JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Criteria for Selecting PEGylation Sites on Proteins for Higher Thermodynamic and Proteolytic Stability

Paul B. Lawrence,[†,∥] Yulian Gavrilov,[‡,∥] Sam S. Matthews,[†] Minnie I. Langlois,[†] Dalit Shental-Bechor,[‡] Harry M. Greenblatt,[‡] Brijesh K. Pandey,[†] Mason S. Smith,[†] Ryan Paxman,[†] Chad D. Torgerson,[†] Jacob P. Merrell,[†] Cameron C. Ritz,[†] Maxim B. Prigozhin,[§] Yaakov Levy,*[,‡] and Joshua L. Price*[,†]

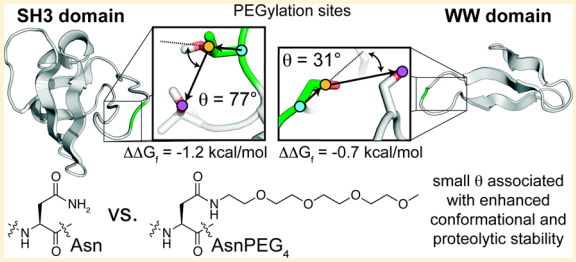
[†]Department of Chemistry and Biochemistry, Brigham Young University, Provo, Utah 84602, United States

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

[‡]Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel

Supporting Information

ABSTRACT: PEGylation of protein side chains has been used for more than 30 years to enhance the pharmacokinetic properties of protein drugs. However, there are no structure- or sequence-based guidelines for selecting sites that provide optimal PEG-based pharmacokinetic enhancement with minimal losses to biological activity. We hypothesize that globally optimal PEGylation sites are characterized by the ability of the PEG oligomer to increase protein conformational stability; however, the current understanding of how PEG influences the conformational stability of proteins is incomplete. Here we use the WW domain of the human protein Pin 1 (WW) as a model system to probe the impact of PEG on protein conformational stability. Using a combination of experimental and theoretical approaches, we develop a structure-based method for predicting which sites within WW are most likely to experience PEG-based stabilization, and we show that this method correctly predicts the location of a stabilizing PEGylation site within the chicken Src SH3 domain. PEG-based stabilization in WW is associated with enhanced resistance to proteolysis, is entropic in origin, and likely involves disruption by PEG of the network of hydrogen-bound solvent molecules that surround the protein. Our results highlight the possibility of using modern site-specific PEGylation techniques to install PEG oligomers at predetermined locations where PEG will provide optimal increases in conformational and proteolytic stability.



INTRODUCTION

PEGylation of protein side chains has been used for more than 30 years to enhance the pharmacokinetic properties of protein drugs.[1−9] Indeed, PEGylated versions of several therapeutic proteins are currently in clinical use.[10−20] Some PEGylated protein drugs are actually heterogeneous mixtures of isoforms that differ in the number and location of the attached PEG oligomers.[21] Others are PEGylated site-specifically at the N-terminus[22,23] or at a surface Cys residue.[24,25] The enhanced pharmacokinetic properties of these proteins are thought to derive from the large hydrodynamic radius of the attached PEG oligomer(s), which shield the protein surface from proteases and antibodies and which inhibit aggregation and clearance of the PEGylated protein through the kidneys.[1−8]

Nonspecific PEGylation can inadvertently place large PEGs near enzyme active sites or protein−protein binding interfaces, where steric hindrance results in decreased biological activity. Site-specific side-chain modification strategies now routinely allow researchers to avoid attaching PEG near such problematic locations.[26−33] However, it can be difficult to choose a suitable PEGylation site from among the many candidate surface-exposed residues that are sufficiently distant from active sites or binding interfaces. Such choices can be important: recent studies reveal that not all candidate PEGylation sites are equally optimal.[31] Are there additional structure- or sequence-based criteria for selecting sites that provide optimal PEG-based pharmacokinetic enhancement with minimal losses to biological activity? Given candidate PEGylation sites that are similarly distant from active sites or binding interfaces, we hypothesize that a distinguishing characteristic of optimal vs suboptimal sites is the ability of PEG to enhance protein conformational stability (i.e., the difference in free energy between the folded and unfolded protein conformations).

Previous reports indicate that conformational stability[34] is fundamentally related to protein aggregation propensity,[35−37] resistance to proteolysis,[38−43] and immunogenic potential[37,44−50] (i.e., exactly the kinds of pharmacokinetic problems PEGylation is supposed to ameliorate). Therefore, it seems reasonable to expect that PEG-based increases to conformational stability should be associated with enhanced protection from aggregation, proteolysis, and immunogenicity. However, the impact of PEGylation on protein conformational stability is incompletely understood. Indeed, PEGylation can in-

Received: September 15, 2014
Published: November 19, 2014

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

crease,[51−76] decrease,[77,78] or have no effect on protein conformational stability.[59,79−83] The molecular basis for these differences is unclear. We seek to identify rational structure-based guidelines for enhancing protein conformational and pharmacokinetic stability via PEGylation.

Here we use a small protein, the WW domain of the human protein Pin 1 (hereafter called WW) as a model system for understanding how PEGylation generally impacts the conformational stability of β-sheet proteins (Figure 1). The WW

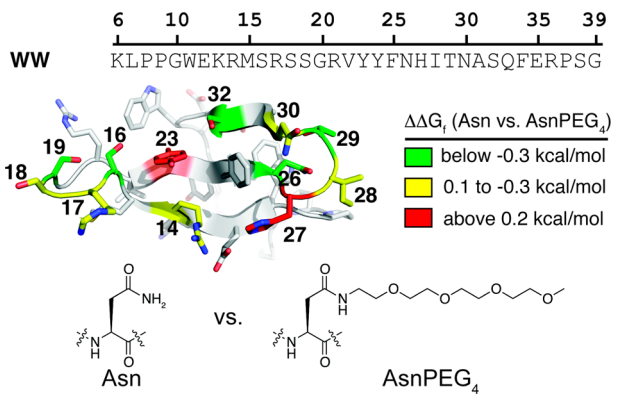


Figure 1. Sequence of the protein WW and ribbon diagram of WW (PDB ID: 1PIN), with side chains shown as sticks. Positions where we incorporated Asn vs $AsnPEG_4$ are highlighted with color, according to the impact of PEGylation on conformational stability. Stabilizing positions are highlighted in green, neutral positions are highlighted in yellow, and destabilizing positions are highlighted in red.

domain is an extensively characterized[84−106] β-sheet protein that contains three antiparallel β-strands connected by two reverse turns.[106] The folding free energy landscape of Pin WW can be approximated as a simple two-state reaction-coordinate diagram in which the unfolded ensemble proceeds through a high-energy transition state to the folded conformation without passing through discrete intermediates.[86] The small size of WW facilitates the direct chemical synthesis of homogeneous site-specifically PEGylated variants.[107,108] WW is much smaller than many of the PEGylated proteins of pharmaceutical interest. However, recent efforts to increase WW conformational stability via glycosylation have been successfully applied in two larger proteins,[104] suggesting that insights gained from WW PEGylation will be applicable to larger therapeutically relevant proteins.

We previously showed that PEGylating an Asn residue within a reverse turn substantially increases WW conformational stability by accelerating folding and slowing unfolding.[107,108] This increase in conformational stability is associated with protection from proteolysis, even though the PEG oligomer is relatively short (four ethylene oxide units), suggesting a link between conformational stability and optimal PEG-based enhancement to protein pharmacokinetic properties.

Here we identify additional locations within WW where PEGylation increases conformational stability and use a combination of experimental and computational approaches to probe the origins of PEG-based stabilization. We use the resulting insights, along with structural information for WW, to identify features that are common to stabilizing PEGylation sites. We use these structural features to develop criteria for predicting stabilizing PEGylation sites, and validate these criteria by correctly predicting the location of a stabilizing PEGylation site within the chicken Src SH3 domain. Finally, we show that PEG-based increases to conformational stability correlate with enhanced resistance to proteolysis. These results highlight the possibility of using modern site-specific PEGylation techniques to install PEG oligomers at locations that lead to optimal increases in conformational and proteolytic stability.

METHODS

WW variants were prepared via microwave-assisted solid-phase peptide synthesis, using a standard Fmoc Nα protection strategy as described previously (see the Supporting Information for details).[103,105,107] Fmoc-protected amino acids were obtained from Advanced Chem Tech, except for PEGylated Asn derivatives Fmoc-$AsnPEG_4$−OH and Fmoc-$AsnPEG_{45}$−OH, which were synthesized as described previously,[107] and Fmoc-D-$AsnPEG_4$−OH, which was synthesized as described in the Supporting Information. 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 (TFA). HPLC fractions containing the desired protein product were pooled, frozen, and lyophilized. Protein identity was confirmed by electrospray ionization time-of-flight mass spectrometry (ESI-TOF), and purity was assessed by analytical HPLC.

Conformational stability and folding kinetics of PEGylated WW variants and their non-PEGylated counterparts were assessed by variable-temperature circular dichroism spectropolarimetry (CD) and by fluorescence-based laser-induced temperature jump experiments, respectively, as described previously[108] (see the Supporting Information for details). Melting temperature, folding free energy, and folding and unfolding rate parameters were derived from global fits of the relevant data to equations based on a two-state folding model (see the Supporting Information for details).

To study the effect of the PEG on WW, we modeled the PEG at each of the experimentally studied sites on WW and the corresponding variants with Asn at these sites. The modeling was done using Coot7 software.[109] The models of Asn and Asn-PEG variants of WW were used to carry out atomistic molecular dynamics simulations. The molecular dynamics (MD) simulations were performed using GROMACS Version 4.5.4.[110] We used the AMBER99SB-ILDN force field,[111] which was modified to incorporate the PEG. All the modeled variants of WW were simulated for 300 ns each.

In addition, we studied the PEGylated WW proteins using a coarse-grained (CG) model based on the native topology of the WW protein (Gō model). This model was used in the past for numerous folding studies, in particular the folding of glycosylated WW proteins.[103] While the atomistic simulations focus on the folded state, the CG model focuses mostly on the unfolded state. All local, secondary, and tertiary native contacts between amino acids are represented by the Lennard-Jones potential without any discrimination between the various chemical types of the interactions. In the model, the PEGylated and non-PEGylated variants at each modified position include the same number of native interactions within WW, which all have the same strength. The simulated PEG can thus interact with the protein via excluded volume interactions only. The Hamiltonian of the system and its parameters can be found elsewhere.[112] The simulations were performed using the GROMACS software package. Multiple trajectories were simulated using the Langevin equation with a friction constant of 0.5 ps^{-1}.

RESULTS AND DISCUSSION

Impact of PEGylation on WW Conformational Stability. As described previously, PEGylation of an Asn residue at position 19 increases WW conformational stability by accelerating folding and slowing unfolding.[108] We wondered whether this effect was unique to position 19 or whether PEGylation might similarly stabilize other positions. To address

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

this question, we generated proteins **14**, **16**, **17**, **18**, **23**, **27**, **28**, **29**, and **32**, in which wild-type residues at positions 14, 16, 17, 18, 23, 27, 28, 29, and 32, respectively, have been changed to Asn (Asn already occupies positions 26 and 30 in the unmodified protein WW; see Figure 1). We also prepared PEGylated proteins **14p**, **16p**, **17p**, **18p**, **23p**, **26p**, **27p**, **28p**, **29p**, **30p**, and **32p**, in which positions 14, 17, 18, 19, 23, 26, 30, and 32, respectively, are occupied by $AsnPEG_4$, a PEGylated Asn derivative in which a four-unit PEG oligomer has been attached to the Asn side-chain amide nitrogen (Figure 1). These PEGylation sites sample the various secondary structural environments present in WW, including reverse turns (positions 16, 17, 18, 26, 27, 28, 29, and 30) and β-strands (positions 14, 23, and 32).

