Published in final edited form as:

Biochemistry. 2017 October 24; 56(42): 5671–5678. doi:10.1021/acs.biochem.7b00215.

The Surface of Protein λ_{6-85} Can Act as a Template for Recurring PEG Structure

Shu-Han Chao[†], Jan Schäfer^{†,§}, and Martin Gruebele^{†,¶,^,*}

[†]Department of Physics, University of Illinois, Urbana, Illinois 61801, 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

Abstract

PEGylated proteins play an increasingly important role in pharmaceutical drug delivery. We recently showed that short PEG chains can affect protein structure, even when they are not making extensive contact with the protein surface. In contrast, PEG is generally assumed to form a relatively unstructured coil, whose compactness depends on solvent conditions. Here we test whether a host protein could allow PEG to form recurrent structural motifs while the PEG chain is in contact with the protein surface. We link a PEG oligomer ($n=45$) to one of two nearly opposite locations on the small alpha-helical protein λ_{6-85} to investigate this question. We first demonstrate experimentally that in these particular positions, PEG does not significantly affect the thermodynamic stability and folding kinetics of λ_{6-85} . We then use several all-atom MD simulations of 1 μ s duration to show how PEG equilibrates between states extending into the solvent, and states packed onto the protein surface. The packing reveals recurring structures, including persistent hydrogen bond and hydrophobic contact patterns that appear multiple times. Some interactions of PEG with surface lysines are best described as an ‘intermittent slithering’ motion of the PEG around the side chain, as seen in short MD movies. Thus, PEG achieves a variety of metastable organized structures on the protein surface, somewhere between a random globule and true folding. We also investigated the PEG-protein interaction in the unfolded state of the protein. We find that PEG has a propensity for stabilizing certain helices of λ_{6-85} , no matter which of the two positions it was attached to. Thus, sufficiently long PEG chains are organized by the protein surface, and in turn interact with certain elements of protein structure more than others, even when PEG is attached to very different sites.

Keywords

PEGylation; protein; thermal melt; molecular dynamics simulation

*Corresponding author: Tel: 001-217-333-6136, mgruebel@illinois.edu.

§Present address: (J. S.) Max Planck Institute for Polymer Research, Mainz, Germany

The Supporting Information is available free of charge on the ACS publications website at <http://pubs.acs.org/XXXXX>.

PDF file describes additional measurements with text and figures. The SI movies contain movies of the molecular dynamics simulations discussed in the text.

Poly(ethylene glycol) (PEG) has been widely used as a ‘stealth polymer’ in the pharmaceutical industry since the 1990s.¹ By PEG attachment, called PEGylation, one or multiple PEG oligomers are covalently coupled with a protein drug to form a PEG-protein conjugate. The molar mass of attached PEG ranges from 400 Da to about 50 kDa, depending on the drug type and its medical application.² PEGylation not only shields protein drugs from the immune system, but it also enhances activity, increases solubility, and prevents aggregation.³ The benefits of PEGylation overall can increase drug circulation time from ten to a hundred times, improving its medical efficacy with little added toxicity. Most of all, it has been shown that PEG oligomers can also stabilize protein itself.⁴ However, the underlying mechanism of PEG-aided stabilization remains elusive.

A recent study of PEGylated human protein Pin1 WW domain⁵ found that PEG length is an important determinant of enhanced stability. Studies show that the excluded volume effect (crowding), in which the larger PEG chains take up surrounding space and hinder protein unfolding, can also account for PEG-aided stabilization.^{6,7} Our recent study on Pin1 WW showed that short PEG chains can restore secondary structure disrupted by mutation even without contacting the protein, and that a longer PEG chain can sample both solvent exposed and protein interaction scenarios, on a 10 ns time scale.⁸

While many studies have been conducted to understand the impact of a conjugated PEG molecule on its host protein, none so far have focused on how the host protein affects the conformation of PEG. In this paper, with the aim of understanding protein-PEG interaction more completely, we studied the problem from the other side, focusing on the conformation change of PEG. We used the well-studied protein λ_{6-85} as host for a 45-unit PEG chain. We studied two different cysteine mutants as PEG-maleimide attachment sites on opposite sides of the protein.

First, we ensured that the thermodynamic stability and folding kinetics of PEGylated λ_{6-85} do not change significantly from the un-PEGylated protein. Upon examining 1 μ s trajectories from all-atom molecular dynamics simulation, we found recurring and cooperative ‘folding’ of PEG 45-mer onto the protein surface. Specific surface regions interacting with the PEG molecule contain both hydrophobic and hydrophilic interactions with protein surface residues. We found that PEG is prone to stabilizing helices 3 and 4, even when its attachment site is far away in sequence and structure. A further examination of the simulation data shows that the PEG chain ‘slithers’ around certain charged side chains on the protein surface. The structure is recurrent, although subunits of PEG shift back and forth around the side chain, as can be expected from the translational symmetry of the PEG homopolymer. Thus the structural organization of PEG on the protein surface lies between random polymer behavior, and highly organized folding into a nearly unique structure.

