MD, Zhong A, Cochran AG. Quantifying beta-sheet stability by phage display. *J Mol Biol* 2002;322:179–88. [PubMed: 12215423]
43. Kauzmann W. Some factors in the interpretation of protein denaturation. *Adv Prot Chem* 1959;14:1–63.
44. Dill KA. Dominant forces in protein folding. *Biochemistry* 1990;29:7133–7155. [PubMed: 2207096]
45. DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendruscolo M. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J Mol Biol* 2004;341:1317–26. [PubMed: 15302561]
46. Esteras-Chopo A, Serrano L, Lopez de la Paz M. The amyloid stretch hypothesis: recruiting proteins toward the dark side. *Proc Natl Acad Sci U S A* 2005;102:16672–7. [PubMed: 16263932]
47. Ohnishi S, Koide A, Koide S. The roles of turn formation and cross-strand interactions in fibrillization of peptides derived from the OspA single-layer beta-sheet. *Protein Sci* 2001;10:2083–92. [PubMed: 11567099]
48. Mian IS, Bradwell AR, Olson AJ. Structure, function and properties of antibody binding sites. *J Mol Biol* 1991;217:133–151. [PubMed: 1988675]
49. Doye JP, Louis AA, Vendruscolo M. Inhibition of protein crystallization by evolutionary negative design. *Phys Biol* 2004;1:P9–13. [PubMed: 16204814]
50. Rousseau F, Schymkowitz J, Serrano L. Protein aggregation and amyloidosis: confusion of the kinds? *Curr Opin Struct Biol* 2006;16:118–26. [PubMed: 16434184]
51. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 2005;435:773–8. [PubMed: 15944695]
52. Haspel N, Zanuy D, Ma B, Wolfson H, Nussinov R. A comparative study of amyloid fibril formation by residues 15–19 of the human calcitonin hormone: a single beta-sheet model with a small hydrophobic core. *J Mol Biol* 2005;345:1213–27. [PubMed: 15644216]
53. Zanuy D, Gunasekaran K, Ma B, Tsai HH, Tsai CJ, Nussinov R. Insights into amyloid structural formation and assembly through computational approaches. *Amyloid* 2004;11:143–61. [PubMed: 15523917]
54. Eakin CM, Berman AJ, Miranker AD. A native to amyloidogenic transition regulated by a backbone trigger. *Nat Struct Mol Biol* 2006;13:202–8. [PubMed: 16491088]
55. Dantas G, Kuhlman B, Callender D, Wong M, Baker D. A large scale test of computational protein design: folding and stability of nine completely redesigned globular proteins. *J Mol Biol* 2003;332:449–60. [PubMed: 12948494]
56. Kraemer-Pecore CM, Lecomte JT, Desjarlais JR. A de novo redesign of the WW domain. *Protein Sci* 2003;12:2194–205. [PubMed: 14500877]
57. Ottesen JJ, Imperiali B. Design of a discretely folded mini-protein motif with predominantly beta-structure. *Nat Struct Biol* 2001;8:535–9. [PubMed: 11373623]
58. Wang W, Hecht MH. Rationally designed mutations convert de novo amyloid-like fibrils into monomeric beta-sheet proteins. *Proc Natl Acad Sci U S A* 2002;99:2760–5. [PubMed: 11880628]
59. Myers JK, Pace CN, Scholtz JM. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 1995;4:2138–2148. [PubMed: 8535251]
60. Hamill SJ, Steward A, Clarke J. The folding of an immunoglobulin-like Greek key protein is defined by a common-core nucleus and regions constrained by topology. *J Mol Biol* 2000;297:165–78. [PubMed: 10704314]

61. Bertagna AM, Barrick D. Nonspecific hydrophobic interactions stabilize an equilibrium intermediate of apomyoglobin at a key position within the AGH region. *Proc Natl Acad Sci U S A* 2004;101:12514–9. [PubMed: 15314218]
62. Lovell SC, Word JM, Richardson JS, Richardson DC. The penultimate rotamer library. *Proteins* 2000;40:389–408. [PubMed: 10861930]
63. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 1998;54:905–921. [PubMed: 9757107]
64. Lee BK, Richards FM. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 1971;55:379–400. [PubMed: 5551392]
65. Lesser GJ, Rose GD. Hydrophobicity of amino acid subgroups in proteins. *Proteins* 1990;8:6–13. [PubMed: 2217164]
66. Rocchia W, Alexov E, Honig B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *Journal of Physical Chemistry B* 2001;105:6507–6514.
67. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997;276:307–26.
68. Collaborative_Computational_Project. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst* 1994;D50:760–763.
69. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997;53:240–55. [PubMed: 15299926]
70. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60:2126–32. [PubMed: 15572765]
71. Hutchinson EG, Thornton JM. PROMOTIF - a program to identify and analyze structural motifs in proteins. *Protein Sci* 1996;5:212–220. [PubMed: 8745398]

Figure 1.