Circular dichroism (CD) spectra of these variants at 25 °C (Figure 2) are generally very similar in shape and magnitude to

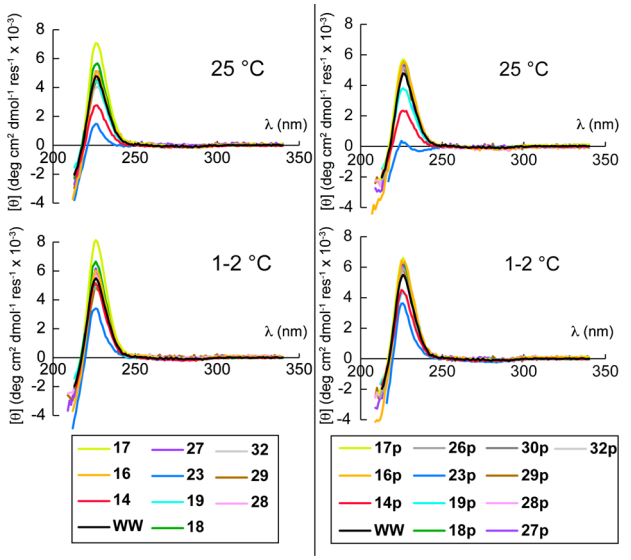


Figure 2. CD spectra of wild-type protein WW; non-PEGylated variants **14**, **16**, **17**, **18**, **19**, **23**, **27**, **28**, **29**, **32**; and PEGylated variants **14p**, **16p**, **17p**, **18p**, **19p**, **23p**, **26p**, **27p**, **28p**, **29p**, **30p**, and **32p** in 20 mM sodium phosphate buffer (pH 7) at 25 °C and at low temperature (i.e., at 2 °C, except for **16**, **27**, **27p**, and **29p**, which were analyzed at 1 °C; variable-temperature CD data for these compounds suggest that each is fully folded at 1 and 2 °C, so these spectra are directly comparable). Spectra were obtained at 100 μM, except for **16**, **16p**, **27**, **27p**, **28**, **28p**, **29**, and **29p**, which were obtained at 50 μM.

that of wild-type unmodified protein WW, suggesting that changing wild-type residues to Asn generally does not introduce dramatic alterations to the folded conformation of the resulting Asn mutants relative to WW. The exceptions to this trend are easily seen in the CD spectra of proteins **14** and **23**, and their PEGylated counterparts **14p** and **23p**, which are similar in shape to WW, though substantially smaller in magnitude.

Variable-temperature CD data for **14**, **14p**, **23**, and **23p** (see below) provide an explanation for this observation: **14**, **14p**, **23**, and **23p** appear to be two-state folders like WW, but are much less stable. Whereas **WW** is fully folded at 25 °C, **14**, **14p**, **23**, and **23p** each exist as equilibrium mixtures of fully folded and fully unfolded conformations at 25 °C. The CD spectrum of a two-state folder under equilibrium conditions is the weighted average of its fully folded and fully unfolded conformations. Therefore, the CD spectra of **14**, **14p**, **23**, and **23p** at 25 °C should be similar in shape but smaller in magnitude than the CD spectrum of WW. This is in fact what we observe. In contrast, variable-temperature CD data indicate that WW, **14**, **14p**, **23**, and **23p** should be each fully folded at 2 °C. Consistent with this expectation, the CD spectra of these variants at 2 °C (Figure 2) are much closer in magnitude to that of WW, suggesting that their fully folded conformations are likewise similar to that of WW. However, without high-resolution structural data, we cannot eliminate the possibility of substantial structural rearrangements in **14**, **14p**, **23**, and **23p**. Therefore, in the discussion below, we avoid using data from these compounds in our efforts to develop structure-based guidelines for identifying optimal PEGylation sites.

We used variable-temperature CD experiments to assess the conformational stability of PEGylated proteins **14p**, **16p**, **17p**, **18p**, **23p**, **26p**, **27p**, **28p**, **29p**, **30p**, and **32p** relative to their non-PEGylated counterparts **14**, **16**, **17**, **18**, **23**, WW, **27**, **28**, **29**, and **32** in 20 mM aqueous sodium phosphate (pH 7.0). We also performed these same measurements on 100 μM solutions of **19p** and **19**, which were characterized previously at 10 and 50 μM.[107,108] The results of this analysis appear in Figure 2 and Table 1. Variable-temperature CD data indicate that each of these variants is a two-state folder like the wild-type WW protein. PEGylation substantially increases WW conformational stability at positions 16, 19, 26, 29, and 32 and moderately increases WW conformational stability at position 17. In contrast, PEGylation has essentially no impact on WW conformational stability at positions 14, 18, 28, and 30 and is substantially destabilizing at positions 23 and 27. No specific secondary structural motif appears to be generally amenable to PEG-based stabilization: stabilizing and destabilizing positions occur within both β-strands and reverse turns.

Van't Hoff analysis allows us to parse the impact of PEGylation on WW conformational stability ($\Delta\Delta G_f$) into enthalpic ($\Delta\Delta H_f$) and entropic terms ($-T\Delta\Delta S_f$). At several positions, large uncertainties in $\Delta\Delta H_f$ and in $-T\Delta\Delta S_f$ preclude further analysis. However, an interesting trend emerges from the data for stabilizing positions 16, 19, 26, and 29 (Table 1). At each of these positions, $-T\Delta\Delta S_f$ is negative (i.e., favorable) whereas $\Delta\Delta H_f$ is positive (i.e., unfavorable). This observation suggests an entropic origin for the PEG-based increases to WW conformational stability at these positions.

The PEG oligomers used in therapeutic proteins are typically much longer than the four-unit oligomer we used in the experiments described above. We previously showed that attaching a 2000 Da PEG oligomer to an Asn at position 19 continues to increase WW conformational stability, even though the 2000 Da oligomer is much longer (~45 ethylene oxide units) than the four-unit oligomer.[108] We wondered whether the energetic impact of the 45-unit PEG at the positions described above would mirror the results described in Table 1 for the four-unit PEG. To test this hypothesis, we prepared WW variants **16p45**, **18p45**, **19p45**, **26p45**, **27p45**, **28p45**, and **29p45**, in which we incorporated an Asn-linked 45-unit PEG ($AsnPEG_{45}$) at positions 16, 18, 19, 26, 27, 28, and 29, respectively. We assessed the conformational stability of these variants relative to their non-PEGylated counterparts using variable-temperature CD experiments. The results of this analysis are shown in Table 2.

Like the 4-unit PEG, the 45-unit PEG increases conformational stability at positions 16, 19, 26, and 29, and decreases stability at position 27. Whereas the 4-unit PEG had no effect at positions 18 and 28, the 45-unit PEG is destabilizing at these

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

Table 1. Impact of the Four-Unit PEG Oligomer on WW Conformational Stability at Various Sites[a]

protein	sequence	T_m (°C)	ΔT_m (°C)	$\Delta\Delta G_f$ (kcal/mol)	$\Delta\Delta H_f$ (kcal/mol)	$-T\Delta\Delta S_f$ (kcal/mol)
14	KLPPGWEK**N**MSRSSGRVYYFNHITNASQFERPSG	34.0 ± 0.8	−0.6 ± 5.3	0.0 ± 0.4	−0.1 ± 3.4	0.1 ± 3.3
14p	KLPPGWEK**N**MSRSSGRVYYFNHITNASQFERPSG	33.4 ± 5.2				
16	KLPPGWEKRM**N**RSSGRVYYFNHITNASQFERPSG	50.6 ± 0.2	10.1 ± 0.3	−0.90 ± 0.03	3.8 ± 1.4	−4.7 ± 1.3
16p	KLPPGWEKRM**N**RSSGRVYYFNHITNASQFERPSG	60.7 ± 0.3				
17	KLPPGWEKRMS**N**SSGRVYYFNHITNASQFERPSG	53.6 ± 0.4	1.9 ± 0.6	−0.18 ± 0.05	−2.2 ± 0.8	2.0 ± 0.8
17p	KLPPGWEKRMS**N**SSGRVYYFNHITNASQFERPSG	55.5 ± 0.5				
18	KLPPGWEKRMSR**N**SGRVYYFNHITNASQFERPSG	56.9 ± 0.2	0.0 ± 0.7	0.00 ± 0.07	−3.7 ± 0.9	3.7 ± 0.9
18p	KLPPGWEKRMSR**N**SGRVYYFNHITNASQFERPSG	57.0 ± 0.7				
19	KLPPGWEKRMSRS**N**GRVYYFNHITNASQFERPSG	55.6 ± 0.2	7.7 ± 0.4	−0.70 ± 0.04	3.6 ± 1.4	−4.3 ± 1.4
19p	KLPPGWEKRMSRS**N**GRVYYFNHITNASQFERPSG	63.3 ± 0.3				
23	KLPPGWEKRMSRSSGRV**N**YFNHITNASQFERPSG	28.5 ± 0.9	−5.2 ± 1.3	0.40 ± 0.10	4.1 ± 1.4	−3.7 ± 1.3
23p	KLPPGWEKRMSRSSGRV**N**YFNHITNASQFERPSG	23.3 ± 1.0				
WW	KLPPGWEKRMSRSSGRVYYF**N**HITNASQFERPSG	58.0 ± 0.7	6.6 ± 0.7	−0.58 ± 0.06	3.4 ± 0.9	−4.0 ± 0.9
26p	KLPPGWEKRMSRSSGRVYYF**N**HITNASQFERPSG	64.6 ± 0.2				
27	KLPPGWEKRMSRSSGRVYYFN**N**ITNASQFERPSG	55.0 ± 0.1	−4.0 ± 0.4	0.38 ± 0.04	0.5 ± 0.9	−0.1 ± 0.9
27p	KLPPGWEKRMSRSSGRVYYFN**N**ITNASQFERPSG	51.0 ± 0.4				
28	KLPPGWEKRMSRSSGRVYYFNH**N**TNASQFERPSG	53.2 ± 0.5	0.0 ± 0.7	0.00 ± 0.07	0.6 ± 0.8	−0.6 ± 0.8
28p	KLPPGWEKRMSRSSGRVYYFNH**N**TNASQFERPSG	53.2 ± 0.5				
29	KLPPGWEKRMSRSSGRVYYFNHI**N**NASQFERPSG	50.0 ± 0.3	4.1 ± 0.4	−0.36 ± 0.04	3.6 ± 1.4	−4.3 ± 1.4
29p	KLPPGWEKRMSRSSGRVYYFNHI**N**NASQFERPSG	54.1 ± 0.3				
WW	KLPPGWEKRMSRSSGRVYYFNHIT**N**ASQFERPSG	58.0 ± 0.7	0.4 ± 0.7	0.00 ± 0.07	−0.5 ± 1.1	0.5 ± 1.1
30p	KLPPGWEKRMSRSSGRVYYFNHIT**N**ASQFERPSG	58.4 ± 0.2				
32	KLPPGWEKRMSRSSGRVYYFNHITNA**N**QFERPSG	45.1 ± 0.2	5.3 ± 0.3	−0.45 ± 0.02	−0.1 ± 0.6	−0.3 ± 0.6
32p	KLPPGWEKRMSRSSGRVYYFNHITNA**N**QFERPSG	50.3 ± 0.2				
16	KLPPGWEKRM**N**RSSGRVYYF**N**HITNASQFERPSG	50.6 ± 0.2	16.7 ± 0.2	−1.38 ± 0.03	8.2 ± 0.9	−9.6 ± 0.9
16p/26p	KLPPGWEKRM**N**RSSGRVYYF**N**HITNASQFERPSG	67.3 ± 0.1				
19	KLPPGWEKRMSRS**N**GRVYYF**N**HITNASQFERPSG	55.6 ± 0.2	14.2 ± 0.2	−1.26 ± 0.02	6.1 ± 0.7	−7.3 ± 0.7
19p/26p	KLPPGWEKRMSRS**N**GRVYYF**N**HITNASQFERPSG	69.8 ± 0.1				
29	KLPPGWEKRMSRSSGRVYYF**N**HI**N**NASQFERPSG	50.0 ± 0.3	6.7 ± 0.4	−0.56 ± 0.04	1.9 ± 0.6	−2.5 ± 0.6
26p/29p	KLPPGWEKRMSRSSGRVYYF**N**HI**N**NASQFERPSG	56.8 ± 0.3				
16/19	KLPPGWEKRM**N**RS**N**GRVYYFNHITNASQFERPSG	56.9 ± 0.1	8.5 ± 0.2	−0.80 ± 0.02	1.2 ± 0.5	−2.0 ± 0.5
16p/19p	KLPPGWEKRM**N**RS**N**GRVYYFNHITNASQFERPSG	65.4 ± 0.1				

[a]Only a selected portion of the entire WW sequence (residues 14−32) is shown here, with amino acids abbreviated according to the standard one-letter code. **N** represents $AsnPEG_4$. Observed data are given ± standard error at 100 μM protein concentration in 20 mM sodium phosphate buffer, pH 7 (except for proteins **27**, **27p**, **28**, **28p**, **29**, and **29p**, which were characterized at 50 μM protein concentration). Observed values of $\Delta\Delta G_f$ were derived from variable-temperature CD experiments at the melting temperature of the corresponding non-PEGylated protein.