EXPERIMENTAL AND COMPUTATIONAL PROCEDURES

Protein expression and purification

We used the stable Y22W/Q33Y/G46,48A mutant of λ_{6-85} , described previously,⁹ as our starting point because it has a large change of Trp fluorescence upon unfolding, and PEGylation is not likely to disrupt its structure much (Figure 1).¹⁰ The DNA fragment

coding for the targeted λ_{6-85} mutation was inserted between the BamHI and NdeI restriction sites of plasmid pET-15b (Genscript, Piscataway, NJ). Mutations N27C and N58C were made using a QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). BL21(DE3) cells (Agilent, Santa Clara, CA) were transformed with the plasmid and grown on an agar plate at 37 °C with 100 mg/L ampicillin. The survival cells were then grown in liquid culture. After reaching an OD₆₀₀ value of 0.6–1, the cells were induced with 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 20 °C for 12 hours, and then were pelleted and lysed using sonication in a solution containing 1 mM phenylmethylsulfonyl fluoride (PMSF) for protease inhibition. 20 μ l deoxyribonuclease (DNase) was added per 10 ml lysis buffer for DNA degradation. The lysate was centrifuged and the supernatant will be filtered with 0.45 μ m and 0.22 μ m syringe filters and purified using a liquid chromatography system (ÄKTA pure 25, GE Healthcare Life Sciences, Little Chalfont, UK). The sample was then dialyzed against 50 mM K₃PO₄, pH = 7.0, at 4 °C.

Protein PEGylation and purification

A $\approx$ 2 kDa methoxypolyethylene glycol maleimide with 45 PEG subunits (mPEG-Mal, Nanocs, Inc. Boston, MA) was covalently attached to the targeted cysteine sites via a thioether linkage. The host protein was first reacted with Tris(2-carboxyethyl) phosphine (TCEP) in a 1:30 molar ratio at room temperature for 2 hrs to reduce disulfide bonds. A 30-fold excess of mPEG-Mal was then added, and reacted with targeted mutant at 20 °C overnight. The reactions were conducted in nitrogen. The PEGylated conjugates were purified using a size-exclusion chromatography column Superdex 200 10/300 GL (GE Healthcare), shown in Figure S1. The result of PEGylation was verified using a mass spectrometry. We refer to the two mutant-PEG combinations as N27C-P45 and N58C-P45.

Equilibrium experiments and data analysis

Circular dichroism (CD) measurements were done on a Jasco-715 spectropolarimeter (Jasco Inc., Easton, MD). The spectra were collected between 200 and 250 nm. The sample concentration was 2.5 μ M. 2 mM 2-Mercaptoethanol (BME) was added to reduce the disulfide bonds. Fluorescence spectra were measured using a fluorescence spectrophotometer FP-8300 (Jasco Inc., Easton, MD). The photomultiplier tube (PMT) voltage was set in between 500–550 V. The excitation wavelength was 280 nm. The emission and excitation bandwidth were both 5 nm. The measurements were taken between 290 nm to 450 nm. The data interval was 1 nm. The scan speed was 500 nm/min. The sample concentration was 10 μ M, which also containing 2 mM BME. In both CD and Fluorescence measurements, each measurement consisted of three accumulations of spectra. The thermal melt was started from 5 °C and increased to 95 °C, with a 3 °C incremental step. The data analysis was done using Python. Fluorescence intensity was calculated by integrating the spectrum from 300 to 450 nm at a specific temperature: $I_{total}(T) = \int I(\lambda, T) d\lambda$. The mean wavelength was calculated according to the formula: $\langle \lambda(T) \rangle = \int \lambda I(\lambda, T) d\lambda / \int I(\lambda, T) d\lambda$. Two-state thermodynamic fits had the form: $F(T) = [F_D / (1 + K_{eq})] + [F_N K_{eq} / (1 + K_{eq})]$, where $K_{eq} = \exp[-G(T)/RT]$, the signals for the states i=D,N (denatured, native) are approximated by linear baselines $F_i = b_i + m_i (T - T_m)$; and $G(T) \cong G(T_m) + S(T_m) (T_m - T)$ is a linear approximation to the folding free energy. We note

that a linear temperature dependence of Trp fluorescence intensity is an approximation used here for simplicity.¹¹