(A) Ribbon representation of the OspA structure. The N- and C-termini and the SLB regions are indicated. (B) The sequence of OspA strands 7–11 represented according to the secondary structure topology. The rectangles indicate β -strand residues as defined with the program Promotif.⁷¹ The side chains of the residues enclosed in the dashed lines point toward the reader, and those of the other β -strand residues point away from the reader. Positively and negatively charged residues are shown in blue and red, respectively. Amino acids that are considered as a "good" β -sheet forming amino acid are shaded in gray. (C and D) The atomic structure of the same region viewed from the opposite faces of the β -sheet. The backbone trace is shown as a cartoon model. The side chains of the β -strand residues on the "front" face are shown as spheres in (C) and those on the "back" face in (D). Residues on a β -sheet edge are labeled. The molecular graphics was made with PyMOL (www.pymol.org).

Figure 2.

$\Delta\Delta G$ values for residues in the strand of the OspA single-layer β -sheet (locations of strands were defined according to Promotif⁷¹). The data are presented according to the strand topology shown in Fig 1A. Hydrophobic residues are in green, polar residues in grey, negatively charged residues in red and positively charged residues in blue.

Figure 3.

Correlations of different factors with $\Delta\Delta G$ values. (A) β -sheet propensity. β -sheet propensity was offset so that the propensity for alanine was zero. (B) Statistical pairwise correlation. (C) Hydrophobic surface burial. (D) Polar surface burial. (E) Statistical propensity for turn and bulge. In (C), the surface burial value calculated from the x-ray crystal structure of F126A is also shown as an open circle.

Figure 4.

Crystal structure of F126A-sm1. (A) Superposition of F126A-sm1 (blue) onto OspA-sm1 (red). The mutated residue (F126) is shown as spheres. (B) The root-mean-squared deviations (RMSD) for the C α atoms are plotted versus residue number. The positions of β -strands and the C-terminal helix are shown as bars. The mutation site is indicated as a red circle. (C) Rearrangements of residues surrounding the F126 site in F126A-sm1. OspA-sm1 is shown in red and F126A-sm1 in blue. The backbone atoms are shown as stick models. The side chains are displayed only for residues in the hydrophobic ladder (F102, L109, F126 (A126 in the mutant) and V132).

Table 1

Thermodynamic parameters for the unfolding transitions of newly characterized OspA mutants

Mutation	ΔG_{UI}^{3M}	ΔG_{NU}^{3M}	m_{NI}	$\Delta\Delta G_{NU}$
wild type	1.74 (0.04)	2.71 (0.03)	-2.41 (0.02)	—
K117A	1.91 (0.18)	3.05 (0.13)	-2.41 (0.02)	-0.34 (0.13)
D118A	0.56 (0.08)	0.72 (0.18)	-2.76 (0.02)	1.99 (0.18)
E131A	0.93 (0.03)	1.80 (0.06)	-2.43 (0.03)	0.91 (0.07)
S133A	1.76 (0.12)	2.75 (0.24)	-2.45 (0.02)	-0.04 (0.24)
E134A	1.66 (0.08)	2.38 (0.12)	-2.47 (0.01)	0.33 (0.12)
K135A	1.18 (0.08)	1.63 (0.07)	-2.69 (0.04)	1.08 (0.08)
I136A	0.53 (0.06)	0.46 (0.20)	-2.76 (0.02)	2.25 (0.20)
T138A	0.63 (0.11)	1.25 (0.29)	-2.61 (0.03)	1.46 (0.29)
S111E	1.04 (0.04)	2.10 (0.07)	-2.36 (0.03)	0.61 (0.07)
K112M	0.98 (0.02)	1.99 (0.05)	-2.48 (0.02)	0.72 (0.06)
S121T	1.76 (0.05)	2.73 (0.05)	-2.50 (0.01)	-0.02 (0.05)
T122V	2.69 (0.03)	4.07 (0.01)	-2.34 (0.01)	-1.36 (0.03)
T122Y	2.19 (0.11)	3.41 (0.15)	-2.44 (0.03)	-0.70 (0.15)
E123Q	1.02 (0.04)	2.06 (0.01)	-2.55 (0.03)	0.65 (0.03)
E124L	2.29 (0.05)	3.25 (0.03)	-2.41 (0.01)	-0.54 (0.04)
E124Q	1.02 (0.06)	1.53 (0.09)	-2.55 (0.02)	1.18 (0.09)
I136K	0.82 (0.14)	1.20 (0.14)	-2.58 (0.02)	1.51 (0.15)
<i>Proteins containing the E160L mutation</i>				
E160L	1.82 (0.02)	-0.33 (0.03)	-2.57 (0.02)	—
F102A [*]	-1.31 (0.02)	-3.26 (0.04)	-2.43 (0.03)	2.93 (0.05)
L109A [*]	0.17 (0.01)	-1.82 (0.11)	-2.43 (0.02)	1.49 (0.12)
V110A [*]	0.61 (0.03)	-1.32 (0.11)	-2.43 (0.01)	0.99 (0.12)
S111A [*]	1.31 (0.02)	-0.69 (0.05)	-2.48 (0.01)	0.36 (0.06)
K113A [*]	0.18 (0.08)	-2.18 (0.26)	-2.58 (0.02)	1.85 (0.26)
T115A [*]	1.81 (0.02)	-0.23 (0.13)	-2.59 (0.01)	-0.11 (0.13)
S116A [*]	1.64 (0.06)	-0.47 (0.25)	-2.55 (0.02)	0.14 (0.26)

