THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Two Structural Scenarios for Protein Stabilization by PEG

Shu-Han Chao,[†] Sam S. Matthews,[‡] Ryan Paxman,[‡] Aleksei Aksimentiev,[†,∥] Martin Gruebele,*[,†,§,∥] and Joshua L. Price*[,‡]

[†]Department of Physics and Center for the Physics of Living Cells, University of Illinois, Urbana, Illinois 61801, United States

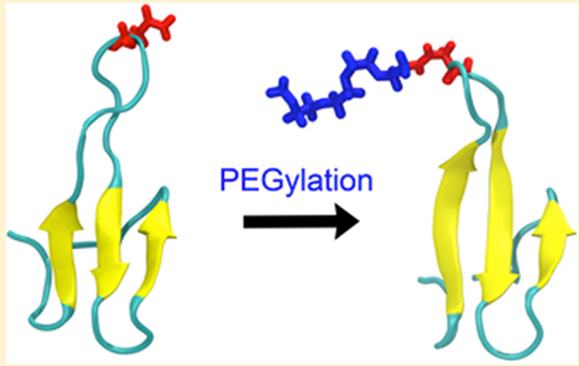
[‡]Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, United States

[§]Department of Chemistry, University of Illinois, Urbana, Illinois 61801, United States

[∥]Center for Biophysics and Quantitative Biology, University of Illinois, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: PEGylation, or addition of poly(ethylene glycol) chains to proteins, is widely used to improve delivery in pharmaceutical applications. Recent studies suggest that stabilization of a protein by PEG, and hence its proteolytic degradability, is sequence-dependent and requires only short PEG chains. Here we connect stabilization by short PEG chains directly to the structural dynamics of the protein and PEG chain. We measured the stability of human Pin1 WW domain with PEG-4 at asparagine 19 for a full mutant cycle at two positions thought to influence PEG−protein interaction: Ser16Ala and Tyr23Phe. We then performed explicit solvent molecular dynamics simulations on all PEGylated and PEG-free mutants. The mutant cycle yields a nonadditive stabilization effect where the pseudo-wild type and double mutant are more stabilized relative to unPEGylated proteins than are the two single mutants. The simulation reveals why: the double mutant suffers loss of β-sheet structure, which PEGylation restores even though the PEG extends as a coil into the solvent. In contrast, in one of the single mutants, PEG preferentially interacts with the protein surface while disrupting the interactions of its asparagine host with a nearby methionine side chain. Thus, PEG attachment can stabilize a protein differentially depending on the local sequence, and either by interacting with the surface or by extending into the solvent. A simulation with PEG-45 attached to asparagine 19 shows that PEG even can do both in the same context.



INTRODUCTION

N-Glycosylation is an important post-translational modification of proteins. It plays a role in signaling and recognition, such as in the innate immune system.[1] A human-made analogue, PEGylation, has been shown to stabilize proteins and lengthen their degradation time when used in therapeutics.[2,3] PEG may impart these useful properties by improving protein solubility, by reducing access of proteolytic enzymes, and by intrinsically stabilizing its host protein. Conventional commercial PEGylation often attaches a linear or branched chain of more than 50 monomer units to lysine residues at the host protein surface in a nonspecific fashion; more recent approaches have targeted the N-terminal α-amino group or a surface cysteine residue for specific modification.

Host protein stabilization is one important reason for PEGylation. How important are PEG length and location for maximizing stabilization? Stabilization can result from crowding, from more specific local interactions between PEG and protein, or from PEG-induced alteration of local protein structure. Crowding requires long PEG chains because a crowding polymer must be about the size of the protein it crowds, but location may be less important. Local contributions to stability could arise for short PEG chains, and may be much more position-sensitive.

Several studies suggest that long PEG chains can stabilize proteins by interacting with the protein surface and reducing the solvent accessible surface area, or by introducing molecular crowding (excluded volume effect) to slow down the unfolding rate of the host protein.[4−7] Other work has shown that N-PEGylation and N-glycosylation (attachment at asparagine) alter protein stability in a sequence-sensitive manner.[8−13] A recent study showed that protein degradation is reduced when a short PEG chain stabilizes a protein.[6] Moreover, chains as short as four PEG units can stabilize a host protein at least as much as a chain of 45 PEG units at the same position,[6] ruling out crowding or reduced surface accessibility as the sole causes of PEG-based protein stabilization.

Context-specific local stabilization could arise from two very different scenarios, or from their combination. In one scenario,

Special Issue: James L. Skinner Festschrift

Received: March 4, 2014
Revised: May 10, 2014
Published: May 12, 2014

   dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388−8395

the PEG chain interacts extensively with the protein surface, for example, by long-lived hydrogen bonds; because of the reduced configurational entropy of surface-bound PEG, the interactions would have to favor enthalpically the folded protein over the unfolded protein. At the other extreme, PEG attachment stabilizes protein structure due to steric effects at the attachment site, but the PEG itself extends into the solvent as a coil; that way the composite PEG–protein system does not pay the entropy penalty of collapsing the PEG onto the protein surface. It is of course possible for a PEG chain to sample both scenarios.

Here we study the local stabilization scenario using short PEG chains of four monomers, which are too short to act by crowding or by large-scale solvent exclusion from the surface. Finally, we contrast the short PEG chain with a longer PEG chain (N = 45 monomers), close in size to our small test protein (N = 34 residues), and closer to commercial PEG chain length.