Kinetic experiments and data analysis

Temperature jumps were performed using an in-house built instrument, which was described in detail previously.¹² Briefly, a sample consisting of the protein in 50 mM K₃PO₄ buffer at pH 7 was loaded in a cuvette made out of a rectangular capillary (Vitrocom) welded on one side. The sample concentrations were 45, 50, 40, and 40 μM, for N27C, N27C-P45, N58C, and N58C-P45, respectively. 2 mM 2-mercaptoethanol (BME) was added to reduce the disulfide bonds. The equilibrium temperature of the sample was maintained using a water flow system. The temperature of the solution was suddenly jumped using a nanosecond pulse of a Q-switched Nd:YAG laser (Continuum). The wavelength of the laser was Raman-shifted from 1,064 nm to 1.9 μm using a 1-m-long chamber with hydrogen compressed to 300 psi. The beam was split using a 50% reflective mirror, and uniformly heated the sample from both sides. Tryptophan was excited every 12.5 ns by a mode-locked Ti:sapphire laser (KMLabs), whose frequency was tripled to 280 nm. Fluorescence emission was collected at a perpendicular direction using an optical waveguide (Oriel), filtered from incident radiation with a bandpass filter (B370; Hoya), and detected by a photomultiplier tube (R7400U-03; Hamamatsu). The signal was digitized with a period of 100 ps (10 GHz; 125 points per fluorescence decay) and a bandwidth of 2.5 GHz using an oscilloscope (DPO7254; Tektronix). The data were analyzed using Igor Pro (Wavemetrics) and Matlab (MathWorks). To increase the signal to noise ratio, 60 kinetic traces were averaged at each temperature. The decays were normalized to yield a parameter $\chi(t)$ such that $\chi(t) = a_1(t)/(a_1(t) + a_2(t))$, where each fluorescence decay is represented as a linear combination of a decay before the temperature jump (f_1) and a decay at the very end of the kinetic trace (f_2), $f(t) = a_1(t) f_1 + a_2(t) f_2$. Thus, every trace begins with $\chi(t=0) = 1$ and ends with $\chi(t=1 \text{ ms}) = 0$. The resulting kinetic traces $\chi(t)$ were then fit to a single-exponential function of the forms $A \exp(-t/\tau)$.

Molecular dynamics simulations and data analysis

The structure of pseudo wild type λ_{6-85} , was obtained from a crystal structure (PDB: 3KZ3; resolution of 1.64 Å¹³). Mutants were prepared using Visual Molecular Dynamics (VMD) software.¹⁴ All-atom molecular dynamics simulations were performed using the program NAMD2¹⁵ with periodic boundary conditions, the CHARMM36 parameter set for the TIP3P water model,¹⁶ ions,¹⁷ protein,¹⁸¹ and PEG,^{19,20} and ion-pair specific corrections to the Lennard-Jones parameter σ .²¹ The force field of the Maleimide linker was developed using Force Field Toolkit (ffTK) plugin of VMD.²² The initial topology information and parameters were assigned by analogy using CHARMM General Force Field (CGenFF 1.0.0). Sodium and chloride ions were added to neutralize the whole system and to mimic approximately the non-zero ionic strength (50 mM phosphate buffer) of the experimental environment. The simulation box was 58 Å and 88 Å in each dimension for unPEGylated proteins and PEGylated proteins, respectively. The number of atoms was $\approx 20,000$ for unPEGylated protein and $\approx 69,000$ for PEGylated proteins, including protein, water molecules, and ions in both cases.

All simulations employed a 2 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 method²³ with a 1.0 Å grid spacing. Following a 2000-step minimization using a conjugate gradient method,²⁴ each system was equilibrated for 1 ns at a constant pressure of 1 atm using a Langevin piston pressure control method,²⁵ with temperature fixed at 293 K. In the unfolded host protein simulations, the systems were first heated to 513 K, causing the protein to unfold. Then the systems were cooled back to 363 K and remain in 363 K for the production run. Production simulations were then performed at constant equilibrium volume in the NVT ensemble for 1 μ s. The simulation trajectory was recorded every 24,000 steps. Trajectory analysis was performed using Tcl scripts with VMD. The protein-PEG distance is calculated as the nearest distance between any protein atom and any of the PEG oxygen atoms. The hydrogen bond formation was calculated with a criterion of donor-acceptor cutoff distance 3.5 Å and angle less than 150°. A 1.4 Å cutoff distance was used for solvent accessible surface area (SASA) calculation.