ΔG_{UI}^{3M} is the free energy difference between the unfolded state and the intermediate state at 3 M urea. ΔG_{NU}^{3M} is the free energy difference between the unfolded state and the native state at 3 M urea. m_{NI} is the m-value for the transition between the native and intermediate states. The wild type data is taken from Yan et al.¹⁶ $\Delta\Delta G_{NU}$ is the difference in ΔG_{NU}^{3M} between the wild type and a mutant ($\Delta\Delta G_{NU} = \Delta G_{NU}^{3M}(\text{wild type}) - \Delta G_{NU}^{3M}(\text{mutant})$). A positive ΔG_{NU} value indicates destabilization by a mutation.

^{*} These mutants also contain the E160L mutation. The $\Delta\Delta G_{NU}$ values for these mutants were calculated using the ΔG_{NU}^{3M} value for E160L as the reference. Errors shown are the greater of the standard deviations from curve fitting and the standard deviations derived from multiple independent experiments.

Table 2

OspA stability changes and stability parameters used in this study^a

Mutation	Structure ^b	$\Delta\Delta G_{exp}$ (kcal/mol)	Propensity (kcal/mol)	$\Delta\Delta G_{pair}$ (kcal/mol)	$\Delta SASA_{np}$ (Å ²)	$\Delta SASA_{pol}$ (Å ²)	ΔE_{coul} (kcal/mol)	ΔE_{solv} (kcal/mol)	$\Delta\Delta E_{Delpol}$ (kcal/mol)
<i>Ala-scan mutants</i>									
E100A	S	1.78 (±0.21)	0.15	1.80	63.4	92.2	21.50	-24.26	-2.76
F102A	S	2.93 (±0.05)	0.58	-0.10	171.9	29	0.16	-2.14	-1.98
L109A	S	1.49 (±0.12)	0.21	0.27	123.8	27.6	0.00	-0.80	-0.80
V110A	B	0.99 (±0.12)	0.4	-	36.7	7.9	0.00	-0.01	-0.01
S111A	S	0.36 (±0.06)	0.56	-0.34	20.8	65.8	-0.37	-0.09	-0.46
K112A	S	1.83 (±0.20)	0.07	2.12	114.1	58.1	17.76	-18.00	-0.24
K113A	S	1.85 (±0.26)	0.07	0.70	125.3	61.4	15.60	-16.77	-1.17
V114A	S	1.69 (±0.13)	0.53	0.02	79.8	29.2	0.00	-0.18	-0.18
T115A	S	-0.11 (±0.13)	0.86	0.52	64.1	55.2	0.18	-1.44	-1.26
S116A	S	0.14 (±0.25)	0.56	1.24	21.3	66.2	1.62	-1.16	0.46
K117A	T	-0.34 (±0.13)	0.07	-	-7.6	28.2	21.98	-22.17	-0.19
D118A	T	1.99 (±0.18)	0.69	-	-12.5	49.5	23.61	-24.01	-0.40
K119A	T	0.24 (±0.20)	0.1	-	-2	71.9	14.30	-13.81	0.49
S120A	S	0.98 (±0.02)	0.56	0.89	23.8	59.9	-1.15	-0.54	-1.69
S121A	S	0.82 (±0.03)	0.56	0.29	21.9	66.7	-1.21	-0.56	-1.77
T122A	S	1.65 (±0.02)	0.855	-0.88	66.6	41	-0.12	-0.25	-0.36
E123A	S	1.01 (±0.10)	0.15	0.47	54.9	45.8	26.33	-25.76	0.57
E124A	S	1.68 (±0.06)	0.15	1.80	63.8	86.4	29.01	-30.18	-1.17
K125A	S	1.21 (±0.05)	0.07	0.62	56.9	63.4	11.37	-12.24	-0.87
F126A	S	3.31 (±0.07)	0.58	-0.24	176.6	17.4	0.14	-1.70	-1.56
N127A	T	2.54 (±0.03)	0.76	-	-24.9	77.7	7.79	-4.39	3.40
E128A	T	0.50 (±0.10)	0.22	-	9.5	14.3	26.77	-25.55	1.22
K129A	T	0.19 (±0.10)	0.14	-	-3.8	35.7	14.89	-15.36	-0.48
E131A	T	0.91 (±0.06)	0.15	-	12.6	47.1	28.97	-31.98	-3.01
V132A	S	1.56 (±0.05)	0.53	-0.05	75.7	26.4	0.00	-0.44	-0.44
S133A	B	-0.04 (±0.24)	0.16	-	-17.5	32.1	-1.60	0.85	-0.75
E134A	S	0.33 (±0.12)	0.15	1.31	60.2	67.1	30.54	-31.45	-0.91
K135A	S	1.08 (±0.07)	0.07	1.03	110.3	68.5	1.67	-4.90	-3.23
I136A	S	2.25 (±0.20)	0.62	0.55	93.8	25.6	0.00	-0.67	-0.67
I137A	S	2.73 (±0.15)	0.62	-0.33	78.9	26	0.00	-1.08	-1.08
T138A	S	1.46 (±0.29)	0.86	0.93	47.5	52.0	0.34	-1.52	-1.18
<i>Non-Ala mutants</i>									
S111E	S	0.68 (±0.07)	-0.38	-	-48.6	14.8	-	-	-
K112M	S	0.83 (±0.05)	-0.36	-	8.1	3.2	18.5	-17.37	1.13
S121T	S	-0.01 (±0.05)	-0.30	-	-27.3	1.0	-	-	-
T122V	S	-1.36 (±0.03)	0.33	-	-19.6	11.0	-0.16	0.22	0.07
T122Y	S	-0.69 (±0.15)	0.15	-	-58.1	-19.9	-	-	-
E123Q	S	0.65 (±0.04)	-0.03	-	2.4	-7.7	22.78	-21.72	1.07
E124Q	S	1.18 (±0.09)	-0.03	-	-0.9	-7.1	31.24	-29.28	1.96
E124L	S	-0.54 (±0.03)	-0.06	-	-62.2	61.5	28.6	-29.47	-0.86
I136K	S	1.51 (±0.14)	0.55	-	33.6	-47.0	-	-	-