In our experiments and molecular dynamics simulations, a small PEG tetramer is attached to loop 1 that connects two β-strands of human Pin1 WW domain. WW is a triple stranded β-sheet protein domain with 34 residues (hPin1 WW, see Figure 1). The use of a PEG tetramer is based on our previous study

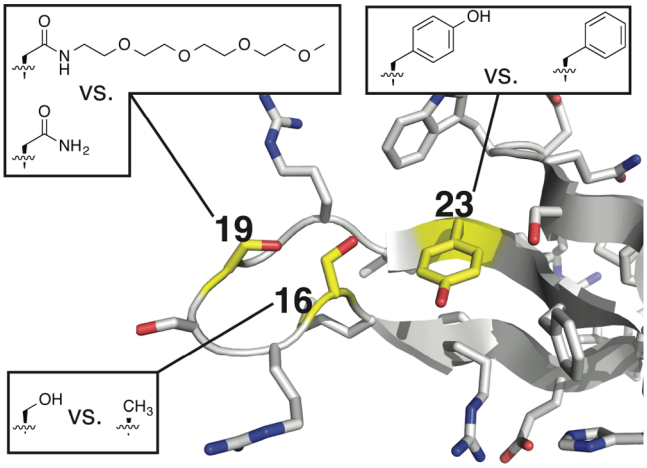


Figure 1. Mutation sites of the WW domain at its N-terminal reverse turn. WW is shown as a ribbon diagram, with side chains shown as sticks (PDB: 1PIN[15]). The mutation sites Ser16, Ser19, Tyr23 are highlighted in yellow, with side chain differences between variants shown in boxes. Asn19 is the PEGylation site.

showing that a tetramer stabilizes this model protein more than shorter (1–3 units) or longer (8–45 units) PEG oligomers.[6] We first examine a Ser19Asn pseudo-wild-type variant with and without PEG attached to the asparagine at position 19. We then study a cycle of two mutations that truncate two hydrogen-bond-forming side chains at the protein surface near the base of the attached PEG: Ser16Ala, Tyr23Phe. Finally, we look at the corresponding double mutant with both hydroxyl groups missing.

Our experiments show that the single Ser16Ala or Tyr23Phe mutant is less stabilized by PEGylation than the pseudo-wild-type variant. This result by itself is consistent with hydrogen bonds between Ser16 or Tyr23 and PEG, which would be disrupted by the mutations. However, the double Ser16Ala Tyr23Phe mutant (which lacks hydroxyl groups at positions 16 and 23) is stabilized just as much by PEGylation as the pseudo-wild-type variant; this result challenges the importance of direct PEG–protein hydrogen bonding. We interpret this unexpected nonadditive result structurally by all-atom explicit solvent molecular dynamics simulations, an excellent tool for protein–PEG interactions.[14] The simulations correlate well with the thermodynamic cycle: the Ser16Ala Tyr23Phe double mutation causes a loss of β-sheet structure in the non-PEGylated protein, but the β-sheet structure is restored when the double mutant protein is PEGylated, even though PEG extends into the solvent. In contrast, PEG creates new local interactions (in the Ser16Ala single mutant), or preferentially interacts with the protein surface (in the Tyr23Phe single mutant).

Is longer PEG necessarily better? No—a recent study showed that PEG-45 linked to Asn19 improves protein stability slightly less than PEG-4.[6] We investigated this structurally with a final simulation of PEG-45 at position 19. We find that long PEG equilibrates evenly between solvent-exposed coil and protein surface-coating coil. All of these observations highlight that a subtle combination of both enthalpic and entropic factors influences whether PEG folds onto a protein surface or not, and that in some cases the protein–PEG construct's overall free energy is lowered more by exclusion of PEG from the protein surface than by protein–PEG interactions.

METHODS

Protein Synthesis, Purification, and Stability. WW variants were prepared via microwave-assisted solid-phase peptide synthesis, as described previously.[6,10] Proteins were purified by preparative reverse-phase high performance liquid chromatography (HPLC) on a C18 column using a linear gradient of water in acetonitrile with 0.1% v/v trifluoroacetic acid. HPLC fractions containing the desired protein product were pooled, frozen, and lyophilized. Protein identity was confirmed by electrospray ionization time-of-flight mass spectrometry, and purity was assessed by analytical HPLC (see Figures S1–S12 in the Supporting Information).[6]

The conformational stability of WW variants was assessed by thermal denaturation and monitored by circular dichroism spectropolarimetry (CD). The melting temperature for each protein was obtained by fitting the sigmoidal unfolding curves to a two-state model with an adjustable enthalpy and heat capacity [see the Supporting Information for raw CD spectra (Figures S13–S15) and for model and fitting parameters].

Molecular Dynamics Simulations. The structure of the wild-type WW domain of human protein Pin1 was obtained from the protein data bank (PDB: 1PIN, resolution of 1.35 Å[15]). All the mutations were made using Visual Molecular Dynamics (VMD) software.[16] All-atom molecular dynamics simulations were performed using the program NAMD2[17] with periodic boundary conditions, the CHARMM36 parameter set for the TIP3P water model,[18] ions,[19] protein,[20–22] and PEG,[23,24] and ion-pair specific corrections to the Lennard-Jones parameter σ.[25] The force field of the PEGylated side chain of asparagine at position 19 was assigned by analogy to similar-structured N-methylacetamide in the CHARMM36 force field. Sodium chloride ions were added to neutralize the whole system and mimic the nonzero ionic strength of the experimental environment. The simulation box measured 70 Å in each dimension and contained around 32 000 atoms, including protein, water molecules, and ions. All simulations employed a 1 fs time step and a 7–8 Å shift-cutoff scheme for van der Waals and short-range electrostatic forces. Long-range electrostatic interactions were computed using the particle mesh Ewald method[26] over an ~1.0 Å resolution grid.

dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388–8395

Table 1. Nomenclature of Non-PEGylated WW Variants, Their Melting Temperature, and Folding Free Energies at Room Temperature[a]

WW variant	mutation	T_m (°C)	ΔG_f (kcal/mol)
S16/N19/Y23	W34F; S19N	55.6 ± 0.2	0.44 ± 0.02
A16/N19/Y23	W34F; S19N; S16A	50.9 ± 0.2	0.85 ± 0.03
S16/N19/F23	W34F; S19N; Y23F	51.4 ± 0.4	0.75 ± 0.04
A16/N19/F23	W34F; S19N; S16A; Y23F	45.7 ± 1.0	1.22 ± 0.13

[a]Tabulated data are given as mean ± standard error at 333 K. Data were obtained from variable temperature CD experiments on 100 μM solutions of protein in 20 mM phosphate, pH 7. Data for **S16/N19/Y23** are from refs 8−13.