Table 2. Impact of PEGylation with the 2000 Da (∼45-unit) Oligomer on WW Conformational Stability at Various Sites[a]

protein	T_m (°C)	ΔT_m (°C)	$\Delta\Delta G_f$ (kcal/mol)	$\Delta\Delta H_f$ (kcal/mol)	$-T\Delta\Delta S_f$ (kcal/mol)
16	54.9 ± 0.1	4.7 ± 0.3	−0.39 ± 0.03	5.1 ± 1.2	−5.5 ± 1.2
16p45	59.6 ± 0.3				
18	55.3 ± 0.8	−0.3 ± 0.9	0.02 ± 0.08	3.3 ± 1.5	−3.3 ± 1.5
18p45	55.1 ± 0.4				
19	55.9 ± 0.2	7.4 ± 0.4	−0.67 ± 0.05	1.7 ± 1.6	−2.4 ± 1.5
19p45	63.3 ± 0.4				
WW	58.0 ± 0.7	3.3 ± 0.7	−0.27 ± 0.06	5.1 ± 0.8	−5.3 ± 0.8
26p45	61.3 ± 0.1				
27	55.0 ± 0.1	−7.0 ± 0.3	0.65 ± 0.04	−0.4 ± 1.0	1.1 ± 1.0
27p45	48.0 ± 0.3				
28	53.2 ± 0.5	−4.5 ± 0.6	0.36 ± 0.05	4.0 ± 1.0	−3.7 ± 1.0
28p45	48.7 ± 0.3				
29	48.6 ± 0.4	4.7 ± 0.5	−0.36 ± 0.04	1.6 ± 1.3	−2.0 ± 1.3
29p45	53.3 ± 0.3				

[a]Data are given ± standard error at 50 μM protein concentration in 20 mM sodium phosphate buffer, pH 7 (except for proteins WW and **26p45**, which were characterized at 100 μM protein concentration) at the melting temperature of the corresponding non-PEGylated protein. Values of T_m, $\Delta\Delta G_f$, $\Delta\Delta H_f$, and $-T\Delta\Delta S_f$ were derived from variable-temperature CD experiments.

positions. Van't Hoff analysis of these results indicates that $-T\Delta\Delta S_f$ is negative (i.e., favorable) and $\Delta\Delta H_f$ is positive (i.e., unfavorable) at stabilizing positions 16, 19, 26, and 29, suggesting that the 45-unit oligomer likewise increases WW

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

Table 3. Effect of Mutagenesis near Selected PEGylation Sites within WW[a]

protein	sequence	T_m (°C)	ΔT_m (°C)	$\Delta\Delta G_f$ (kcal/mol)	$\Delta\Delta H_f$ (kcal/mol)	$-T\Delta\Delta S_f$ (kcal/mol)
19	KLPPGWEKRMSRS**N**GRVYYFNHITNASQFERPSG	55.6 ± 0.2	7.7 ± 0.4	−0.70 ± 0.04	3.6 ± 1.4	−4.3 ± 1.4
19p	KLPPGWEKRMSRS**N**GRVYYFNHITNASQFERPSG	63.3 ± 0.3				
D-19	KLPPGWEKRMSRS**n**GRVYYFNHITNASQFERPSG	55.4 ± 0.3	−0.2 ± 0.4	0.01 ± 0.04	0.7 ± 0.5	−0.7 ± 0.5
D-19p	KLPPGWEKRMSRS**n**GRVYYFNHITNASQFERPSG	55.2 ± 0.3				
19-16A	KLPPGWEKRM**A**RS**N**GRVYYFNHITNASQFERPSG	51.0 ± 0.2	5.8 ± 0.3	−0.51 ± 0.02	2.5 ± 0.5	−3.0 ± 0.5
19p-16A	KLPPGWEKRMARS**N**GRVYYFNHITNASQFERPSG	56.8 ± 0.1				
19-23F	KLPPGWEKRMSRS**N**GRV**F**YFNHITNASQFERPSG	51.4 ± 0.4	5.0 ± 0.5	−0.43 ± 0.03	1.8 ± 0.6	−2.2 ± 0.6
19p-23F	KLPPGWEKRMSRS**N**GRV**F**YFNHITNASQFERPSG	56.5 ± 0.1				
19-32A	KLPPGWEKRMSRS**N**GRVYYFNHITN**A**A**Q**FERPSG	54.5 ± 0.3	8.4 ± 0.3	−0.71 ± 0.03	2.2 ± 0.4	−2.9 ± 0.4
19p-32A	KLPPGWEKRMSRS**N**GRVYYFNHITN**A**A**Q**FERPSG	62.8 ± 0.1				
19-16A,23F	KLPPGWEKRM**A**RS**N**GRV**F**YFNHITNASQFERPSG	45.8 ± 1.0	8.0 ± 1.1	−0.72 ± 0.08	−4.2 ± 1.5	3.5 ± 1.5
19p-16A,23F	KLPPGWEKRMARS**N**GAV**F**YFNHITNASQFERPSG	53.7 ± 0.3				
16	KLPPGWEKRM**N**RSSGRVYYFNHITNASQFERPSG	50.6 ± 0.2	10.1 ± 0.3	−0.90 ± 0.04	3.8 ± 1.4	−4.7 ± 1.3
16p	KLPPGWEKRM**N**RSSGRVFYFNHITNASQFERPSG	60.7 ± 0.3				
16-Y23F	KLPPGWEKRM**N**RSSGRV**F**YFNHITNASQFERPSG	50.9 ± 0.6	5.5 ± 0.7	−0.45 ± 0.06	1.7 ± 1.5	−2.1 ± 1.5
16p-Y23F	KLPPGWEKRM**N**RSSGRV**F**YFNHITNASQFERPSG	56.4 ± 0.3				
16-S32A	KLPPGWEKRM**N**RSSGRVYYFNHITN**A**A**Q**FERPSG	55.0 ± 0.3	6.6 ± 0.3	−0.58 ± 0.03	2.3 ± 0.5	−2.9 ± 0.5
16p-S32A	KLPPGWEKRM**N**RSSGRVYYFNHITN**A**A**Q**FERPSG	61.6 ± 0.1				
WW	KLPPGWEKRMSRSSGRVYYF**N**HITNASQFERPSG	58.0 ± 0.7	6.6 ± 0.7	−0.58 ± 0.06	3.4 ± 0.9	−4.0 ± 0.9
26p	KLPPGWEKRMSRSSGRVYYF**N**HITNASQFERPSG	64.6 ± 0.2				
WW-T29A	KLPPGWEKRMSRSSGRVYYF**N**HI**A**NASQFERPSG	40.4 ± 0.7	4.8 ± 0.8	−0.32 ± 0.06	4.0 ± 0.7	−4.4 ± 0.7
26p-T29A	KLPPGWEKRMSRSSGRVYYF**N**HI**A**NASQFERPSG	45.2 ± 0.2				
29	KLPPGWEKRMSRSSGRVYYFNHI**N**NASQFERPSG	50.0 ± 0.3	4.1 ± 0.4	−0.36 ± 0.04	0.3 ± 0.5	−0.6 ± 0.5
29p	KLPPGWEKRMSRSSGRVYYFNHI**N**NASQFERPSG	54.1 ± 0.3				
29-S32A	KLPPGWEKRMSRSSGRVYYFNHI**N**N**A**A**Q**FERPSG	40.4 ± 0.5	11.6 ± 0.8	−0.88 ± 0.06	−4.1 ± 1.7	3.3 ± 1.7
29p-S32A	KLPPGWEKRMSRSSGRVYYFNHI**N**N**A**A**Q**FERPSG	52.0 ± 0.7				
32	KLPPGWEKRMSRSSGRVYYF**N**HITNANQFERPSG	45.1 ± 0.2	5.3 ± 0.3	−0.45 ± 0.02	−0.1 ± 0.6	−0.3 ± 0.6
32p	KLPPGWEKRMSRSSGRVYYFNHITNA**N**QFERPSG	50.3 ± 0.2				
32-Y23F	KLPPGWEKRMSRSSGRV**F**YFNHITNA**N**QFERPSG	30.0 ± 1.3	10.3 ± 1.5	−0.61 ± 0.11	7.7 ± 2.9	−8.3 ± 2.9
32p-Y32F	KLPPGWEKRMSRSSGRVFY**F**NHITNA**N**QFERPSG	40.3 ± 0.7				

[a]Data are given ± standard error at 100 μM protein concentration in 20 mM sodium phosphate buffer (pH 7) at the melting temperature of the corresponding non-PEGylated protein. **N** represents AsnPEG$_4$. Values of T_m, $\Delta\Delta G_f$, $\Delta\Delta H_f$, and $-T\Delta\Delta S_f$ were derived from variable-temperature CD experiments.

stability via an entropic effect. The close correlation between the position-dependent results for the 4-unit and 45-unit oligomers in these experiments suggests that insights gained from the 4-unit oligomer should be reasonably predictive for longer oligomers that more closely resemble those currently used in therapeutic proteins.

PEGylation at Two Stabilizing Sites. We next wondered whether simultaneously PEGylating two of the identified "stabilizing" positions would result in more substantial increases to WW conformational stability. To address this question, we prepared doubly PEGylated proteins **16p/26p**, **19p/26p**, **26p/29p**, and **16p/19p** and compared them to their non-PEGylated counterparts (**16**, **19**, **16/19**, and **29**, respectively; see Table 1 for the sequences of these proteins). We assessed the conformational stability of these variants using variable-temperature CD experiments; results are shown in Table 1. Doubly PEGylated compounds **16p/26p**, **19p/26p**, **26p/29p**, and **16p/19p** are −1.38 ± 0.03, −1.26 ± 0.02, −0.56 ± 0.04, and −0.80 ± 0.02 kcal mol^{-1} more stable, respectively, than their non-PEGylated counterparts.

Double-mutant cycle analysis of **19p/26p** and its mono- and non-PEGylated counterparts (see the Supporting Information for details) indicates that the two PEG oligomers at positions 19 and 26 contribute independently and additively to WW stability. In contrast, the two PEG oligomers in **16p/26p**, **19p/26p**, **26p/29p**, and **16p/19p** do not contribute additively to WW conformational stability: the overall impact of the two PEG oligomers is smaller than what one would expect on the basis of the impact of each PEG oligomer individually (see the Supporting Information for details). These observations could reflect steric clashes between the two PEG oligomers due to their proximity to each other. For example, in **26p/29p**, the two PEG oligomers are close to each other in primary sequence and are each part of the same reverse turn. The same is true in **16p/19p**. Alternatively, it is possible that in these variants, the two PEG oligomers have to compete with each other for the same favorable interactions with nearby residues.

Mechanistic Origins of PEG-Based Stabilization. We next used temperature-jump kinetic experiments to assess the contribution of folding and unfolding kinetics to the PEG-based changes in conformational stability described above, with the goal of gaining insights into how PEG can stabilize proteins. At stabilizing positions 19 and 26, and to a lesser extent at position 17, PEGylation accelerates folding and slows unfolding (see the Supporting Information for details). In contrast, at neutral positions 14, 18, and 30, PEGylation slows both folding and unfolding by similar amounts, resulting in no overall change to folding thermodynamics. Accelerated folding and slowed unfolding could be consistent with simultaneous stabilization of the native state and the transition state, with the native state

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

experiencing greater stabilization. Alternatively, these observations are also consistent with simultaneous destabilization of the unfolded ensemble and transition state, with the unfolded ensemble experiencing greater destabilization.

To help discern between these two mechanistic possibilities, we studied PEGylated proteins **14p**, **16p**, **17p**, **18p**, **19p**, **23p**, **26p**, and **30p** and their non-PEGylated counterparts using a coarse-grained native-topology-based (CG) model in which only the heavy atoms of the protein and the PEG conjugate are included. We have previously used similar models[113−116] to study the impact of glycosylation, ubiquitination, and myristoylation on protein folding. In this CG approach, the PEG is modeled as an excluded-volume polymer, which is exposed to the solvent and cannot form favorable interactions with protein side-chain or backbone groups. Therefore, PEG-based changes to WW conformational stability in this model are assumed to come from changes in the free energy of (1) the unfolded ensemble, which might not be as compact as the native state and might therefore be more affected by an excluded-volume PEG oligomer, or (2) the native state, due to unfavorable steric interactions between the PEG oligomer and the protein.