RESULTS

Thermodynamic and kinetic validation of λ_{6-85} mutants and their PEGylated conjugates

We targeted asparagine residues as our PEGylation sites for comparative N-PEGylation studies in the future. There are 5 asparagine residues on the pseudo wild type λ_{6-85} surface. N27 and N58 had the largest and second-largest solvent accessible surface areas (SASAs) in a preliminary molecular dynamics simulation (see Figure S2). We investigated these two sites further, as being least likely to affect protein structure and stability. The two sites are also on opposite sides of the protein surface (Figure 1a), which allows us to study the PEG-protein interaction when PEG is attached on separate regions of its host protein. The sequence of the mutants is shown in Figure 1b. We call the mutants “N27C” and “N58C,” and add “-P45” for the PEG-labeled versions.

The CD and fluorescence thermal melt data are shown in Figure 2 For N58C and Figure S3 for N27C. The parameters from two-state fitting (Methods) are summarized in Table S1. The CD spectra at 20 °C (Figure 2a, Figure S3a) show that PEGylation causes no significant secondary structural change in either mutant at 20 °C. Temperature melt data shows that N27C (Figure S3b–d) is about 5 °C more stable than N58C (Figure 2b–d). PEGylation causes only a 1–2 °C melting temperature shift as probed by CD thermal melt data (Figure 2b, Figure S3b), and even less by fluorescence thermal melt data (Fig. 2cd and S3cd). Comparing to published data for the pseudo wild type λ_{6-85} (without Asn-Cys mutations), the N27C mutant has almost the same thermodynamic stability as pseudo wild type, while N58C is slightly less stable (Table S1). In addition, we found that N58C is more prone to dimerization than N27C in the absence of BME, so its cysteine is more reactive.²⁶

The temperature jump relaxation kinetics also shows no significant difference before and after PEGylation for both mutants (Figure S4, S5). At 60 °C, all mutants fold with a relaxation time in the 19 to 24 μ s range.

Recurring formation of cooperative PEG structure on the templating protein surface

By examining the molecular dynamics trajectories of PEGylated mutants, we observe conformational transitions of PEG molecule, as we reported in our previous study.⁸ The radius of gyration R_g , which is a measure of PEG size, derived from the simulation data of N58C, shows this fluctuation over time (Figure 3a, movie S1). The PEG molecule constantly switches back and forth between coating a localized area on the protein surface, which results in a smaller R_g , and extending into the solvent, which results in a larger R_g (Figure 3a, d, e). PEG is also able to coat a larger area of the protein surface, which results in an intermediate R_g . For reference, the radius of gyration of free PEG chains in aqueous solution fits well to $R_g = 0.181N^{0.58}$ (nm),²⁷ which for $N=45$ yields 16.5 Å. The R_g we observe for extended contact periods with the protein surface is 10–12 Å in Figure 3a.

This intermittent ‘folding’ of PEG onto the surface can be monitored via hydrogen bonds and surface contacts as well. In Fig. 3b, hydrogen bonds form and break cooperatively, usually 2–4 at a time. In addition, there are long periods of extensive hydrogen bonding followed by periods with little hydrogen bonding (Fig. S6). The number of PEG-protein H-bonds and the number of PEG units making contact with the protein surface (within 3.5 Å) are both anti-correlated with R_g . When PEG collapses onto the protein surface, it has lower R_g , and higher number of hydrogen bonds and surface contacts (Fig 3 a–c, highlighted in purple); when PEG extends into the solvent, it has higher R_g , but lower number of hydrogen bonds and surface contacts (Figure 3a, c, highlighted in orange). However, for some other parts of the simulation, PEG can also form contacts with far-apart residues on the protein surface, yielding a high R_g for some bound PEG conformations (See SI movie S1).

The step-wise transitions from zero hydrogen bonds to multiple hydrogen bonds, and the extended periods for which hydrogen bonds persist, then rapidly dissolve, show that PEG binds to protein surface with cooperativity (Figure 3b, Figure S6). When PEG moves close to the surface residues, it first binds a single hydrophilic surface residue by hydrogen bond formation. The initial binding event rapidly forces nearby PEG monomers to also interact with neighboring residues more frequently, and quickly form more hydrogen bonds with them.

The contact map for PEG-protein surface binding patches

From the contact map between PEG and protein residues along the 1 μ s simulations, we observed specific binding patches on the protein surface (Figure 4). For mutant N58C, PEG was attached close to helix 4. PEG residues repeatedly bind to the residues in helix 4 (Figure 4a, blue boxes), primarily with hydrophilic –OH and –NH₂ groups on Y60, K67, and K70. PEG units also closely contact with a few hydrophobic residues like A63 and L64. We also observe repeated binding patterns in helix 1–3, but not as frequently as in helix 4.