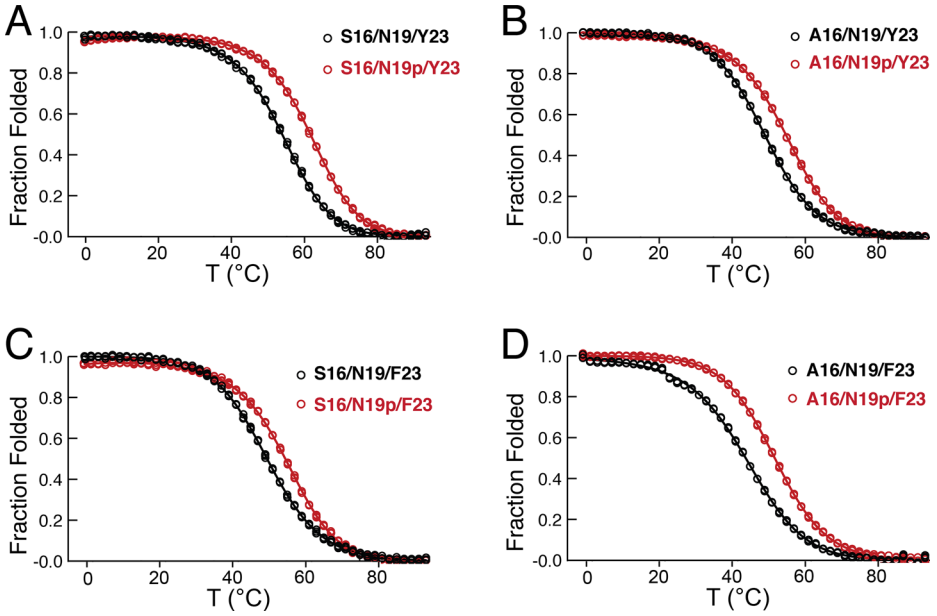


Figure 2. Stability characterization of each WW mutant. CD melting curves for mutant (A) **S16/N19/Y23**, (B) **A16/N19/Y23**, (C) **S16/N19/F23**, and (D) **A16/N19/F23** with (red curve) and without (black curve) PEG attachment are shown.

Following a 2400-step minimization using a conjugate gradient method,[27] each system was equilibrated for 1 ns at a constant pressure of 1 atm using a Langevin piston pressure control method[28] with temperature fixed at 293 K. Production simulations were then performed at constant equilibrium volume in the NVE ensemble for 100 ns. To allow for further equilibration, only the last 60 ns in each trajectory were taken for further analysis. Results were visualized with VMD software (see Figures S16−S23 in the Supporting Information for additional simulation details).

Structural Analysis of the Simulations. Structural information, like secondary structure classification and dihedral angles, was obtained using the *Timeline* plugin in VMD. The protein−PEG distance is the nearest distance between any protein atom and any of the PEG oxygen atoms. The hydrogen bond formation was calculated with a criterion of the donor−acceptor cutoff distance 3.5 Å and an angle less than 150°. Both distance and H-bond calculations were done using tcl scripts. The root-mean-square distance (RMSD) value of asparagine at position 19 relative to the initial conformation of pseudo-wild type (Figure S24, Supporting Information) based on crystal structure (PDB: 1PIN) was taken as an indicator of PEG orientation. The protein molecule in each simulation was first aligned with the crystal structure based on its most rigid β-sheet structure (residue 11 to 15, 22 to 26, and 32 to 33). The RMSD of Asn19 was then calculated for all atoms except hydrogen throughout the whole trajectory. The alignment and calculation were both done using the *RMSD Trajectory Tool* of VMD.

RESULTS

The Impact of Mutations on PEG-Aided Stabilization. We chose an hPin1 WW domain variant with mutations Trp34Phe and Ser19Asn (which we call pseudo-wild type or **S16/N19/Y23**) as our model system (Figure S24, Supporting Information), because glycosylation and PEGylation with PEG-4 at that site have already been investigated extensively in this variant.[6,9,10,12] The Ser19Asn mutation allows N-PEGylation at a natural asparagine site, in analogy to N-glycosylation. As in previous experiments,[6,9,10,12] the C-terminal tryptophan at position 34 was mutated to phenylalanine so that the fluorescence signal would only be determined by the N-terminal tryptophan at position 11. These two mutations together slightly decrease stability, causing the melting temperature to decrease from 58.6 to 55.6 °C.[8−13,29]

Structural data for the unmodified Pin WW domain shows that position 19 is close to a number of polar side chains, in particular Ser16 and Tyr23 (Figure 1). We wondered if the PEG oligomer oxygen atoms are engaging in stabilizing H-bond interactions with the side chain hydroxyl groups of these residues. To test this hypothesis, we prepared a variant of **S16/N19/Y23** (Table 1), in which Ser16 was truncated to Ala (**A16/N19/Y23**), along with a variant in which Tyr23 was truncated to Phe (**S16/N19/F23**); these variants lack the OH group either at position 16 or at position 23.

As shown in Table 1, on the basis of a two-state fit (Methods and the Supporting Information) of variable temperature CD

dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388−8395

Table 2. Stabilizing Effect of PEG[a]

WW variant	ΔT_m (°C)	$\Delta\Delta G_f$ (kcal/mol)	$\Delta\Delta H_f$ (kcal/mol)	$-T\Delta\Delta S_f$ (kcal/mol)
S16/N19p/Y23	7.7 ± 0.4	−0.76 ± 0.04	3.2 ± 1.4	−4.0 ± 1.4
A16/N19p/Y23	5.8 ± 0.3	−0.54 ± 0.03	−1.2 ± 0.8	0.6 ± 0.8
S16/N19p/F23	5.0 ± 0.4	−0.42 ± 0.04	−3.3 ± 1.0	2.9 ± 1.0
A16/N19p/F23	8.0 ± 1.1	−0.63 ± 0.13	−0.9 ± 2.8	0.2 ± 2.9

[a]The "**p**" indicates PEGylation. Tabulated data are given as mean ± standard error relative to the corresponding non-PEGylated compound at 333 K. Data were obtained from variable temperature CD experiments on 100 μM solutions of protein in 20 mM phosphate, pH 7. Data for **S16/N19/Y23** are from refs 8−13.

data (Figure 2), the single mutations reduce the melting temperature T_m of the protein by about 5 °C. The double mutation reduces the melting temperature by about an additional 5 °C, so the two truncations have an approximately additive effect on protein stability.