^a All values were derived by subtracting the value for the wild type from that for the mutant. The unit is kcal/mol unless otherwise specified. The $\Delta\Delta G_{exp}$ values are ΔG_{NTU} values taken from Table 1 and from Yan et al.¹⁶

^b Structure classification: S, β -strand; T, turn; and B, β -bulge. β -Prop is averaged β -sheet propensity (see main text). $\Delta\Delta G_{\text{pair}}$ is cross-strand pairwise correlation (see main text). $\Delta SASA_{\text{pp}}$ and $\Delta SASA_{\text{pol}}$ are changes in buried nonpolar and polar surface areas, respectively, defined as $SASA^{\text{mutant}} - SASA^{\text{wild type}}$. ΔE_{coul} and $\Delta\Delta E_{\text{Delphi}}$ are changes in Coulombic, solvation and net electrostatic energies, respectively, calculated using the program Delphi.

Table 3
Data collection and refinement statistics for F126A-sm1

<i>Data collection statistics</i>	
Space group	P2 ₁
Cell parameters	a=34.78
	b=53.84
	c=65.50
	β =99.32
Beamline	APS-19BM
Wavelength	0.9791 Å
Resolution (Å) (highest resolution shell) ^a	50–1.2(1.24–1.20)
Completeness(%)	95.2(91.6)
I/σ (I)	15.59(2.43)
R _{merge} ^b	0.071(0.265)
Average redundancy	2.3(1.8)
<i>Refinement statistics</i>	
Resolution range (Å)	20.0–1.2
Reflections used (free)	68343(3599)
R factor ^c	0.148
R _{free} ^d	0.176
RMS deviations	
Bonds (Å)	0.011
Angles (°)	1.397
No. protein residues	246
No. waters	595
Average B factor (Å ²)	10.58
<i>Ramachandran plot statistics</i>	
Most favored (%)	92.8
Additionally allowed (%)	7.2
Generally allowed (%)	0.0

^a Highest resolution shell is shown in parenthesis.

^b R-merge = $\sum_{hkl} \sum_i |I(hkl)_i - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i \langle I(hkl) \rangle$ over *i* observations of a reflection hkl.

^c R-factor = $\sum |F(obs) - F(calc)| / \sum |F(obs)|$.

^d R_{free} is R with 5% of reflections sequestered before refinement.