For the N27C mutant, PEG is attached close to helix 1. Close contact was observed between PEG and helix 1 and 2. Surprisingly, in the N58C simulation PEG also extensively contacted a patch on helix 4, which is on the opposite side of the protein from the PEGylation site. This implies that helix 4 could be a strong binding partner for PEG. Could PEG differentially stabilize helices 1 and 4 in the unfolded state?

PEG has differential stability effects on denatured state structure

Helices 1 and 4 of lambda repressor fragment have been found both computationally and experimentally to orient properly and form native-like contacts early in folding.²⁸ We ran four simulations of unfolded proteins for 1 μ s each after equilibrating at 363 K (Methods): both mutants with and without PEG attached. We found that helix 4 has a tendency to form in all simulations, in agreement with previous findings that helices 1 and 4 form early on during folding.^{28–30} In contrast, helix 1 was suppressed in the two PEG simulations, irrespective of the attachment site of the PEG to λ_{6-85} , whereas helix 1 is relatively stable without PEG in Fig. 5. Thus PEG induces differential stability in secondary structure, leaving helix 4 relatively unaffected, while disfavoring helix 1 irrespective of PEGylation site. For PEGylated N27C, we observed instead some non-native β sheet formation near the attachment site during the simulation period (yellow in Figure 5b), indicating that PEG may stabilize sheet structure rather than helix structure near its attachment site on helix 1. In contrast, Fig. S7 shows that helices 3 and 4 appear just before (10–20 ns) PEG interacts with the helices, so the interaction with PEG may either help kinetically trap, or even stabilize these helices after they form transiently.

A closer look at the interaction between PEG and specific amino acid side chains on the templating protein surface

We visualize binding between PEG and protein surface from MD simulations (see Movie S2). As we described in the previous section, PEG is particularly prone to bind to helix 4 of λ_{6-85} , with conformations that show some polymer translational symmetry as the PEG chain slides around amino acid side chains on the protein surface (Figure 6).

The binding patch on helix 4 consists of both hydrophobic and hydrophilic residues (Figure 4a, Figure 6a–c). The hydrophilic residues Y60 and K67 and the hydrophobic residues A63 and L64 form a “valley-like” binding patch (Figure 6c), where Y60 and K67 are two hills and A63 and L64 are the bottom of the valley. In this binding patch, oxygen atoms of PEG units form hydrogen bonds with K67, and CH groups have hydrophobic interaction with hydrophobic residues and the aromatic ring of Y60. PEG is prone to wrap around lysine residues (Figure 6ab). We can find PEG wrapped around K39, K70, and K67 in Figure 6ab. Since each PEG unit is identical, they can freely slither around the side chain to form hydrogen bonds with lysine, resulting in a translation of the wrapped PEG back and forth (Figure 6e–i, Movie S2). Similarly, we also found PEG sliding in the valley of the binding patch (Movie S3).

DISCUSSION

PEGylation of proteins is desirable for pharmaceutical applications because PEG confers solubility and reduces degradation, yet it does so without inordinately affecting protein stability or function. NMR and crystallographic studies of small protein-PEG conjugates,³¹ like the λ_{6-85} -P45 conjugate we present here, show that PEG does not form permanently organized structure that could easily be seen as electron density in crystals or via NOEs or chemical shift effects in solution. However, such studies do not preclude PEG forming

highly organized recurring structure at the protein surface, as it equilibrates between solvent and surface.

Attaching a PEG to a protein surface effectively forces the two into closer contact. Studies with a surface force apparatus have shown that when PEG is forced against protein surfaces to which it is not attached, eventually a strong attractive interaction results.³² This interaction has been presumed to be non-specific, but is it?

Here we find that PEG structure can form transiently on the protein surface, and that it recurs in similar patterns, as PEG interacts with particular secondary structure elements and amino acid side chains on the surface of the protein. Due to the translational symmetry of the PEG homopolymer chain, it slides around binding patches, but with locally conserved binding motifs. Thus PEG structure equilibrates between solvent and protein surface, and when on the surface, occupies a realm between random coils and a highly structured fold.