Variable temperature CD data shown in Figure 2 and summarized in Table 2 indicate that PEGylation at position 19 with PEG-4 increases the stability of each variant. The stabilizing effect is nearly 8 °C on the pseudo-wild-type **S16/N19/Y23**. The stabilizing effect drops to only about 5−6 °C for the two single mutants. These results would certainly suggest that Ser16 and Tyr23 are important mediators of the stabilizing impact of PEGylation, and could be consistent with a direct interaction between PEG and these residues.

If specific PEG−Ser16 and PEG−Tyr23 interactions contribute additively to the increased stability, we would expect that simultaneously changing Ser16 to Ala and Tyr23 to Phe should decrease the stabilizing impact of PEGylation to approximately 2 °C or only −0.20 ± 0.08 kcal/mol. To test this hypothesis, we prepared double mutant **A16/N19/F23** and its PEGylated version **A16/N19p/F23**. We were surprised to observe that PEGylation increases the double mutant's conformational stability by 8 °C or about −0.63 ± 0.13 kcal/mol, the same level of stabilization observed for the pseudo-wild-type variant **S16/N19/Y23**. Though the individual Ser16Ala and Tyr23Phe experiments suggest that Ser16 and Tyr23 side chains are important, the surprising and counter-intuitive PEG-4 stabilization of the double mutant indicates that their importance is not via independent direct H-bonding of their hydroxyl groups to PEG.

To gain additional insights, we used van't Hoff analysis to parse the observed PEGylation-based increases of the conformational stability ($\Delta\Delta G_f$) into enthalpic ($\Delta\Delta H_f$) and entropic components ($-T\Delta\Delta S_f$) (Table 2). The increased stability of PEGylated **S16/N19p/Y23** relative to the pseudo-wild-type variant **S16/N19/Y23** is favored entropically based on experiment. In contrast, the PEG-based stabilization of **S16/N19p/F23** relative to **S16/N19/F23** is favored enthalpically. The other variants are too close to call.

The Structural Impact of Asn19 on Loop 1. Molecular dynamics simulation provides atomic resolution to interpret structurally the stabilization mechanism of PEGylation we measured. In return, consistency with experimental data can validate the simulations.

We started by comparing the local conformation of the N-terminal reverse turn between the pseudo-wild-type **S16/N19/Y23** and the WW domain before addition of the asparagine, **S16/S19/Y23** (PDB: 1PIN). The structural changes at the reverse turn are shown in Figure 3A. The crystal structure has the Ser19 side chain pointing toward the binding side ("front") of the protein. It stays in that orientation throughout the whole simulation. In contrast, the Asn19 side chain points toward the back. This structural difference can also be distinguished by the backbone dihedral angles of these two residues: Ser19 lies in the β-sheet region, while Asn19 sits in the α_{Left}-helix region (Figure 3B).

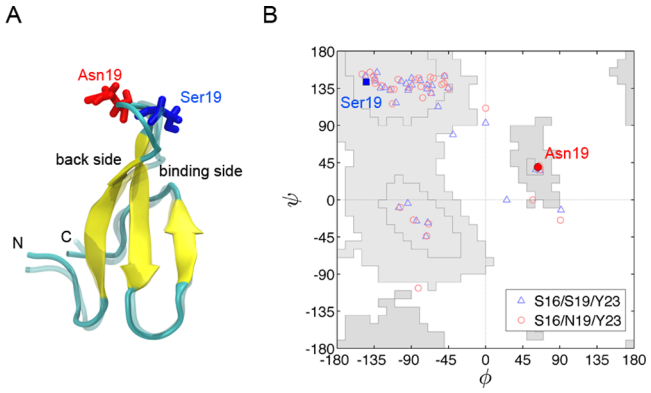


Figure 3. (A) Structural change from the **S16/S19/Y23** hPin1 WW domain (transparent, PDB: 1PIN[15]) to our pseudo-wild-type **S16/N19/Y23**. (B) Ramachandran plot, showing all 34 residues of **S16/S19/Y23** (blue triangles) and of **S16/N19/Y23** (red circles). Each point in the plot is the average dihedral angle of a 60 ns equilibrium simulation. The points for the Ser vs Asn side chains at position 19 are colored according to part A.

The 16/23 Double Mutation Disrupts Secondary Structure. Positions 16 and 23 lie right at the interface of strands 1 and 2 and loop 1. Both mutations reduce stability by altering local loop structure (discussed further below) and local β-sheet structure. **A16/N19/Y23** has longer strands 1 and 2 (Figure S17A, Supporting Information), while **S16/N19/F23** has both strands less stable (Figure S18A, Supporting Information). A much greater structural breakdown happens in double mutant **A16/N19/F23**: the β-sheet formed by strands 1 and 2 becomes unzipped from the loop (Figure S19A, Supporting Information), creating a longer and more flexible loop 1. Moreover, the turn shifts one position from residues 16−20 to residues 17−21. The extensive destabilization of the β-sheet structure observed in the simulation agrees with the experimental observation that the double mutant is 10 °C less stable than the pseudo-wild type.

PEG Does Not Hydrogen Bond with S16 and Y23. To examine our initial hypothesis of PEG interaction with serine 16 and tyrosine 23, the minimum distance between PEG-4 and residues on the protein surface was calculated (Figure 4, hydrogen bonds in red). In the context of the pseudo-wild type, PEG interacts very little with residues 16 and 23, and there is very little hydrogen bonding from PEG to any side chains. In contrast, PEG in both single mutants actually forms hydrogen bonds but with a completely different residue, arginine 21. The double mutant looks more like the pseudo-wild type in terms of PEG extending into solution and very little H-bonding. None

dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388−8395

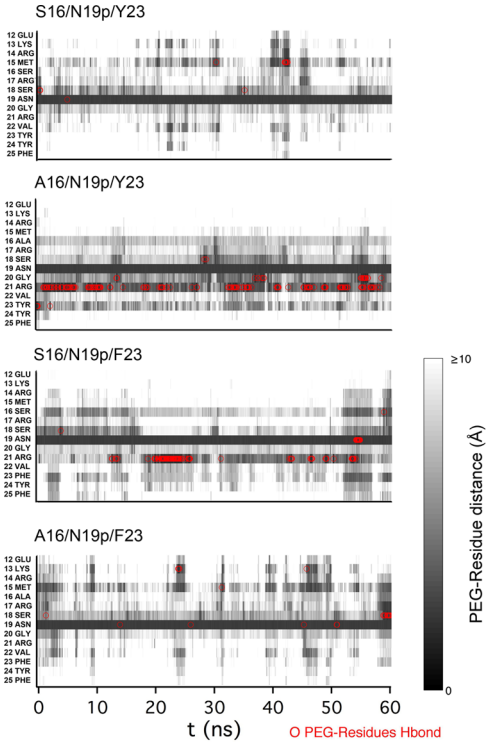


Figure 4. Closest PEG−residue distance for each residue near loop 1 along a 60 ns equilibrium trajectory. The distance is indicated by a gray scale bar. Hydrogen bond formation between PEG and individual residues is indicated by a red circle. The cutoff for H-bonds was 3.5 Å.