The CG model captures the observed destabilization of **23p** and **27p** relative to **23** and **27**, respectively (see the Supporting Information for details). For the variants where PEGylation has no substantial observed impact on conformational stability (**14p**, **18p**, **28p**, **30p**), the CG model simulations also predict a minimal effect, with the exception of position 14, where the CG model predicts strong destabilization. The small effect of PEG on stability for these variants is also reflected by the kinetic rates predicted from the CG simulations (see the Supporting Information for details). A more substantial disagreement between the CG model and experimental observations is seen at stabilizing positions 16, 19, 26, and 29. The CG model predicts that PEGylation will strongly destabilize **16p** and **26p** relative to **16** and WW, respectively, and have a minimal effect on **19p** and **29p** relative to **19** and **29**, respectively. In contrast, we observe strong stabilization at each of these positions. The limited predictive power of the CG model suggests that the observed PEG-based stabilization and acceleration of folding do not come from an excluded volume effect.

An alternative to this mechanistic hypothesis is that PEG-based increases to conformational stability come from stabilization of the transition state and native state relative to the unfolded ensemble. This scenario could potentially involve favorable PEG−protein interactions in the transition state and in the native state. In the crystal structure of the parent WW domain, the side chain at position 19 appears to be oriented toward several nearby OH-containing side chains, including Ser16, Tyr23, and Ser32 (Figure 1). We wondered whether interactions between PEG and nearby OH groups contribute to the observed PEG-based stabilization. If so, the orientation of the side chain at position 19 should also be an important factor.

To test this hypothesis, we prepared proteins D-**19** and D-**19p**, in which D-Asn or D-AsnPEG$_4$ occupy position 19, respectively (D-AsnPEG$_4$ is the enantiomer of AsnPEG$_4$ shown in Figure 1). Incorporating D-Asn or D-AsnPEG$_4$ should invert the orientation of the side chain at this position. Previous work by Kelly and co-workers indicates that WW can tolerate D-amino acids within this reverse turn without substantial disruption of secondary and tertiary structure.[87] The observed similarity of the CD spectra of D-**19** and D-**19p** to those of their counterparts **19** and **19p** (see the Supporting Information) is consistent with this assertion, as are the nearly identical melting temperatures of **19** and D-**19** (55.6 ± 0.3 and 55.4 ± 0.3 °C, respectively. Whereas PEGylation of Asn at position 19 increases WW conformational stability by −0.70 ± 0.04 kcal/mol, PEGylation of D-Asn at position 19 has no effect ($\Delta\Delta G_f$ = 0.01 ± 0.04 kcal/mol), suggesting that side-chain orientation is an important feature of stabilizing PEGylation sites.

We recently probed the extent to which the Asn-linked PEG-oligomer at position 19 engages in favorable interactions with nearby Ser16 and Tyr23 side chains.[117] For convenience, these previously reported data are also shown in Table 3. Removing the OH group at position 16 by replacing Ser with Ala reduces the stabilizing impact of PEGylation from −0.70 ± 0.04 kcal mol^{-1} (compare **19p** vs **19**) to −0.51 ± 0.02 kcal mol^{-1} (compare **19p-16A** vs **19-16A** in Table 3). Similarly, replacing Tyr23 with Phe reduces the stabilizing impact of PEGylation to −0.43 ± 0.03 (compare **19p-23F** vs **19-23F** in Table 3). In contrast, we observed here that replacing Ser32 with Ala has no significant effect (compare **19p-32A** vs **19-32A** in Table 3). In principle, these results could be interpreted in terms of direct favorable interactions between PEG at position 19 and the OH groups at positions 16 and 23, but not at position 32, presumably because of its distance from position 19.

However, direct PEG−OH interactions are absent from previously reported MD simulations of **19p**,[117] suggesting that such interactions are not a significant component of PEG-based stabilization. Instead, the simulations show the PEG oligomer at position 19 extending predominantly into the solvent, with a high degree of conformational entropy. The flexible PEG oligomer also appears to increase the conformational entropy of amino acid residues within **19p** relative to **19** (as measured by root-mean-square deviations of the simulated structures for **19p** and **19** vs the crystal structure of the parent WW protein), but without substantially disrupting the native-state interactions present in the reverse turns and β-strands of **19p**.[117] These simulations are consistent with our observations that PEG-based stabilization at position 19 is associated with an unfavorable increase in enthalpy, which is offset by a favorable increase in entropy (Table 1). Moreover, the simulations imply that the influence of nearby OH groups on PEG-based stabilization at position 19 must occur via an indirect mechanism rather than via direct PEG−OH contacts.

We wondered whether OH groups near other "stabilizing" PEGylation sites might be similarly (though indirectly) important to the observed PEG-based stabilization. To address this question, we identified one or more OH-containing side chains (Ser, Thr, or Tyr) near positions 16, 26, 29, and 32, and replaced these residues individually with Ala or Phe. The results of this analysis are shown in Table 3. Replacing Tyr23 or Ser32 with Phe or Ala, respectively, decreases the stabilizing impact of PEGylation at position 16 (Table 3; compare **16p-23F** vs **16-23F**, and **16p-32A** vs **16-32A**). Similarly, replacing Thr29 with Ala decreases the stabilizing impact of PEGylation at position 26 (Table 3; compare **26p-29A** vs **26-29A**). In contrast, PEG-based stabilization actually increases at positions 29 and 32 upon removal of OH groups at Ser32 and Tyr23, respectively (Table 3; compare **29p-32A** vs **29-32A**, and **32p-23F** vs **32-23F**). Interpretation of these last two results is complicated by the strong destabilizing impact of the Ser32Ala and Tyr23Phe mutations in these variants. In any case, these mutagenesis experiments are difficult to rationalize on the basis of direct favorable PEG−OH interactions and hint at a more indirect influence.

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

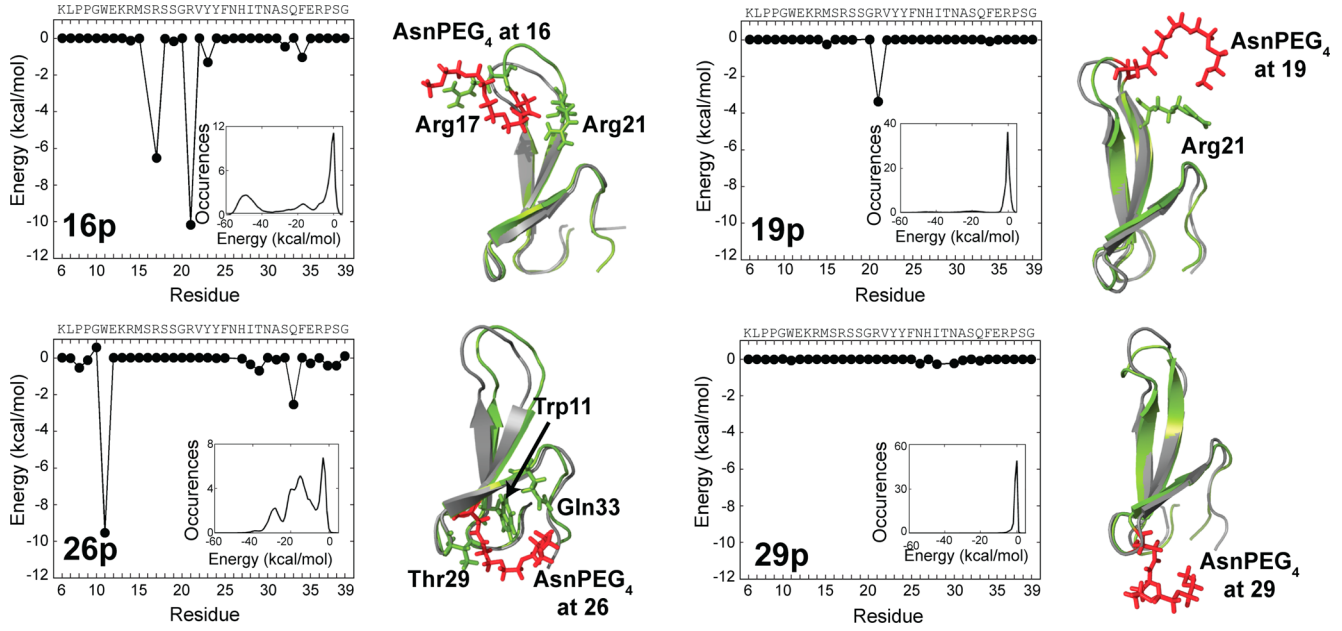


Figure 3. Results of atomistic simulations of WW variants PEGylated at stabilizing positions 16, 19, 26, and 29. Large plots show the average interaction energy between PEG and every other residue within WW. Insets show a histogram of this interface for each variant: a sharp peak near 0 kcal mol^{-1} denotes a mostly solvent exposed PEG oligomer that does not engage extensively in PEG−protein interactions. A histogram with lower energy peaks denotes the presence of more stable PEG−protein interactions. A broad histogram suggests a diverse, low-specificity PEG−protein interaction interface. Snapshots from each simulation of each variant (green) are overlaid with the crystal structure of the unmodified protein WW (gray). $AsnPEG_4$ is shown as red sticks. Side chains that appear to engage in interactions with the PEG are shown as green sticks. The simulated PEG−protein interactions are sometimes relatively transient; therefore, snapshots are not necessarily representative of the entire simulation.

In agreement with this conclusion, fully atomistic simulations of PEGylated proteins **16p**, **19p**, **26p**, and **29p** provide no evidence for strong direct PEG−OH interactions. Figure 3 shows the results of these simulations. For each variant, we calculated the average interaction energy between PEG and every residue within WW (Figure 3, large graphs), along with the total energy of PEG−protein interface during the simulation (Figure 3, insets). Snapshots from the simulation of each variant are also shown in Figure 3, though these do not indicate the lifetime of individual interactions, which in some cases are relatively transient.

For **16p**, the simulations suggest the presence of strong interactions between PEG and Arg17 and Arg21. Similarly, the PEG in **26p** appears to interact strongly with Trp11 and Gln33. However, strong, tight PEG−protein interfaces also occur in simulations of destabilized variants **23p** and **27p** (see the Supporting Information), suggesting that favorable PEG−protein interactions are not sufficient for increasing the overall conformational stability of WW. Moreover, simulations of **19p** and **29p** show that PEG-based stabilization can occur even in the absence of strong PEG−protein interactions (Figure 3), indicating that direct PEG−OH interactions are not responsible for the observed impact of OH groups on PEG-based stabilization. Other factors, including the conformational entropy of PEG as well the solvation of WW surface residues, must also make important contributions.

For example, the favorable PEG−protein interactions present in the rigid interface between PEG and **23p** or between PEG and 27p may not be sufficient to compensate for the reduced entropy of the PEG oligomer upon binding tightly to the protein surface, resulting in net PEG-based destabilization. In contrast, the broader, more flexible PEG−protein interfaces present in **16p** and **26p** (indicated by the broad interface energy distributions for these compounds; see Figure 3) may not reduce the entropy of the PEG oligomer as much, allowing for net PEG-based stabilization.

However, the ability of a highly flexible PEG to stabilize some WW variants even in the absence of strong PEG−protein interactions, together with the observed impact of nearby OH groups described above, suggests the possibility that changes in WW solvation may also play a role in PEG-based stabilization. One possibility is that differential solvation of these nearby OH groups in the presence or absence of PEG affects protein conformational stability.

We investigated this possibility in the simulations of **16p** and **26p** by analyzing the organization of water near residues that interact strongly with PEG. Figure 4 shows plots of the radial distribution function of water about Tyr23 or Phe34 in **16p** vs **16** and about Trp11 or Thr29 in **26p** vs **26**. These radial distribution function plots show the density of water molecules as a function of the distance from the indicated residues in **16p** and **26p** vs **16** and WW, respectively. For proteins **16p** vs **16**, PEGylation results in lower water density (and therefore higher water disorder) around Tyr23 and Phe34. Interestingly, this change in water molecule organization is long-range and can extend out to 10 Å from the protein, indicating that PEG not only affects the first hydration shell but also more distant shells. We observe similar effects in the water around Trp11 and Thr29 in proteins **26p** vs WW. The insets in each panel of Figure 4 show histograms of the number of water molecules observed in the simulations at a distance <3 Å from the indicated side chain (i.e., the first hydration shell). In **26p**, PEG results in a decrease in the number of water molecules (i.e., dehydration) in the first solvation shell around Trp11. A smaller amount of dehydration occurs about Phe34 in **16p**.