Comparison of our two mutants, which attach PEG-45 to opposite sides on the protein, shows that preferential interactions recur, no matter the attachment site. Helices 3 and 4 of λ_{6-85} -P45 interact preferentially with PEG, whether it is attached at position 27 or 58 of the protein. On a more local scale, lysine side chains such as K39 and K67 organize PEG structure, which can 'slither' around the side chains when there are also local hydrophobic contacts available nearby. Other evidence for specific interactions of PEG with protein surfaces exists. For example, Cleland and Randolph showed that only a specific partly hydrophobic, partly solvent exposed site on a folding intermediate of carbonic anhydrase binds PEG, whereas the fully folded or unfolded states do not.³³

A previous study of much shorter PEG conjugated to WW domain showed that PEG can also stabilize proteins when it extends into solution⁸ (as in our Figure 3d). It does so by modifying local secondary structure propensity at the attachment site, increasing the stability of hairpin 1 β -sheets in WW domain. This propensity is not inconsistent with our findings here: Although the tail of PEG-45 attached at N27 preferentially interacts with helices 3 and 4 far away from the binding site, PEG-45 can increase the beta sheet propensity near its attachment site as seen in Fig. 5b. Thus different parts of the PEG molecule can stabilize different secondary structures on the protein surface.

We showed that PEG can become particularly structured around lysine residues surrounded by more hydrophobic small patches, such as the K67/A63/L64 or the K70/L29 combinations in Fig. 6. It has been shown that surface lysines and arginines are well-tolerated as surface mutations.^{34,35} Thus it may become possible to engineer protein surfaces to increase the efficacy of local PEG binding, by placing lysine side chains to hydrogen bond with the PEG oxygen, near hydrophobic patches that can interact with PEG methylene groups. Based on our results, the charged/hydrophobic pattern could be important: for example, helix 1 has three adjacent lysines, but these interact much less with PEG than the N-terminal amino group in Fig. 4a. Rather, the patterning of PEG relies on a conjunction of hydrogen bonding and hydrophobic interactions at the surface, as illustrated in Figure 6. This could be useful for improving the masking of sensitive sites on the protein surface in PEG-protein

bioconjugates,³⁶ and such transient but partly structured interactions can even allow PEG to replace part of a protein backbone on the surface.³⁷

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Basilio Cieza Huaman and Taras Pogorelov for helpful discussions, and Yi Zhang for help with running denatured protein simulations. The authors declare no competing financial interest.

Funding

Financial support was provided by the National Institutes of Health grant GM093318. The authors gladly acknowledge supercomputer time provided through XSEDE Allocation Grant MCA05S028 and the Taub Cluster (UIUC), as well as Anton Computer time at the Pittsburgh Supercomputing Center.

ABBREVIATIONS

PEG	poly-ethylene glycol
HPLC	high-performance liquid chromatography
VMD	Visual Molecular Dynamics
CHARMM	Chemistry at Harvard Molecular Mechanics
NMR	nuclear magnetic resonance
BME	beta-mercaptoethanol
CD	circular dichroism

References

1. Nischan N, Hackenberger CPR. Site-specific PEGylation of proteins: recent developments. *J Org Chem.* 2014; 79:10727–10733. [PubMed: 25333794]
2. Knop K, Hoogenboom R, Fischer D, Schubert US. Poly(ethylene glycol) in drug delivery: pros and cons as well as potential alternatives. *Angew Chem Int Ed.* 2010; 49:6288–6308.
3. Harris JM, Chess RB. Effect of pegylation on pharmaceuticals. *Nat Rev Drug Discov.* 2003; 2:214–221. [PubMed: 12612647]
4. Kinstler OB, Brems DN, Lauren SL, Paige AG, Hamburger JB, Treuheit MJ. Characterization and stability of N-terminally PEGylated rhG-CSF. *Pharm Res.* 1996; 13:996–1002. [PubMed: 8842035]
5. Pandey BK, Smith MS, Torgerson C, Lawrence PB, Matthews SS, Watkins E, Groves ML, Prigozhin MB, Price JL. Impact of Site-Specific PEGylation on the Conformational Stability and Folding Rate of the Pin WW Domain Depends Strongly on PEG Oligomer Length. *Bioconjug Chem.* 2013; 24:796–802. [PubMed: 23578107]
6. 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. [PubMed: 23106398]
7. Price JL, Powers ET, Kelly JW. N-PEGylation of a Reverse Turn Is Stabilizing in Multiple Sequence Contexts, unlike N-GlcNAcylation. *ACS Chem Biol.* 2011; 6:1188–1192. [PubMed: 21939258]
8. Chao SH, Matthews SS, Paxman R, Aksimentiev A, Gruebele M, Price JL. Two Structural Scenarios for Protein Stabilization by PEG. *J Phys Chem B.* 2014; 118:8388–8395. [PubMed: 24821319]