of these variants consistently form hydrogen bonds between PEG and either Ser16 or Tyr23.

Effect of PEGylation on Loop Orientation and New Side Chain Contacts. If PEG does not interact directly with the mutated residues, how does it stabilize the protein by different amounts for different mutants? Very specific structural changes occur upon PEGylation. Each non-PEGylated variant already favors different loop conformations, and when the PEG-4 oligomer is linked to the protein, some variants continue to follow that propensity while others reverse it completely.

Figure 5 summarizes the effect of mutations on loop orientation before and after PEG is attached. The histograms plot the RMSD of heavy atoms at position 19 relative to their initial positions in the pseudo-wild-type (**S16/N19/Y23**) structure shown in Figure 3 and Figure S24 (Supporting Information). In the pseudo-wild type (A) and double mutant (D), the loop is flipped to the back of the hPin1 WW domain with or without PEG-4 attached, and the PEG chain predominantly extends into the solvent (jp502234s_si_002.avi, Supporting Information). In these two cases, the loop points in the opposite direction from the Ser19 wild type in Figure 3. In the **A16/N19/Y23** single mutant, the bare loop is also flipped to the back, but the addition of PEG inverts that propensity to a more wild-type-shaped loop (Figure 5B), and the PEG interacts strongly with arginine 21. Finally, in the **S16/N19/F23** single mutant, the loop flips even more strongly to the binding side than in the Ser19 wild type in Figure 3, and when PEG-4 is attached, this trend persists. It is worth pointing out that PEG has little effect on the flexibility of the pseudo-wild-type loop (the widths of the red and blue histograms are identical in Figure 5A), but it reduces loop fluctuations in all three mutants (blue distribution is narrower than red distribution).

In the single mutants, a very specific interaction of PEG-4 with Arg21 is easy to discern. Figure 6A shows this "side chain" scenario for the **A16/N19/Y23** mutant. In the non-PEGylated protein, the loop is flipped into a non-native "back" conformation due to a persistent interaction between the side chains of Asn19 and Met15, which fluctuate in and out of contact with an equilibrium constant K_{eq} close to unity as judged by the trajectory (jp502234s_si_003.avi, Supporting Information). Once PEGylated, Asn19 ceases to interact with Met15 and instead flips to the binding side where PEG-4 interacts with Arg21 (jp502234s_si_004.avi, Supporting Information).

In the double mutant, PEG stabilizes the local secondary structure instead of interacting with protein side chains. Figure

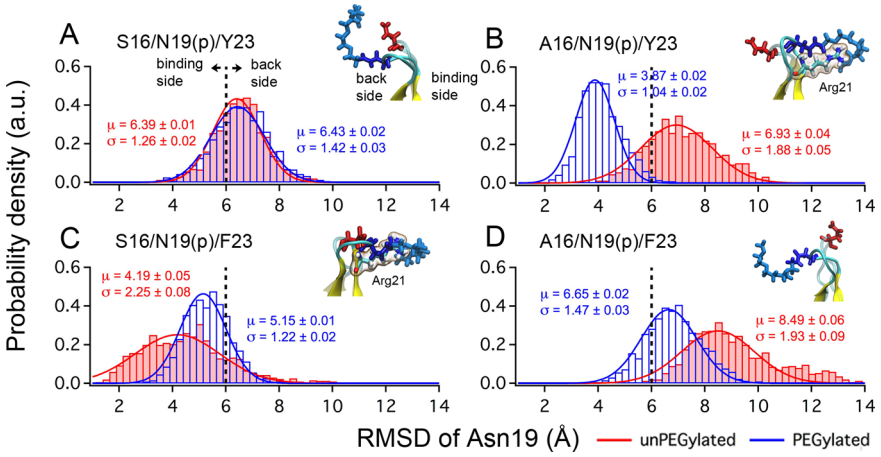


Figure 5. RMSD histogram of asparagine 19 relative to its initial position in the pseudo-wild-type structure before relaxation. Non-PEGylated variants are shown in red, and PEGylated variants, in blue. (A) **S16/N19/Y23**, (B) **A16/N19/Y23**, (C) **S16/N19/F23**, and (D) **A16/N19/F23F**. The histograms are taken from 60 ns trajectories (Figure S20, Supporting Information) and fitted to a Gaussian distribution with mean μ and standard deviation σ shown in each panel. The dashed line indicates the difference of RMSD level between the loop oriented toward the back side (>6 Å) and the binding side (<6 Å). The insets show the representative snapshots of the loop before and after PEGylation, with Asn19 (red when not PEGylated and blue when PEGylated) and PEG (cyan) shown explicitly. Arginine 21 is also shown in parts B and C to highlight its interaction with PEG. See the Supporting Information movies for full trajectories of the mutants in (A) and (B).

dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388−8395

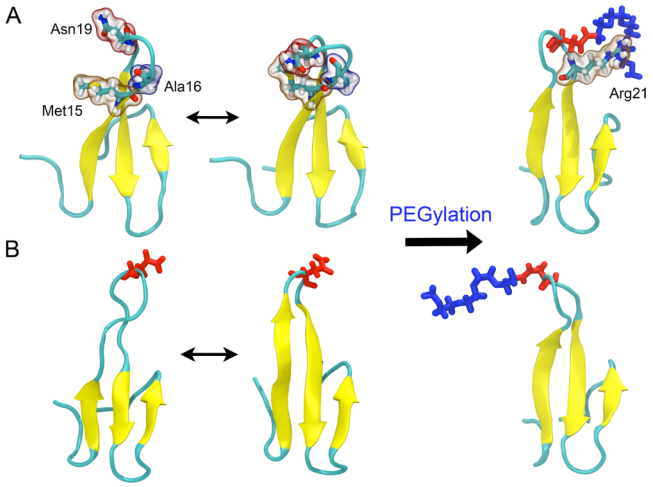


Figure 6. Two PEG stabilization scenarios. (A) PEG disrupts a side chain interaction to acquire its own preferred side chain interaction in the **A16/N19/Y23** mutant. (B) PEG restores β-sheet structure and reduces β-sheet fluctuation in the **A16/N19/F23** double mutant. PEG additions are shown in blue.