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

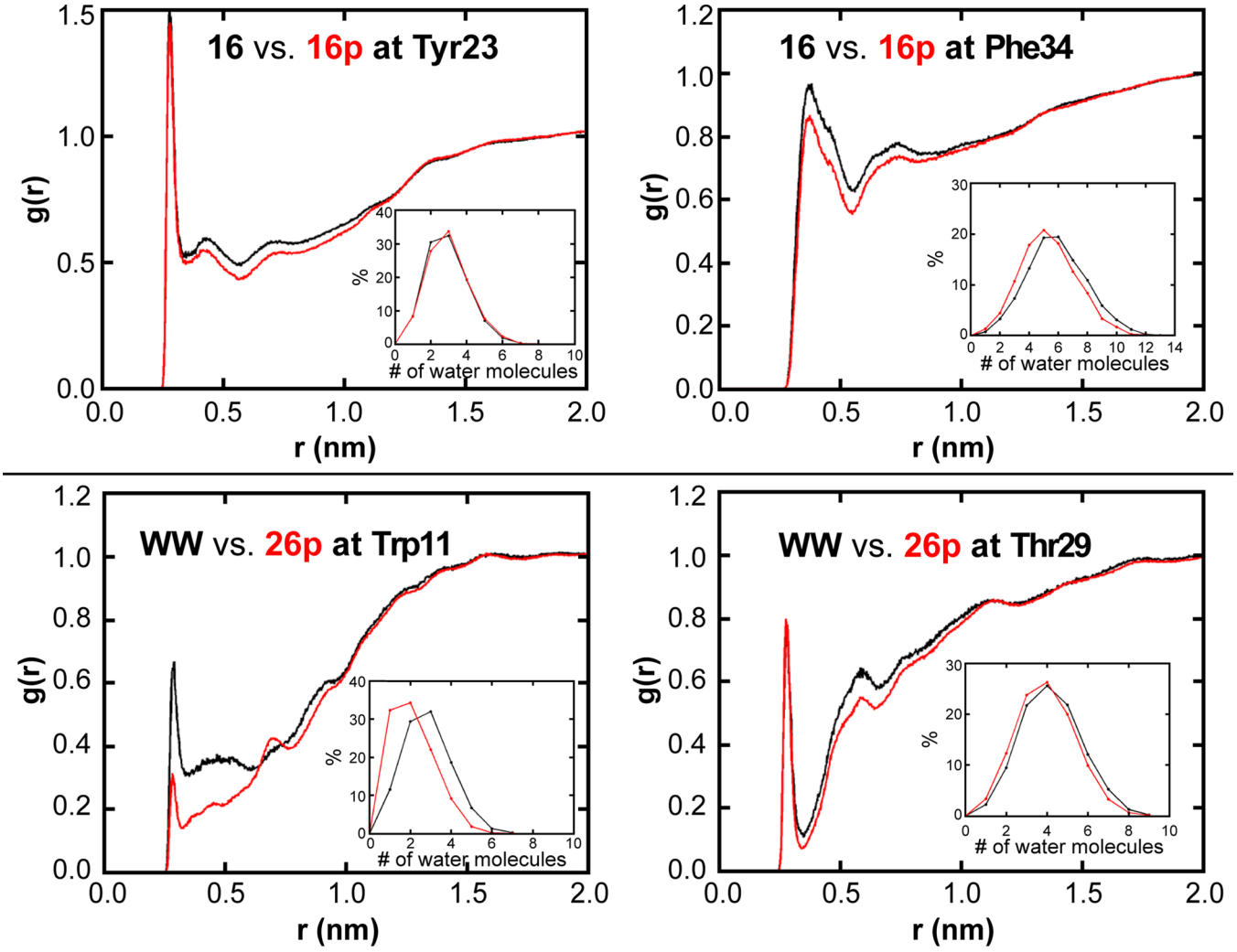


Figure 4. Simulated radial distribution function of water about Tyr23 or Phe34 acids in **16** vs **16p** (top panels) or about Trp11 or Thr29 in WW vs **26p** (bottom panels). Insets show histograms of the number of water molecules at a distance of <3 Å from the indicated side chains (i.e., the first hydration shell).

We would expect this PEG-based dehydration (i.e., release of water from the protein surface to bulk solvent) to be entropically favorable, offset by a smaller increase in enthalpy due to the loss of protein−water hydrogen bonds, an expectation consistent with our earlier observations that PEG-based stabilization is entropic in origin (Table 1; compare **19p** vs **19**: $\Delta\Delta H_f = 3.6 \pm 1.4$ kcal mol^{-1}, $-T\Delta\Delta S_f = -4.3 \pm 1.4$ kcal mol^{-1}). We speculate that this dehydration effect is more pronounced near water-binding OH groups and is the origin of the observed impact of OH groups on the PEG-based stabilization of WW.

We explored this possibility experimentally by assessing the impact of increasing amounts of heavy water (D_2O) on the conformational stability of **19p** vs **19**. The results of this analysis are shown in Figure 5A. Non-PEGylated **19** is -0.47 ± 0.05 kcal mol^{-1} more stable in buffer containing 98% D_2O than in buffer containing no D_2O. In contrast, a similar increase in D_2O only increases the stability of **19p** by -0.25 ± 0.04 kcal mol^{-1}. Previous studies indicate that D_2O decreases the internal flexibility and increases the conformational stability of proteins and suggest that the origin of this effect lies in the increased strength of the noncovalent O−D···X interaction (i.e., a deuterium bond) relative to the noncovalent O−H···X interaction (i.e., a hydrogen bond).[118] This difference in strength provides an energetic incentive for the oxygen atoms within D_2O to engage in more solvent−solvent deuterium bonds and fewer solvent−protein hydrogen bonds. This effect increases the compactness of the folded protein and makes unfolding less favorable. We hypothesize that increasing D_2O concentration to 98% stabilizes non-PEGylated **19** more profoundly than PEGylated **19p** because **19p** is less solvated than **19**, with fewer solvent−protein hydrogen bonds to replace with stronger solvent−solvent deuterium bonds.

Increasing the D_2O concentration also affects the heat capacity change due to folding (ΔC_p) for **19** and **19p** (Figure 4B). For **19**, ΔC_p increases from -0.66 ± 0.07 kcal mol^{-1} K^{-1} (no D_2O) to 0.01 ± 0.12 kcal mol^{-1} K^{-1} (98% D_2O). In contrast, the ΔC_p for **19p** (-0.74 ± 0.05 kcal mol^{-1} K^{-1} in H_2O) is not substantially affected by increasing amounts of D_2O. In the context of protein folding, negative values of ΔC_p are associated with folding processes that decrease the amount of solvent-accessible surface area by burying nonpolar side chains (or, alternatively, with unfolding processes that increase solvent-accessible surface area by exposing nonpolar side chains to solvent).[119] We hypothesize that increasing D_2O concentration makes ΔC_p of **19** less negative because the unfolded

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

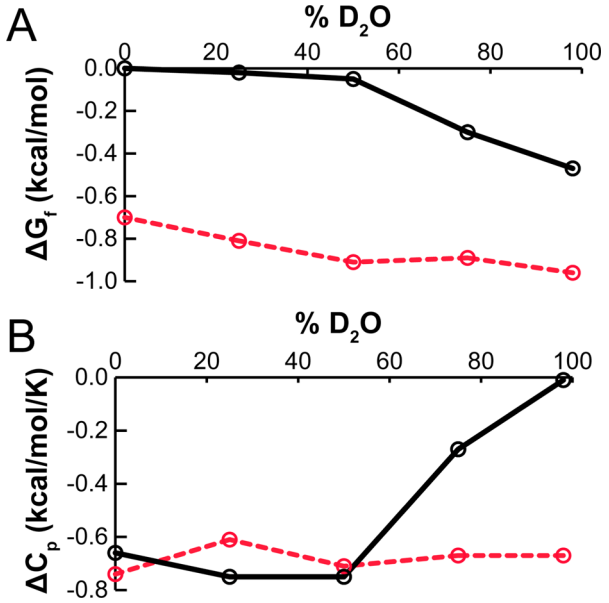


Figure 5. Change in (A) conformational stability (ΔG_f) or (B) heat capacity (ΔC_p) associated with folding of **19** (black solid line) or **19p** (red dashed line) in the presence of increasing amounts of D_2O in 20 mM phosphate buffer (pH 7) at 25 °C and at a concentration of 100 μM.

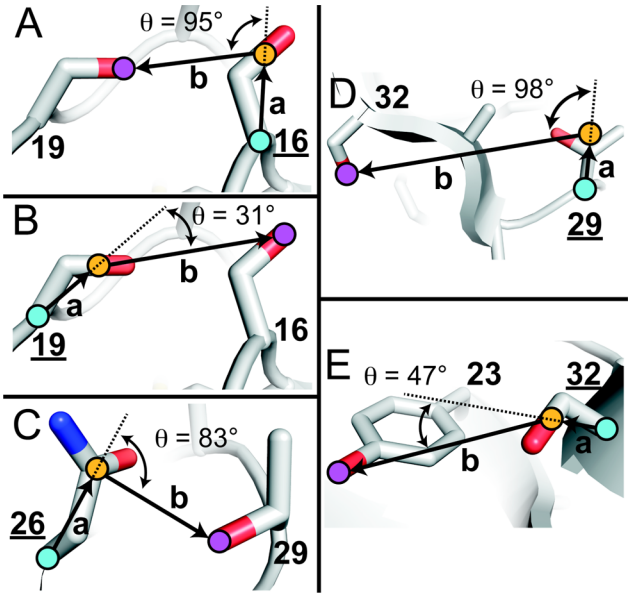


Figure 6. Angle θ between vectors **a** and **b** at positions (A) 16, (B) 19, (C) 26, (D) 29, and (E) 32. Alpha carbons (Cα), side-chain centers-of-mass (COM), and oxygens of the nearest OH-containing side chain are highlighted with blue-, orange-, and purple-filled circles, respectively.

conformation of **19** in D_2O is more compact, with less exposed nonpolar surface area than the unfolded conformation of **19** in H_2O (i.e., the stronger network deuterium bonds in D_2O more effectively constrains the unfolded conformation of **19** than does the weaker network hydrogen bonds in H_2O). In contrast, we hypothesize that the ΔC_p of **19p** is independent of D_2O concentration because PEG disrupts the strong network of deuterium bonds surrounding the protein, thereby attenuating the penalty for unfolding in D_2O.

Structure-Based Selection of Stabilizing PEGylation Sites. On the basis of these mechanistic insights, we wondered whether (1) side chain orientation and (2) the presence of nearby OH groups could be used as structure-based criteria to identify positions most likely to experience substantial entropic PEG-based stabilization. To this end, we analyzed each of the PEGylation sites discussed above in the X-ray crystal structure of the parent WW domain from which **16p**, **17p**, **18p**, **19p**, **26p**, **27p**, **28p**, **29p**, **30p**, **32p** and their non-PEGylated counterparts were derived. We limited this analysis to these variants because their CD spectra indicate close structural similarity to the parent WW domain.

At each PEGylation site, we defined vectors **a** and **b** (Figure 6): vector **a** begins with the backbone α carbon and ends at the side-chain center-of-mass (determined by averaging the *x*,*y*,*z*-coordinates of each side-chain atom); vector **b** begins with the side-chain center-of-mass and ends at side-chain oxygen atom of the nearest Ser, Thr, or Tyr residue. We then measured the angle θ between vectors **a** and **b** at each position using the following relationship: $\cos\theta = \mathbf{a}\cdot\mathbf{b}/(|\mathbf{a}|\cdot|\mathbf{b}|)$. Small values of θ indicate that a side chain is oriented toward the nearest Ser, Thr, or Tyr residue, whereas large values of θ indicate orientation away from the nearest Ser, Thr, or Tyr residue. Values of θ for all the positions investigated are shown in Table 4.

Next, we examined the relationship between the angle θ and the PEG-based stabilization ($\Delta\Delta G_f$) of **16p**, **17p**, **18p**, **19p**, **26p**, **27p**, **28p**, **29p**, **30p**, and **32p**, relative to their non-PEGylated counterparts. Figure 7 indicates that PEGylation

Table 4. Angle θ at Various PEGylation Sites within WW[a]

PEGylation site	native residue	$\Delta\Delta G_f$ (kcal/mol)	θ (deg)
16	Ser	−0.90 ± 0.03	95
17	Arg	−0.18 ± 0.05	121
18	Ser	0.00 ± 0.07	145
19	Ser	−0.70 ± 0.04	31
26	Asn	−0.58 ± 0.06	83
27	His	0.38 ± 0.04	155
28	Ile	0.00 ± 0.07	128
29	Thr	−0.36 ± 0.04	98
30	Asn	0.00 ± 0.07	150
32	Ser	−0.45 ± 0.02	47

[a] $\Delta\Delta G_f$ values associated with PEGylation at each position are from Table 1.