9. Yang WY, Gruebele M. Rate-temperature relationships in lambda-repressor fragment lambda(6-85) folding. *Biochemistry*. 2004; 43:13018–13025. [PubMed: 15476395]
10. Ghaemmaghami S, Word JM, Burton RE, Richardson JS, Oas TG. Folding kinetics of a fluorescent variant of monomeric lambda repressor. *Biochemistry*. 1998; 37:9179–9185. [PubMed: 9636065]
11. Naganathan AN, Munoz V. Determining denaturation midpoints in multiprobe equilibrium protein folding experiments. *Biochemistry*. 2008; 47:6752–6761. [PubMed: 18540681]
12. Prigozhin MB, Gruebele M. The fast and the slow: Folding and trapping of lambda6-85. *J Am Chem Soc*. 2011; 133:19338–19341. [PubMed: 22066714]
13. Liu F, Gao YGG, Gruebele M. A survey of lambda repressor fragments from two-state to downhill folding. *J Mol Biol*. 2009; 397:789–798.
14. Humphrey WF, Dalke A, Schulten K. VMD Visual Molecular Dynamics. *J Mol Graph*. 1996; 14:33–38. [PubMed: 8744570]
15. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26:1781–1802. [PubMed: 16222654]
16. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys*. 1983; 79:926–935.
17. Beglov D, Roux B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J Chem Phys*. 1994; 100:9050–9063.
18. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell Alexander DJ. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ 1 and χ 2 Dihedral Angles. *J Chem Theory Comput*. 2012; 8:3257–3273. [PubMed: 23341755]
19. Vorobyov I, Anisimov VM, Greene S. Additive and Classical Drude Polarizable Force Fields for Linear and Cyclic Ethers. *J Chem Theory Comput*. 2007; 3:1120–1133. [PubMed: 26627431]
20. Lee H, Venable RM, MacKerell AD, Pastor RW. Molecular dynamics studies of polyethylene oxide and polyethylene glycol: hydrodynamic radius and shape anisotropy. *Biophys J*. 2008; 95:1590–1599. [PubMed: 18456821]
21. 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.
22. Mayne CG, Saam J, Schulten K, Tajkhorshid E, Gumbart JC. Rapid parameterization of small molecules using the force field toolkit. *J Comput Chem*. 2013; 34:2757–2770. [PubMed: 24000174]
23. 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.
24. Payne MC, Teter MP, Allan DC, Arias TA, Joannopoulos JD. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev Mod Phys*. 1992; 64:1045–1097.
25. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys*. 1994; 101:4177–4189.
26. Bocedi A, Fabrin R, Pedersen JZ, Federici G, Iavarone F, Martelli C, Castagnola M, Ricci G. The extreme hyper-reactivity of selected cysteines drives hierarchical disulfide bond formation in serum albumin. *FEBS J*. 2016; 283:4113–4127. [PubMed: 27685835]
27. Kawaguchi S, Imai G, Suzuki J, Miyahara A, Kitano T. Aqueous solution properties of oligo- and poly(ethylene oxide) by static light scattering and intrinsic viscosity. *Polymer (Guildf)*. 1997; 38:2885–2891.
28. Larios E, Pitera JW, Swope WC, Gruebele M. Correlation of early orientational ordering of engineered λ 6–85 structure with kinetics and thermodynamics. *Chem Phys*. 2006; 323:45–53.
29. Liu YX, Strumpfer J, Freddolino PL, Gruebele M, Schulten K, Liu J, Freddolino PL, Gruebele M, Schulten KYXS, Liu YX, Strumpfer J, Freddolino PL, Gruebele M, Schulten K. Structural Characterization of lambda-Repressor Folding from All-Atom Molecular Dynamics Simulations. *J Phys Chem Lett*. 2012; 3:1117–1123. [PubMed: 22737279]
30. Portman JJ, Takada S, Wolynes PG. Microscopic theory of protein folding rates. II Local reaction coordinates and chain dynamics. *J Chem Phys*. 2001; 114:5082–5096.