6B shows the "secondary structure" scenario for **A16/N19/F23**. As discussed earlier, the double mutant without PEG repeatedly loses and regains the top of β-strands 1 and 2 during the simulation. Once PEGylated, the protein is locked into a more sheet-rich structure throughout the same length simulation. The PEG oligomer rigidifies the loop, maintains zipped strands, and reduces structural fluctuations, restoring the double mutant to a "pseudo-wild-type-like" state. However, a simulation of the PEGylated protein starting with unraveled β-strands does not restore the pseudo-wild-type structure in ≤80 ns. Thus, the barrier between beta-rich and beta-poor states is smaller in the non-PEGylated double mutant than in the PEGylated double mutant.

The restoration of the β-strand structure in the double mutant by PEG-4 can also be seen in the Ramachandran angles of residues 18 and 20 in the loop, neighboring the attachment site. Figure 7 shows that Ser18 is in the Lα basin and Gly20 is in the polyproline-II region in the pseudo-wild type. These two residues flip upon double mutation, but they are retained in the PEGylated protein. The importance of local Ramachandran angles highlights that PEG can act by sterically constraining amino acids near its attachment site.

A Longer PEG Chain Samples Both Solvent Exposed and Protein Interaction Scenarios. We saw for PEG-4 that the pseudo-wild type and double mutant favor a solvent exposed PEG chain, while the single mutants favored new interactions of PEG with protein side chains not observed in the pseudo-wild type. Can a longer PEG chain display both scenarios simultaneously? We computed a long equilibrium trajectory of PEG-45 attached to Asn19 for comparison. Although still somewhat shorter than typical commercial PEG chains (N > 50, often branched), PEG-45 is similar in size to its hPin1 WW host protein. We find that PEG-45 can indeed sample both scenarios, and coat the protein sometimes, extend as a coil into the solvent at other times, with an equilibrium constant close to unity. We discuss this in more detail below.

DISCUSSION

Glycosylation and PEGylation can stabilize or destabilize proteins depending on the attachment site.[8−13] In addition, the context of the local sequence is important, with certain sequences favoring stabilization over others.[12,13] We showed here that one reasonable proposed sequence-dependent

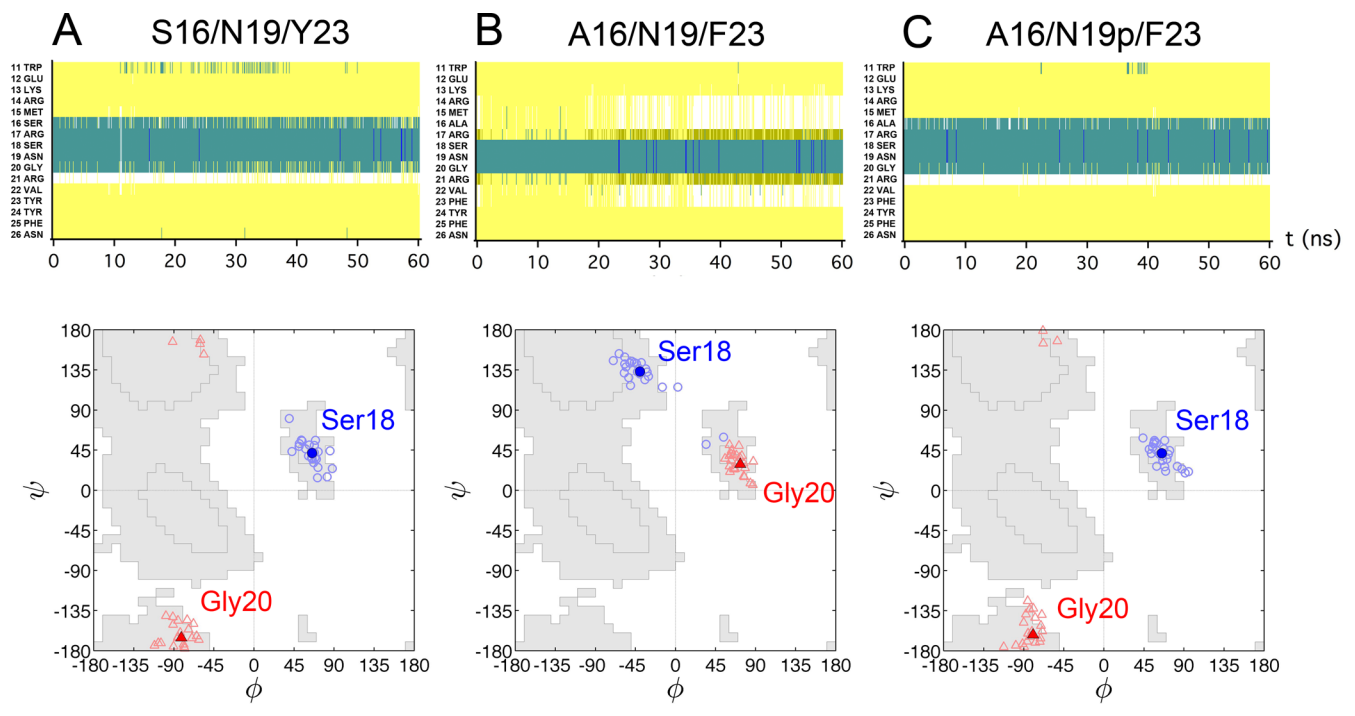


Figure 7. Secondary structure from residue 11 to 26 (upper) and Ramachandran plot for Ser18 and Gly20 (lower) of (A) pseudo-wild-type **S16/N19/Y23**, (B) double mutant **A16/N19/F23**, and (C) its PEGylated counterpart **A16/N19p/F23**. Secondary structure is calculated by VMD, with the following color code: turn, cyan; coil, white; β-sheet, yellow; 3-10 helix, blue. Dihedral angles are plotted every 2.5 ns, with the average values highlighted for Gly20 and Ser18 as shown.