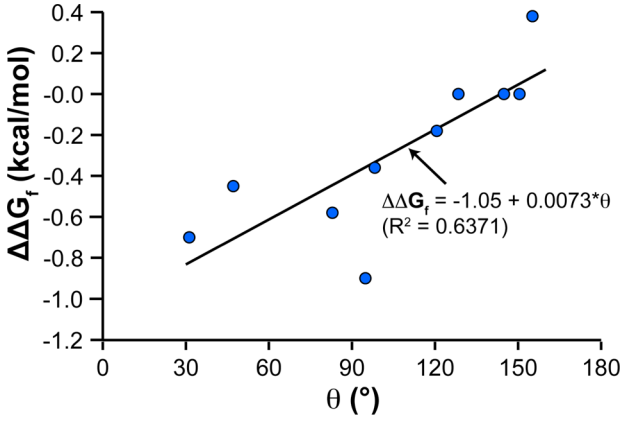


Figure 7. Relationship between the angle θ and PEG-based stabilization at a given site ($\Delta\Delta G_f$).

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

tends to be most stabilizing at positions with smaller values of θ (i.e., at positions that are oriented toward nearby Ser, Thr, or Tyr side chains). Most importantly, had we used the correlation line shown in Figure 7 prospectively, we would have correctly predicted that PEGylation at positions 16, 19, 26, 29, and 32 would result in substantial (<−0.30 kcal mol^{-1}) increases to conformational stability.

We tested the utility of the angle θ as a predictor of PEG-based stabilization within other proteins by calculating θ for each residue within the chicken Src SH3 domain (hereafter called SH3, Figure 8A).[120] Thr20 and Thr22 are within the same loop near the N-terminus of SH3 and are only 3.7 Å apart from each other. More importantly, Thr20 is oriented toward Thr22, with θ = 77° (Figure 8A). The correlation observed between θ and PEG-based stabilization in the context of WW (Figure 7) led us to predict that PEGylation of an Asn residue at position 20 of SH3 would enhance conformational stability.

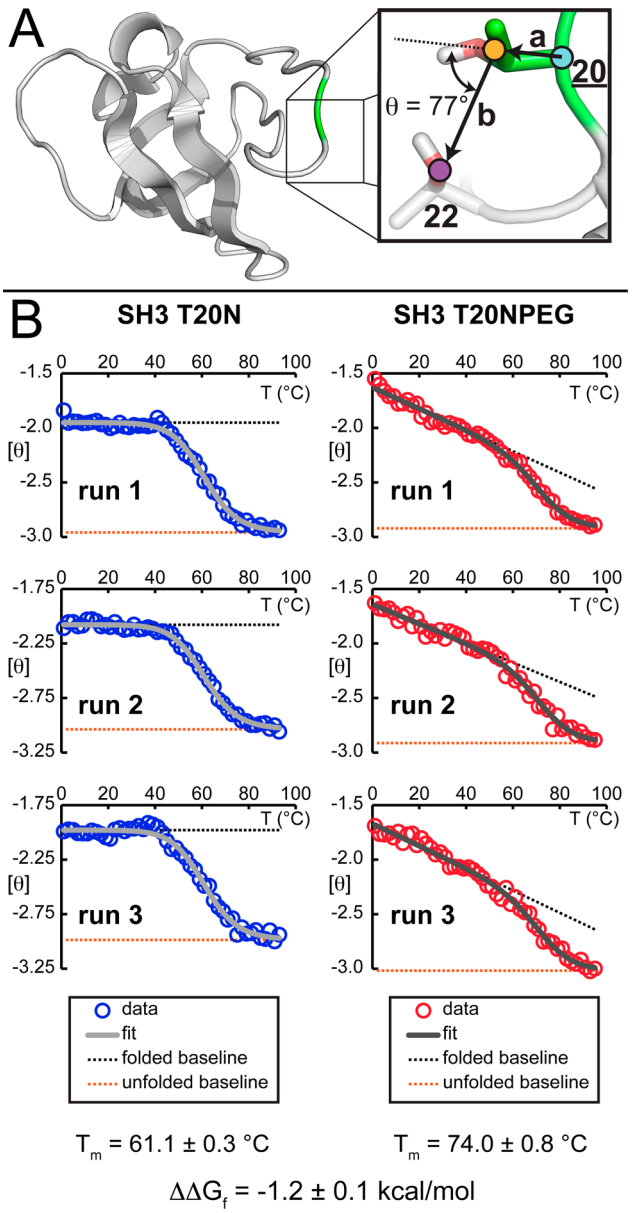


Figure 8. (A) Ribbon diagram of chicken Src SH3 (PDB ID: 1SRL), with Thr20 highlighted in green. The inset shows the angle θ between vectors **a** and **b** at position 20. Alpha carbons (Cα), side-chain centers-of-mass (COM), and oxygens of the nearest OH-containing side chain are highlighted with blue-, orange-, and purple-filled circles, respectively. (B) Variable-temperature CD data for SH3 T20N and SH3 T20NPEG at 50 μM in 20 mM sodium phosphate buffer (pH 7) run in triplicate for each variant.

To test this hypothesis, we used solid-phase peptide synthesis to prepare SH3 T20N and SH3 T20NPEG, in which Thr20 has been replaced by Asn or AsnPEG$_4$, respectively. We assessed the conformational stability of these variants using variable-temperature CD experiments, run in triplicate for each variant (Figure 8B). Data from these experiments are readily fit to equations derived from a two-state folding−unfolding model (see the Supporting Information). The melting temperature T_m of SH3 T20NPEG (74.0 ± 0.8 °C) is 13.0 ± 0.9 °C higher than that of non-PEGylated SH3 T20N (T_m = 61.1 ± 0.3 °C), corresponding to an increase in stability of −1.2 ± 0.1 kcal mol^{-1} at 61.1 °C. Atomistic simulations suggest that PEGylation of SH3 T20NPEG is associated with lower water density (and higher water disorder) around nearby polar side chains, including Thr22, suggesting a similar origin for PEG-based stabilization in SH3 and in WW (see Supporting Information). It is remarkable that a 4-unit PEG oligomer can have such a large effect on the stability of the 56-residue SH3 domain. This substantial increase in SH3 conformational stability is consistent with the predictions derived from our studies on WW, suggesting that the angle θ (i.e., the orientation of a side chain with respect to nearby OH groups) is a reasonable predictor of PEG-based stabilization in β-sheet- and reverse-turn-containing proteins.

Correlated Conformational and Proteolytic Stabilization. Finally, we wondered whether the PEG-based stabilization observed at selected positions in WW would be associated with enhanced protection from proteolytic degradation. To address this question, we assessed the resistance of the WW variants described above to degradation by pronase and by proteinase K.[121] We previously showed that PEGylation at stabilizing position 19 protected WW from proteolysis.[108] Recall that PEGylation at position 19 is stabilizing ($\Delta\Delta G_f$ = −0.70 ± 0.04 kcal mol^{-1}). The half-life of PEGylated **19p** in pronase is 3.6 ± 0.3 times longer than that of non-PEGylated **19**; we observed a similar effect in proteinase K (Table 5). At neutral position 18 (where PEGylation does not change conformational stability), the same four-unit PEG oligomer provides much less protection against proteolysis: the half-lives of **18p** in pronase and proteinase K are only 1.23 ± 0.06 and 1.7 ± 0.3 longer, respectively, than those of non-PEGylated **18** (Table 5).

Proteolysis experiments with the other PEGylated WW variants and their non-PEGylated counterparts described above follow a similar trend (see Table 5), with increased resistance to proteolysis observed in cases where PEG is most strongly stabilizing. For example, the half-lives of doubly PEGylated **19p**/**26p** in pronase and proteinase K are 5.9 ± 0.9 and 6.9 ± 1.9 times longer, respectively, than those of non-PEGylated **19**/**26**. This trend is illustrated in Figure 9, which shows the ratio of the half-lives of each matched pair of PEGylated vs non-PEGylated WW variants plotted against the conformational stability of each PEGylated variant relative to its non-PEGylated counterpart. The plot suggests that beyond a certain basal level, the proteolytic protection imparted by the four-unit oligomer is substantially enhanced at positions where PEG also increases WW conformational stability. These data

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

Table 5. Impact of PEGylation with PEG4 at Various Sites on Resistance of WW Variants to Proteolysis[a]

site	$\Delta\Delta G_f$ (kcal/mol)	pronase $t_{1/2}$ ratio	proteinase K $t_{1/2}$ ratio
16	−0.90 ± 0.03	2.0 ± 0.2	2.3 ± 0.3
17	−0.18 ± 0.05	1.19 ± 0.07	
18	0.00 ± 0.07	1.23 ± 0.06	1.7 ± 0.3
19	−0.70 ± 0.04	3.6 ± 0.3	3.4 ± 0.4
26	−0.58 ± 0.06	1.7 ± 0.2	2.7 ± 0.3
27	0.38 ± 0.04	0.86 ± 0.05	0.9 ± 0.1
28	0.00 ± 0.07	1.1 ± 0.1	1.0 ± 0.1
29	−0.36 ± 0.04	1.7 ± 0.2	2.0 ± 0.3
16/26	−1.38 ± 0.01	2.3 ± 0.2	4.1 ± 0.5
19/26	−1.26 ± 0.02	5.9 ± 0.9	6.9 ± 1.9
16/19	−0.80 ± 0.01	3.1 ± 0.3	
26/29	−0.56 ± 0.02	3.2 ± 0.4	

[a]Tabulated data are given ± standard error. Values of $\Delta\Delta G_f$ are presented as given in Table 1. The $t_{1/2}$ ratio for a given site in pronase or proteinase K is the ratio of the half-life of the PEGylated WW variant to the half-life of the corresponding non-PEGylated WW variant in the indicated protease. Proteolysis experiments were performed at 50 μM protein concentration in 20 mM sodium phosphate buffer (pH 7).

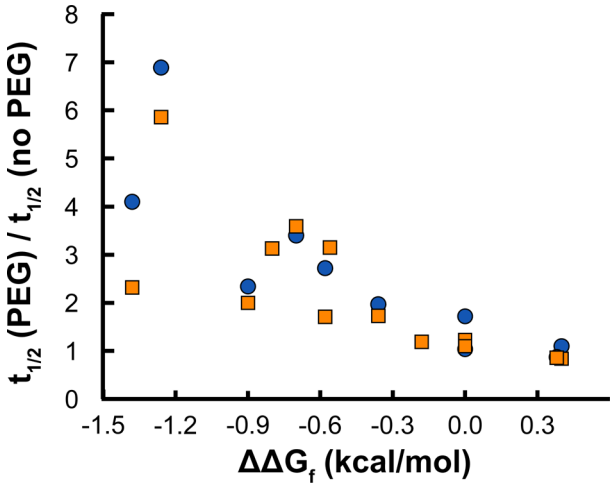


Figure 9. Plot of PEG-based proteolytic stability (expressed as the ratio of half-life of a given PEGylated WW variant to the half-life of its sequence-matched non-PEGylated counterpart) in the presence of pronase (blue circles) or proteinase K (orange squares) vs PEG-based conformational stability ($\Delta\Delta G_f$).

provide support for our hypothesis that globally optimal PEGylation sites are characterized by the ability of the PEG oligomer to increase protein conformational stability.