31. Cattani G, Voegelé L, Crowley PB. Structure of a PEGylated protein reveals a highly porous double-helical assembly. *Nat Chem*. 2015; 7:823–828. [PubMed: 26391082]
32. Sheth SR, Leckband D. Measurements of attractive forces between proteins and end-grafted poly(ethylene glycol) chains. *Proc Natl Acad Sci*. 1997; 94:8399–8404. [PubMed: 9237988]
33. Cleland JL, Randolph TW. Mechanism of polyethylene-glycol interaction with the molten globule folding intermediate of bovine carbonic anhydrase-b. *J Biol Chem*. 1992; 267:3147–3153. [PubMed: 1310682]
34. Strub C, Alies C, Lougarre A, Ladurantie C, Czaplicki J, Fournier D. Mutation of exposed hydrophobic amino acids to arginine to increase protein stability. *BMC Biochem*. 2004; 5(1–6):9. [PubMed: 15251041]
35. Sine SM, Wang HL, Bren N. Lysine scanning mutagenesis delineates structural model of the nicotinic receptor ligand binding domain. *J Biol Chem*. 2002; 277:29210–29223. [PubMed: 12011092]
36. Veronese F, editor PEGylated Protein Drugs: Basic Science and Clinical Applications 2009 PEGylated Protein Drugs: Basic Science and Clinical Applications; 1287
37. Reinert ZE, Musselman ED, Elcock AH, Horne WS. A PEG-based oligomer as a backbone replacement for surface-exposed loops in a protein tertiary structure. *Chem Bio Chem*. 2012; 13:1107–1111.

Figure 1. Structure and sequences of λ_{6-85} mutants

(a) Crystal structure of the pseudo wild type λ_{6-85} (PDB: 3KZ3 (8)) with Trp–Tyr pair 22 (red) and 33 (yellow), and mutation sites 27 (cyan) and 58 (cyan) highlighted. (b) Sequence of pseudo wild type λ_{6-85} and two mutants N27C and N58C.

Figure 2. Equilibrium thermal denaturation of mutant N58C and its PEGylated counterparts (a) CD spectra at 20 °C for PEGylated (solid line) and unPEGylated (dashed line) N58C. (b) Mean residue ellipticity (MRE) at 222 nm as a function of temperature for PEGylated (circles=data, solid line=fit) and unPEGylated (triangles=data, dashed line=fit) N58C. (c) Mean wavelength of fluorescence spectra as a function of temperature for PEGylated (circles=data, solid line=fit) and unPEGylated (triangles=data, dashed line=fit) N58C. (d) Integrated fluorescence intensity as a function of temperature for PEGylated (circles=data, solid line=fit) and unPEGylated (triangles=data, dashed line=fit) N58C.

Figure 3. Conformational fluctuations of PEG-45 conjugated to λ_{6-85} mutant N58C

(a) Radius of gyration of PEG versus time. The black dots are raw data from every 48 ps. The blue line is a smoothed trace. (b) Number of hydrogen bond between PEG and the protein surface residues versus time. (c) Number of PEG units within 3.5 Å of the protein surface versus time. The faint gray line is raw data from every 48 ps. The black line is a smoothed trace. Orange and purple transparent blocks in (a)(b)(c) mark the corresponding regions in the same time frames for better visual comparison. (d)(e) The conformation of N58C-P45 conjugate at the moments pointed out in (a). The PEG molecule and maleimide linker are shown as blue sticks, with hydrogen atoms not shown. The protein is shown in a cartoon representation with helices in orange, coils and turns in white, and surface in transparent gray. The oxygen atoms of PEG within 3.5 Å of the protein surface are shown as gray spheres. A full trajectory analysis from which this highlight was taken is in SI Figure S6.

Figure 4. Closest PEG-residue distance for each residue

(a) N58C-P45 and (b) N27C-P45. The surface contacts are shown in gray. Hydrogen bond formation between PEG and individual residues is indicated by small red circles. The cutoff for H-bonds was 3.5 Å. The native 5-helix regions, color-coded as in Figure 1A, are shown next to the y-axis for reference.

Figure 5. Secondary structure for each residue in unfolded PEGylated and unPEGylated λ_{6-85} (a) N58C-P45 and N58C; (b) N27C-P45 and N27C. Secondary structure is calculated by using VMD,¹⁴ with the following color code: turn, cyan; β -sheet, yellow; isolated bridge, dark yellow; alpha helix, pink; 3–10 helix, blue; pi helix, red; coil, gray. The native helix regions are shown next to the left y-axis, sequence position next to the right y axis. The production run was at 363 K at constant equilibrium volume in the NVT ensemble for 1 μ s.

Figure 6. The interaction between PEG and the protein surface residues

(a)(b) Two snapshots showing that PEG partially collapsed onto the protein surface, with different conformations. The all-atom model of N58C and PEG are shown in surface and tube representative, respectively. The hydrophilic residues are colored in pink. The hydrophobic residues are colored in white. (c) Valley-like binding patch on the protein surface. The PEG units not in the binding patch are shown in transparent. (e–i) A series of snapshots showing PEG circling around a lysine side chain (K39). The time difference between each snapshot is 0.05 ns. PEG units around K39 are colored in light and dark blue to visually show the rotation of the PEG. The remaining PEG units are shown in transparent. See SI Movie S3 for a full visualization of this ‘slithering’ motion.