dx.doi.org/10.1021/jp502234s | J. Phys. Chem. B 2014, 118, 8388−8395

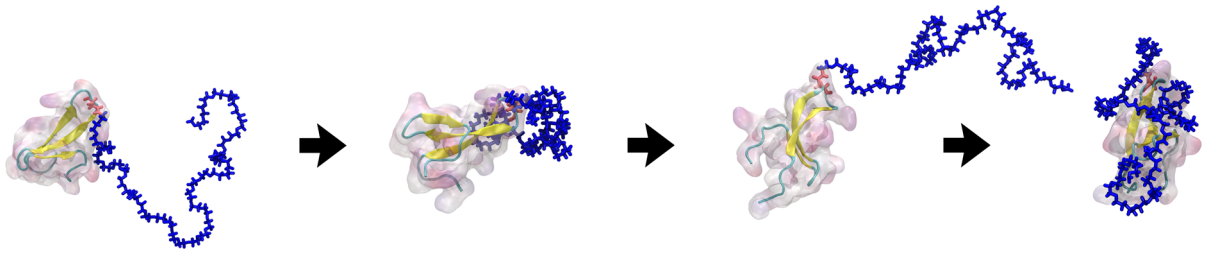


Figure 8. Conformational fluctuations of PEG-45 attached at Asn19. The conformation of PEG-45 alternates between solvated coil and protein surface-coating states. Each black arrow corresponds to ∼15 ns. An animation illustrating the full molecular dynamics trajectory is available in the Supporting Information (movie jp502234s_si_005.avi).

stabilization mechanism, direct hydrogen bonding of PEG to nearby OH groups on Ser16 and Tyr23, does not occur. Experimental measurement of hPin1 WW domain stabilization by PEG reveals a highly nonadditive effect in a truncation cycle of two OH groups near the PEG attachment site. Molecular dynamics simulation further reveals that PEG does not interact with the OH-bearing Ser16 and Tyr23 residues before or after any of the mutations.

Instead, the mutations affect the local loop and β-sheet structure, and PEG acts upon these mutants in two distinct ways. In one case (Ala16 single mutant), PEGylation restores a wild-type-like loop 1 orientation disrupted by a Met15−Asn19 interaction absent in the wild-type **S16/S19/Y23**; PEG favors hydrogen bonding to Arg21 on the binding side of the protein, in an orientation more similar to the wild-type Ser19 residue. In another case (double mutant), PEGylation allows the protein to retain a native-like β-sheet structure, disrupted by the double mutation when PEG is absent; PEG does so even though it hangs into the solvent, by local steric interaction at the attachment site (Figure 7). In two other cases, PEGylation maintains the status quo (pseudo-wild type and Phe23 single mutant).

The central conclusion is that PEG does not stabilize host proteins via a single general mechanism. Instead, whether PEG extends into the solvent to act only by local steric interaction, or whether the whole PEG chain interacts strongly with the protein surface, is highly context-dependent. One point mutation can switch the mechanism from one to the other, and it does not even have to be a mutation at a site that PEG strongly interacts with.

Although our simulations are not sufficiently long to directly compute free energies, together with the experimental data, we propose the following scenario for the unusually large stabilization (8 °C) of the double mutant by PEG. The double mutant is significantly less stable than either single mutant, and loses part of β-strands 1 and 2, while acquiring a more flexible loop 1. PEG restores native-like β-sheet content, and this leads to the additional stabilization and nonadditive effect. Thus, PEG makes up for the ca. 5 °C loss of stability of the double mutant relative to the single mutants by restoring structure, in addition to providing a similar stabilization as it does for the single mutants.

Clearly, the effect of PEGylation is very context sensitive, even when specific hydrogen bonds are not involved. This is evident from the thermodynamic data in Table 2. The stabilization of the pseudo-wild-type **S16/N19p/Y23** by PEG is entropic, while the stabilization of the **S16/N19p/F23** single mutant is enthalpic (Table 2). This difference is plausible in terms of Figure 5A and C: the pseudo-wild type has a native state with PEG coiled into the solution, maintaining high configurational entropy of PEG. In contrast, the PEG chain of the Phe23 mutant has low configurational entropy because it collapses onto the binding side of the protein to interact with Arg21. Of course the simulations cannot provide quantitative information beyond these propensities, as sampling is insufficient to compute entropies or free energies.

The observation that PEG-4 can stabilize its host protein by either coiling into the solvent or interacting with the protein surface suggests that enthalpic and entropic mechanisms of stabilization are almost balanced. This balance is a common observation in biomolecular interactions, and makes it harder to come up with simple "local contact" explanations for how mutations or PEG placement affect the free energy. Will this balance shift in favor of one or the other scenario when the PEG chain length is increased? Will PEG now either coat the protein surface, or extend as a coil into solution?

To investigate this question, we performed a molecular dynamics simulation on the **S16/N19p/Y23** pseudo-wild type with a PEG-45 attached. It has been shown that, relative to PEG-4, PEG-45 leaves the unfolding rate unchanged but decreases the folding rate: the longer chain still stabilizes the protein but a little bit less than PEG-4.[6] Figure 8 shows the result (jp502234s_si_005.avi, Supporting Information). The longer PEG constantly equilibrates back and forth between coating the protein surface and coiling into the solvent. Thus, the dichotomy between enthalpic and entropic stabilization mechanisms persists for PEG chain lengths approaching the type used in commercial drug PEGylation. It remains to be seen in future studies whether the context sensitivity of PEG-4 also persists in proteins modified with longer PEG chains. In particular, if PEG-based protection from protein degradation is as context sensitive as is the impact of PEG on conformational stability, careful rather than random or N-terminal placement could produce major improvements in peptide and protein drug half-life in the bloodstream.