However, one could argue that this observed dependence of proteolytic resistance on PEG-based conformational stabilization (Table 5, Figure 9) is a result of local effects that are important for four-unit PEGs, but which are insignificant for longer PEGs. Indeed, one might expect steric hindrance to be the dominant contributor to the proteolytic resistance associated with longer PEGs, independent of conformational stabilization. To test this hypothesis, we assessed the ability of a 45-unit PEG to protect WW variants **16p45**, **19p45**, **26p45**, **28p45**, and **29p45** from proteolysis (Table 6). This 45-unit PEG is clearly much shorter than the 20−40 kDa PEGs typically found in PEGylated protein drugs. However, WW is a small protein (~4 kDa); the 45-unit PEG (~2 kDa) comprises ~33% of the total masses of these PEGylated WW variants, a

Table 6. Impact of PEGylation with PEG45 at Various Sites on Resistance of WW Variants to Proteolysis[a]

proteins	$\Delta\Delta G_f$ (kcal/mol)	pronase $t_{1/2}$ ratio
16p45 vs **16**	−0.39 ± 0.03	2.3 ± 0.5
19p45 vs **19**	−0.67 ± 0.05	5.6 ± 1.2
26p45 vs WW	−0.27 ± 0.06	2.8 ± 0.5
28p45 vs **28**	0.36 ± 0.05	1.1 ± 0.3
29p45 vs **29**	−0.59 ± 0.08	2.8 ± 0.5

[a]Tabulated data are given ± standard error. Values of $\Delta\Delta G_f$ are presented as given in Table 2. The $t_{1/2}$ ratio for a given site in pronase is the ratio of the half-life of the PEGylated WW variant to the half-life of the corresponding non-PEGylated WW variant. Proteolysis experiments were performed at 50 μM protein concentration in 20 mM sodium phosphate buffer (pH 7).

PEG/protein composition approaching that of many PEGylated protein drugs (pegfilgrastim, for example, is 50% PEG: a ~20 kDa PEG attached to a ~20 kDa protein).

If steric hindrance were the only significant contributor to the proteolytic resistance associated with longer PEGs, one would expect PEG-based changes in protein conformational stabilization to matter less and less with increasing PEG/protein ratios. For example, one would expect proteolytic resistance in the 33:67 PEG/protein conjugates (e.g., the 45-unit PEG WW variants) to be less dependent on conformational stability than in the 5:95 PEG/protein conjugates (e.g., the 4-unit PEG WW variants). Instead, we find that a PEG/protein ratio of 33:67, PEG-based increases to proteolytic resistance remain strongly correlated with the impact of the 45-unit PEG on conformational stability. At "stabilizing" positions 16, 19, 26, and 29, the increases in conformational stability associated with the 45-unit PEG oligomer are accompanied by 2.3-, 5.6-, 2.8-, and 2.8-fold increases in half-life, respectively, in the presence of pronase. However, at "destabilizing" position 28 ($\Delta\Delta G_f$ = 0.36 ± 0.05 kcal mol^{-1}), the 45-unit PEG oligomer has no substantial impact on proteolytic stability. These observations are not consistent with the hypothesis that steric hindrance is the only significant factor contributing to PEG-based proteolytic resistance and suggest that PEG-based changes to conformational stability also play an important role for PEG/protein conjugates approaching the compositions typical of PEGylated protein drugs.

Conclusion. Advances in protein chemistry now allow site-specific PEGylation of any arbitrary position on the protein surface. Why pursue predictive tools for identifying optimal PEGylation sites when one can simply scan a PEGylated side chain through a list of potential sites and pick the one(s) that provide the best balance between enhanced pharmacokinetic properties and biological function?[31,122] Such a trial-and-error approach is unsatisfying from a scientific point of view, is both time- and resource-intensive (site-specific side-chain modification is much more challenging to carry out than alanine-scanning mutagenesis, for example) and may therefore be limited by practical considerations to a subset of potential surface sites, and must be repeated for each new protein of interest. In contrast, rational structure-based guidelines for identifying optimal PEGylation sites have the potential to circumvent this time-consuming step in PEGylated protein drug development.

We have developed a structure-based method for predicting which sites within the WW domain are most likely to experience PEG-based stabilization and have shown that

dx.doi.org/10.1021/ja5095183 | J. Am. Chem. Soc. 2014, 136, 17547−17560

PEG-based stabilization is associated with enhanced resistance to proteolysis. We developed this method on the basis of mutagenesis experiments, which showed that side-chain orientation and the presence of nearby OH groups can modulate PEG-based stabilization at a given site. MD simulations suggest that stabilization cannot always be explained by favorable PEG−protein interactions because the formation of a tight PEG−protein interface is coupled by an entropic loss in many cases. While direct PEG−OH interactions cannot explain the increased thermodynamic stability, it is likely that nearby OH groups may instead exert a more indirect influence, involving the network of hydrogen-bound solvent molecules surrounding the protein. The simulations indicate that PEG can increase the disorder of water molecules around nearby residues. Solvent isotope experiments are consistent with this possibility, as are our observations that PEG-based stabilization is entropic in origin, with beneficial increases in entropy compensating for unfavorable increases in enthalpy.

We find that 45- and 4-unit PEGs have a similar impact on WW conformational and proteolytic stability, suggesting that the structure-based model developed using the 4-unit PEG will apply in the context of the larger oligomers typically used in therapeutically relevant proteins. Most importantly, we have also shown that our structure-based method can correctly predict a location within the Src SH3 domain (another β-sheet protein) where PEGylation enhances conformational stability. We look forward to applying this method to larger therapeutically relevant proteins.

■ ASSOCIATED CONTENT

Supporting Information

Complete experimental methods, compound characterization, variable temperature circular dichroism data, temperature jump kinetic data, and proteolysis assay data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

koby.levy@weizmann.ac.il

jlprice@chem.byu.edu

Author Contributions

[‖]P.B.L. and Y.G. contributed equally to the manuscript, which was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

J.L.P. acknowledges the support of start-up funds from the Department of Chemistry and Biochemistry at Brigham Young University. Y.L. acknowledges The Morton and Gladys Pickman Professional Chair in Structural Biology. When this work was done, M.B.P. was a Howard Hughes Medical Institute International Student Research Fellow.

■ REFERENCES

(1) Abuchowski, A.; Vanes, T.; Palczuk, N. C.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3578−3581.

(2) Abuchowski, A.; Mccoy, J. R.; Palczuk, N. C.; Vanes, T.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3582−3586.

(3) Harris, J. M.; Chess, R. B. *Nat. Rev. Drug Discovery* **2003**, *2*, 214−221.

(4) Frokjaer, S.; Otzen, D. E. *Nat. Rev. Drug Discovery* **2005**, *4*, 298−306.

(5) Fishburn, C. S. *J. Pharm. Sci.* **2008**, *97*, 4167−4183.

(6) Veronese, F. M.; Mero, A. *Biodrugs* **2008**, *22*, 315−329.

(7) *PEGylated Protein Drugs: Basic Science and Clinical Applications*; Veronese, F. M., Ed.; Birkhauser Verlag: Basel, 2009.

(8) Jevsevar, S.; Kunstelj, M.; Porekar, V. G. *Biotechnol. J.* **2010**, *5*, 113−128.

(9) Alconcel, S. N. S.; Baas, A. S.; Maynard, H. D. *Polym. Chem.* **2011**, *2*, 1442−1448.

(10) Molineux, G. *Curr. Pharm. Des.* **2004**, *10*, 1235−1244.

(11) Rajender Reddy, K.; Modi, M. W.; Pedder, S. *Adv. Drug Delivery Rev.* **2002**, *54*, 571−586.

(12) Grace, M.; Youngster, S.; Gitlin, G.; Sydor, W.; Xie, L.; Westreich, L.; Jacobs, S.; Brassard, D.; Bausch, J.; Bordens, R. *J. Interferon Cytokine Res.* **2001**, *21*, 1103−1115.

(13) Wang, Y.-S.; Youngster, S.; Grace, M.; Bausch, J.; Bordens, R.; Wyss, D. F. *Adv. Drug Delivery Rev.* **2002**, *54*, 547−570.

(14) Hershfield, M. S.; Buckley, R. H.; Greenberg, M. L.; Melton, A. L.; Schiff, R.; Hatem, C.; Kurtzberg, J.; Markert, M. L.; Kobayashi, R. H.; Kobayashi, A. L.; Abuchowski, A. *N. Engl. J. Med.* **1987**, *316*, 589−596.

(15) Kopchick, J. J.; Parkinson, C.; Stevens, E. C.; Trainer, P. J. *Endocr. Rev.* **2002**, *23*, 623−646.

(16) Graham, M. L. *Adv. Drug Delivery Rev.* **2003**, *55*, 1293−1302.

(17) Muller, A. F.; Kopchick, J. J.; Flyvbjerg, A.; van der Lely, A. J. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 1503−1511.

(18) Topf, J. M. *Expert. Opin. Pharmacother.* **2008**, *9*, 839−849.

(19) Ganson, N.; Kelly, S.; Scarlett, E.; Sundy, J.; Hershfield, M. *Arthritis Res. Ther.* **2006**, *8*, R12.

(20) Bourne, T.; Fossati, G.; Nesbitt, A. *Biodrugs* **2008**, *22*, 331−337.

(21) Park, Y. K.; Abuchowski, A.; Davis, S.; Davis, F. *Anticancer Res.* **1981**, *1*, 373−376.

(22) Huang, Z. F.; Ye, C. H.; Liu, Z. J.; Wang, X. J.; Chen, H. B.; Liu, Y. L.; Tang, L.; Zhao, H. X.; Wang, J. F.; Feng, W. K.; Li, X. K. *Bioconjugate Chem.* **2012**, *23*, 740−750.

(23) Wu, L.; Ho, S. V.; Wang, W.; Gao, J. P.; Zhang, G. F.; Su, Z. G.; Hu, T. *Int. J. Pharm.* **2013**, *453*, 533−540.

(24) Resch, G.; Moreillon, P.; Fischetti, V. *AMB Express* **2011**, *1*, 29.

(25) Kunstelj, M.; Fidler, K.; Škrajnar, Š.; Kenig, M.; Smilović, V.; Kusterle, M.; Caserman, S.; Zore, I.; Porekar, V. G.; Jevševar, S. *Bioconjugate Chem.* **2013**, *24*, 889−896.

(26) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.

(27) Rosendahl, M. S.; Doherty, D. H.; Smith, D. J.; Carlson, S. J.; Chlipala, E. A.; Cox, G. N. *Bioconjugate Chem.* **2005**, *16*, 200−207.

(28) Dirksen, A.; Dawson, P. E. *Curr. Opin. Chem. Biol.* **2008**, *12*, 760−766.

(29) Brocchini, S.; Godwin, A.; Balan, S.; Choi, J.-w.; Zloh, M.; Shaunak, S. *Adv. Drug Delivery Rev.* **2008**, *60*, 3−12.

(30) Van de Vijver, P.; Suylen, D.; Dirksen, A.; Dawson, P. E.; Hackeng, T. M. *Biopolymers* **2010**, *94*, 465−474.

(31) Cho, H.; Daniel, T.; Buechler, Y. J.; Litzinger, D. C.; Maio, Z.; Putnam, A. M.; Kraynov, V. S.; Sim, B. C.; Bussell, S.; Javahishvili, T.; Kaphle, S.; Viramontes, G.; Ong, M.; Chu, S.; Becky, G. C.; Lieu, R.; Knudsen, N.; Castiglioni, P.; Norman, T. C.; Axelrod, D. W.; Hoffman, A. R.; Schultz, P. G.; DiMarchi, R. D.; Kimmel, B. E. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9060−9065.

(32) Tada, S.; Andou, T.; Suzuki, T.; Dohmae, N.; Kobatake, E.; Ito, Y. *PLoS One* **2012**, *7*, e49235.

(33) Levine, P. M.; Craven, T. W.; Bonneau, R.; Kirshenbaum, K. *Org. Lett.* **2014**, *16*, 512−515.

(34) Frokjaer, S.; Otzen, D. E. *Nat. Rev. Drug Discovery* **2005**, *4*, 298−306.

(35) Dobson, C. M. *Nature* **2003**, *426*, 884−890.

(36) Clark, P. L. *Trends Biochem. Sci.* **2004**, *29*, 527−534.

(37) Wang, W.; Roberts, C. J. *Aggregation of Therapeutic Proteins*; Wiley: Hoboken, NJ, 2010.

(38) Daniel, R. M.; Cowan, D. A.; Morgan, H. W.; Curran, M. P. *Biochem. J.* **1982**, *207*, 641−644.
(39) Parsell, D. A.; Sauer, R. T. *J. Biol. Chem.* **1989**, *264*, 7590−7595.
(40) Klink, T. A.; Raines, R. T. *J. Biol. Chem.* **2000**, *275*, 17463−17467.
(41) Ahmad, S.; Kumar, V.; Ramanand, K. B.; Rao, N. M. *Protein Sci.* **2012**, *21*, 433−446.
(42) Imoto, T.; Yamada, H.; Ueda, T. *J. Mol. Biol.* **1986**, *190*, 647−649.
(43) Cline, L. L.; Waters, M. L. *Biopolymers* **2009**, *92*, 502−507.
(44) Moore, W. V.; Leppert, P. *J. Clin. Endocrinol. Metab.* **1980**, *51*, 691−697.
(45) Robbins, D. C.; Cooper, S. M.; Fineberg, S. E.; Mead, P. M. *Diabetes* **1987**, *36*, 838−841.
(46) Schernthaner, G. *Diabetes Care* **1993**, *16*, 155−165.
(47) Braun, A.; Kwee, L.; Labow, M. A.; Alsenz, J. *Pharm. Res.* **1997**, *14*, 1472−1478.
(48) Hermeling, S.; Crommelin, D.; Schellekens, H.; Jiskoot, W. *Pharm. Res.* **2004**, *21*, 897−903.
(49) Ohkuri, T.; Nagatomo, S.; Oda, K.; So, T.; Imoto, T.; Ueda, T. *J. Immunol.* **2010**, *185*, 4199−4205.
(50) Thai, R.; Moine, G.; Desmadril, M.; Servent, D.; Tarride, J.-L.; Menez, A.; Leonetti, M. *J. Biol. Chem.* **2004**, *279*, 50257−50266.
(51) Shu, J. Y.; Lund, R.; Xu, T. *Biomacromolecules* **2012**, *13*, 1945−1955.
(52) Shu, J. Y.; Tan, C.; DeGrado, W. F.; Xu, T. *Biomacromolecules* **2008**, *9*, 2111−2117.
(53) Jain, A.; Ashbaugh, H. S. *Biomacromolecules* **2011**, *12*, 2729−2734.
(54) Baillargeon, M.; Sonnet, P. *J. Am. Oil Chem. Soc.* **1988**, *65*, 1812−1815.
(55) Basri, M.; Ampon, K.; Yunus, W. M. Z. W.; Razak, C. N. A.; Salleh, A. B. *J. Chem. Technol. Biotechnol.* **1995**, *64*, 10−16.
(56) Longo, M. A.; Combes, D. *J. Chem. Technol. Biotechnol.* **1999**, *74*, 25−32.
(57) Hernaiz, M. J.; Sanchez-Montero, J. M.; Sinisterra, J. V. *Enzyme Microb. Technol.* **1999**, *24*, 181−190.
(58) Gaertner, H. F.; Puigserver, A. J. *Enzyme Microb. Technol.* **1992**, *14*, 150−155.
(59) Monfardini, C.; Schiavon, O.; Caliceti, P.; Morpurgo, M.; Harris, J. M.; Veronese, F. M. *Bioconjugate Chem.* **1995**, *6*, 62−69.
(60) Zhang, Z.; He, Z.; Guan, G. *Biotechnol. Technol.* **1999**, *13*, 781−786.
(61) Treetharnmathurot, B.; Ovartlarnporn, C.; Wungsintaweekul, J.; Duncan, R.; Wiwattanapatapee, R. *Int. J. Pharm.* **2008**, *357*, 252−259.
(62) Chiu, K.; Agoubi, L. L.; Lee, I.; Limpar, M. T.; Lowe, J. W.; Goh, S. L. *Biomacromolecules* **2010**, *11*, 3688−3692.
(63) Topchieva, I. N.; Efremova, N. V.; Khvorov, N. V.; Magretova, N. N. *Bioconjugate Chem.* **1995**, *6*, 380−388.
(64) Castellanos, I. J.; Al-Azzam, W.; Griebenow, K. *J. Pharm. Sci.* **2005**, *94*, 327−340.
(65) Rodriguez-Martinez, J. A.; Solá, R. J.; Castillo, B.; Cintron-Colon, H. R.; Rivera-Rivera, I.; Barletta, G.; Griebenow, K. *Biotechnol. Bioeng.* **2008**, *101*, 1142−1149.
(66) Rodríguez-Martínez, J.; Rivera-Rivera, I.; Solá, R.; Griebenow, K. *Biotechnol. Lett.* **2009**, *31*, 883−887.
(67) Garcia, D.; Ortéga, F.; Marty, J.-L. *Biotechnol. Appl. Biochem.* **1998**, *27*, 49−54.
(68) López-Cruz, J. I.; Viniegra-Gonzalez, G.; Hernandez-Arana, A. *Bioconjugate Chem.* **2006**, *17*, 1093−1098.
(69) Nie, Y.; Zhang, X.; Wang, X.; Chen, J. *Bioconjugate Chem.* **2006**, *17*, 995−999.
(70) Hinds, K. D.; Kim, S. W. *Adv. Drug Delivery Rev.* **2002**, *54*, 505−530.
(71) Yang, C.; Lu, D. N.; Liu, Z. *Biochemistry* **2011**, *50*, 2585−2593.
(72) Meng, W.; Guo, X.; Qin, M.; Pan, H.; Cao, Y.; Wang, W. *Langmuir* **2012**, *28*, 16133−16140.
(73) Shu, J. Y.; Tan, C.; DeGrado, W. F.; Xu, T. *Biomacromolecules* **2008**, *9*, 2111−2117.
(74) Dhalluin, C.; Ross, A.; Leuthold, L. A.; Foser, S.; Gsell, B.; Muller, F.; Senn, H. *Bioconjugate Chem.* **2005**, *16*, 504−517.
(75) Basu, A.; Yang, K.; Wang, M.; Liu, S.; Chintala, R.; Palm, T.; Zhao, H.; Peng, P.; Wu, D.; Zhang, Z.; Hua, J.; Hsieh, M.-C.; Zhou, J.; Petti, G.; Li, X.; Janjua, A.; Mendez, M.; Liu, J.; Longley, C.; Zhang, Z.; Mehlig, M.; Borowski, V.; Viswanathan, M.; Filpula, D. *Bioconjugate Chem.* **2006**, *17*, 618−630.
(76) Ramon, J.; Saez, V.; Baez, R.; Aldana, R.; Hardy, E. *Pharm. Res.* **2005**, *22*, 1374−1386.
(77) Garcia-Arellano, H.; Valderrama, B.; Saab-Rincon, G.; Vazquez-Duhalt, R. *Bioconjugate Chem.* **2002**, *13*, 1336−1344.
(78) Plesner, B.; Fee, C. J.; Westh, P.; Nielsen, A. D. *Eur. J. Pharm. Biopharm.* **2011**, *79*, 399−405.
(79) Yang, Z.; Williams, D.; Russell, A. J. *Biotechnol. Bioeng.* **1995**, *45*, 10−17.
(80) Yang, Z.; Domach, M.; Auger, R.; Yang, F. X.; Russell, A. J. *Enzyme Microb. Technol.* **1996**, *18*, 82−89.
(81) Callahan, W. J.; Narhi, L. O.; Kosky, A. A.; Treuheit, M. J. *Pharm. Res.* **2001**, *18*, 261−266.
(82) Plesner, B.; Westh, P.; Nielsen, A. D. *Int. J. Pharm.* **2011**, *406*, 62−68.
(83) Rodríguez-Martínez, J. A.; Rivera-Rivera, I.; Griebenow, K. *J. Pharm. Pharmacol.* **2011**, *63*, 800−805.
(84) Koepf, E. K.; Petrassi, H. M.; Ratnaswamy, G.; Huff, M. E.; Sudol, M.; Kelly, J. W. *Biochemistry* **1999**, *38*, 14338−14351.
(85) Koepf, E. K.; Petrassi, H. M.; Sudol, M.; Kelly, J. W. *Protein Sci.* **1999**, *8*, 841−853.
(86) Jäger, M.; Nguyen, H.; Crane, J. C.; Kelly, J. W.; Gruebele, M. *J. Mol. Biol.* **2001**, *311*, 373−393.
(87) Kaul, R.; Angeles, A. R.; Jäger, M.; Powers, E. T.; Kelly, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 5206−5212.
(88) Deechongkit, S.; Kelly, J. W. *J. Am. Chem. Soc.* **2002**, *124*, 4980−4986.
(89) Kaul, R.; Deechongkit, S.; Kelly, J. W. *J. Am. Chem. Soc.* **2002**, *124*, 11900−11907.
(90) Kowalski, J. A.; Kiu, K.; Kelly, J. W. *Biopolymers* **2002**, *63*, 111−121.
(91) Nguyen, H.; Jäger, M.; Moretto, A.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3948−3953.
(92) Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. *Nature* **2004**, *430*, 101−105.
(93) Nguyen, H.; Jäger, M.; Kelly, J. W.; Gruebele, M. *J. Phys. Chem. B* **2005**, *109*, 15182−15186.
(94) Sekijima, Y.; Wiseman, R. L.; Matteson, J.; Hammarström, P.; Miller, S. R.; Sawkar, A. R.; Balch, W. E.; Kelly, J. W. *Cell* **2005**, *121*, 73−85.
(95) Jäger, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 10648−10653.
(96) Jäger, M.; Dendle, M.; Fuller, A. A.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 2306−2313.
(97) Jäger, M.; Nguyen, H.; Dendle, M.; Gruebele, M.; Kelly, J. W. *Protein Sci.* **2007**, *16*, 1495−1501.
(98) Jäger, M.; Deechongkit, S.; Koepf, E. K.; Nguyen, H.; Gao, J.; Powers, E. T.; Gruebele, M.; Kelly, J. W. *Biopolymers* **2008**, *90*, 751−758.
(99) Liu, F.; Du, D.; Fuller, A. A.; Davoren, J. E.; Wipf, P.; Kelly, J. W.; Gruebele, M. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2369−2374.
(100) Fuller, A. A.; Du, D.; Liu, F.; Davoren, J. E.; Bhabha, G.; Kroon, G.; Case, D. A.; Dyson, H. J.; Powers, E. T.; Wipf, P.; Gruebele, M.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11067−11072.
(101) Gao, J.; Bosco, D. A.; Powers, E. T.; Kelly, J. W. *Nat. Struct. Mol. Biol.* **2009**, *16*, 684−690.
(102) Jäger, M.; Dendle, M.; Kelly, J. W. *Protein Sci.* **2009**, *18*, 1806−1813.
(103) Price, J. L.; Shental-Bechor, D.; Dhar, A.; Turner, M. J.; Powers, E. T.; Gruebele, M.; Levy, Y.; Kelly, J. W. *J. Am. Chem. Soc.* **2010**, *132*, 15359−15367.

(104) Culyba, E. K.; Price, J. L.; Hanson, S. R.; Dhar, A.; Wong, C. H.; Gruebele, M.; Powers, E. T.; Kelly, J. W. *Science* **2011**, *331*, 571−575.
(105) Price, J. L.; Powers, D. L.; Powers, E. T.; Kelly, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 14127−14132.
(106) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. *Cell* **1997**, *89*, 875−886.
(107) Price, J. L.; Powers, E. T.; Kelly, J. W. *ACS Chem. Biol.* **2011**, *6*, 1188−1192.
(108) 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. *Bioconjugate Chem.* **2013**, *24*, 796−802.
(109) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D* **2010**, *66*, 486−501.
(110) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(111) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950−1958.
(112) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. *Proteins* **2009**, *75*, 430−441.
(113) Shental-Bechor, D.; Levy, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8256−8261.
(114) Shental-Bechor, D.; Levy, Y. *Curr. Opin. Struct. Biol.* **2009**, *19*, 524−533.
(115) Hagai, T.; Levy, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2001−2006.
(116) Shental-Bechor, D.; Smith, M. T. J.; MacKenzie, D.; Broom, A.; Marcovitz, A.; Ghashut, F.; Go, C.; Bralha, F.; Meiering, E. M.; Levy, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17839−17844.
(117) Chao, S.-H.; Matthews, S. S.; Paxman, R.; Aksimentiev, A.; Gruebele, M.; Price, J. L. *J. Phys. Chem. B* **2014**, *118*, 8388−8895.
(118) Cioni, P.; Strambini, G. B. *Biophys. J.* **2002**, *82*, 3246−3253.
(119) Prabhu, N. V.; Sharp, K. A. *Annu. Rev. Phys. Chem.* **2004**, *56*, 521−548.
(120) Yu, H.; Rosen, M. K.; Schreiber, S. L. *FEBS Lett.* **1993**, *324*, 87−92.
(121) Horne, W. S.; Boersma, M. D.; Windsor, M. A.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2008**, *47*, 2853−2856.
(122) Zhang, C.; Yang, X.-l.; Yuan, Y.-h.; Pu, J.; Liao, F. *Biodrugs* **2012**, *26*, 209−215.