■ ASSOCIATED CONTENT

Supporting Information

The PDF file describes additional measurements with text and figures. The movies show the molecular dynamics simulations discussed in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*Phone: 001-217-333-1624. E-mail: mgruebel@illinois.edu.
*E-mail: jlprice@chem.byu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support was provided by the National Science Foundation grant MCB 1019958 (M.G. and S.-H.C), the National Science Foundation grant PHY 0822613 (A.A. and S.-H.C), and start-up funds from BYU (S.S.M., R.P., and J.L.P.). The authors gladly acknowledge supercomputer time provided through XSEDE Allocation Grant MCA05S028 and the Taub Cluster (UIUC).

■ ABBREVIATIONS

PEG, poly(ethylene glycol); HPLC, high-performance liquid chromatography; Visual Molecular Dynamics, VMD; CHARMM, Chemistry at Harvard Molecular Mechanics

■ REFERENCES

(1) Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. Glycosylation and the Immune System. *Science* **2001**, *291*, 2370−2376.
(2) Harris, J. M.; Chess, R. B. Effect of Pegylation on Pharmaceuticals. *Nat. Rev. Drug Discovery* **2003**, *2*, 214−221.
(3) Veronese, F. M.; Pasut, G. Pegylation, Successful Approach to Drug Delivery. *Drug Discovery Today* **2005**, *10*, 1451−1458.
(4) Meng, W.; Guo, X.; Qin, M.; Pan, H.; Cao, Y.; Wang, W. Mechanistic Insights into the Stabilization of Srcsh3 by Pegylation. *Langmuir* **2012**, *28*, 16133−16140.
(5) Hamed, E.; Xu, T.; Keten, S. Poly(Ethylene Glycol) Conjugation Stabilizes the Secondary Structure of A-Helices by Reducing Peptide Solvent Accessible Surface Area. *Biomacromolecules* **2013**, *14*, 4053−4060.
(6) Pandey, B. K.; Smith, M. S.; Torgerson, C.; Lawrence, P. B.; Matthews, S. S.; Watkins, E.; Groves, M. L.; Prigozhin, M. B.; Price, J. L. Impact of Site-Specific Pegylation on the Conformational Stability and Folding Rate of the Pin Ww Domain Depends Strongly on Peg Oligomer Length. *Bioconjugate Chem.* **2013**, *24*, 796−802.
(7) Shental-Bechor, D.; Levy, Y. Effect of Glycosylation on Protein Folding: A Close Look at Thermodynamic Stabilization. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8256−8261.
(8) Chen, M. M.; Bartlett, A. I.; Nerenberg, P. S.; Friel, C. T.; Hackenberger, C. P. R.; Stultz, C. M.; Radford, S. E.; Imperiali, B. Perturbing the Folding Energy Landscape of the Bacterial Immunity Protein Im7 by Site-Specific N-Linked Glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22528−22533.
(9) Price, J. L.; Shental-Bechor, D.; Dhar, A.; Turner, M. J.; Powers, E. T.; Gruebele, M.; Levy, Y.; Kelly, J. W. Context-Dependent Effects of Asparagine Glycosylation on Pin Ww Folding Kinetics and Thermodynamics. *J. Am. Chem. Soc.* **2010**, *132*, 15359−15367.
(10) Price, J. L.; Powers, E. T.; Kelly, J. W. N-Pegylation of a Reverse Turn Is Stabilizing in Multiple Sequence Contexts, Unlike N-Glcnacylation. *ACS Chem. Biol.* **2011**, *6*, 1188−1192.
(11) Ellis, C. R.; Maiti, B.; Noid, W. G. Specific and Nonspecific Effects of Glycosylation. *J. Am. Chem. Soc.* **2012**, *134*, 8184−8193.
(12) Price, J. L.; Powers, D. L.; Powers, E. T.; Kelly, J. W. Glycosylation of the Enhanced Aromatic Sequon Is Similarly Stabilizing in Three Distinct Reverse Turn Contexts. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 14127−14132.
(13) Culyba, E. K.; Price, J. L.; Hanson, S. R.; Dhar, A.; Wong, C.-H.; Gruebele, M.; Powers, E. T.; Kelly, J. W. Protein Native-State Stabilization by Placing Aromatic Side Chains in N-Glycosylated Reverse Turns. *Science* **2011**, *331*, 571−575.
(14) *Poly(Ethylene Glycol) Chemistry*; Lim, K., Herron, J. N., Eds.; Springer: New York, 1992.
(15) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggests Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997**, *89*, 875−886.
(16) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(17) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with Namd. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(18) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(19) Beglov, D.; Roux, B. Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **1994**, *100*, 9050−9063.
(20) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(21) MacKerell, A. D.; Feig, M.; Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.
(22) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, J. A. D. Optimization of the Additive Charmm All-Atom Protein Force Field Targeting Improved Sampling of the Backbone Φ, Ψ and Side-Chain X 1and X 2dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.
(23) Vorobyov, I.; Anisimov, V. M.; Greene, S. Additive and Classical Drude Polarizable Force Fields for Linear and Cyclic Ethers. *J. Chem. Theory Comput.* **2007**, *3*, 1120−1133.
(24) Lee, H.; Venable, R. M.; MacKerell, A. D.; Pastor, R. W. Molecular Dynamics Studies of Polyethylene Oxide and Polyethylene Glycol: Hydrodynamic Radius and Shape Anisotropy. *Biophys. J.* **2008**, *95*, 1590−1599.
(25) Yoo, J.; Aksimentiev, A. Improved Parametrization of Li+, Na+, K+, and Mg2+ Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. *J. Phys. Chem. Lett.* **2012**, *3*, 45−50.
(26) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N • Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(27) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. Iterative Minimization Techniques for Ab Initio Total-Energy Calculations: Molecular Dynamics and Conjugate Gradients. *Rev. Mod. Phys.* **1992**, *64*, 1045−1097.
(28) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(29) Jäger, M.; Dendle, M.; Kelly, J. W. Sequence Determinants of Thermodynamic Stability in a WW Domain—an All-*β*-Sheet Protein. *Protein Sci.* **2009**, *18*, 1806−1813.