MANUSCRIPT

Solvent Reorganization in Stabilized Protein-Polymer Conjugates Visualized by 2D IR and NMR Spectroscopy

Raiza Maia[1,†], Xiaobing Chen[1,†], Emma Mulry[2], Matthew T. Eddy[2,*], and Carlos R. Baiz[1,*]

[1]Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, USA

[2]Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA

[†]Authors contributed equally

*Corresponding authors: cbaiz@cm.utexas.edu and matthew.eddy@ufl.edu

ABSTRACT

PEGylation is essential for effective function of biologics, shielding them from rapid degradation and clearance in the complex environment of the human body. Despite its significance, a mechanistic understanding of PEGylation's role in enhancing protein stability is incomplete, limiting the ability to design PEGylated proteins with predictable properties. Solvation, a well-known driving force in protein folding and stability, is hypothesized to play a central role in protein stabilization via PEGylation, but molecular mechanisms underlying solvent-driven stabilization are not well understood. Here, we investigated solvent dynamics and the interactions of solvent with the PEGylated carbohydrate recognition domain of human Galectin-3 (Gal3C) in aqueous solutions. Two-dimensional infrared (2D IR) spectroscopy, which captures subpicosecond molecular ensembles, revealed polymer-length dependent differences in protein dynamics and solvent dynamics for PEGylated Gal3C. Slower solvent dynamics correlated with increased conjugate thermal stability. Complementing these data, multi-dimensional NMR spectroscopy provided evidence that Gal3C conjugated to longer PEG forms a non-covalent interaction "shroud", which correlated with changes in dynamics of the solvent and protein backbone. Molecular dynamics (MD) simulations supported an interpretation of the experimental results that PEGylation did not reduce the protein's solvent accessible surface area. The integration of these data challenge the idea that PEGylation stabilizes conjugated proteins by dehydrating a protein's surface. Instead, these data support a mechanism whereby PEGylation improves protein stability by stabilizing the protein's solvation shell. These insights offer guidance for optimizing polymer length to achieve the desired thermal stability in biologics.

INTRODUCTION

Biologics are one of the fastest-growing sectors in the pharmaceutical industry[1,2], accounting for over 40% of all FDA-approved oncology drugs from 2000 to 2022 and have predominant roles in the development of biomarkers and many other disease treatments.[3] However, the molecular complexity providing their target specificity also makes biologics more vulnerable and less robust than small molecules. One of the most promising approaches to improving the stability of biologics is chemical conjugation with polymers, particularly poly(ethylene glycol) (PEG), to form protein-polymer conjugates.[4,5] While PEGylation has been shown to improve pharmacokinetic properties and prolong therapeutic efficacy of biologics,[6–11] its effects on protein stability and activity have been unpredictable, with studies reporting both an improvement and reduction in stability. [12–15] Further, much of the knowledge surrounding PEGylation's effects remains empirical rather than mechanistically understood. [9–13,16] These challenges underscore the need for predictive models that elucidate the molecular mechanisms governing PEGylation-driven protein stabilization.

Solvent interactions with proteins are well known to be among the most important driving forces in protein folding and unfolding.[17–21] In line with this view, growing evidence suggests that PEGylation enhances the thermal stability of conjugated proteins by modifying the solvation network surrounding the protein and altering its solvent-accessible surface area (SASA).[22] Computational simulations with PEGylated polypeptides and small proteins observed that PEGylation-drive increase in protein stability correlated with a reduction in SASA.[23,24] A combined experimental and computational study of a PEGylated small protein Pin1 indicated that PEGylation increased the disorder of water molecules in the protein solvation shell.[25] Coupled to this question is the related question of the extent of interaction between the protein and polymer and how the two are spatially organized. In general, two models are frequently invoked to describe PEG-protein interactions: the "dumbbell model" and the "shroud model." These models differ in the extent of interaction between the protein and conjugated polymer. In the dumbbell model, the protein and covalently attached PEG remain loosely connected and behave as largely independent molecules.[26–29]

In contrast, the shroud model adopts more persistent and extensive protein-PEG interactions, forming a more interwoven conformation.[16,30,31] We hypothesize that the more extensive protein-polymer interactions in the shroud model have a greater impact on the protein's solvation network, thereby conferring greater stability to the conjugated protein.

To test this hypothesis, we study the dynamics and solvent interactions of the PEGylated carbohydrate recognition domain of human Galectin-3 (Gal3C), a human β-galactoside-binding lectin that plays a significant role in cell adhesion[32] and potential cancer biology. [33,34] Our investigation compared Gal3C conjugated with PEG molecules of two different molecular weights, 550 Da and 5 kDa, to investigate the relationship between the conjugated polymer and properties of the conjugated protein. Combining data from two-dimensional infrared (2D IR) spectroscopy, nuclear magnetic resonance (NMR) spectroscopy, and molecular dynamics (MD) simulations, we compared the impact of the different-sized PEG molecules on the properties of conjugated Gal3C to test their effects on solvent accessibility and conjugate thermal stabilization. A unique strength of 2D IR spectroscopy is its ability to capture ultrafast molecular processes, which are leveraged to directly investigate protein-solvent interactions, solvent exposure, and the dynamics and organization of solvation networks surrounding proteins.[35] We used 2D IR to investigate how PEGylation altered the backbone of conjugated Gal3C and investigate the impact of PEGylation on protein-solvent interactions and solvent exposure. NMR spectroscopy provided complementary insights into the influence of conjugated PEG on the protein backbone at slower timescales and was also used to quantify the diffusion properties of PEGylated Gal3C. These measurements enabled us to correlate changes in solvation networks with the dumbbell and shroud models of protein-PEG interactions. MD simulations further complemented our experimental data by visualizing the conjugated polymer and providing a global characterization of solvation environments and backbone flexibility[36–38]. Together, these integrated approaches enabled us to correlate polymer length with changes in the solvation network of the conjugated

protein and its stability, providing mechanistic insights into the intentional design of protein-polymer conjugates with predictable properties.

RESULTS AND DISCUSSION

Conjugation of Gal3C with a PEG polymer.

To examine the impact of PEG on the conformational dynamics and solvent organization of conjugated Gal3C, we compared unconjugated Gal3C with two Gal3C conjugates, each linked to poly (ethylene glycol) (PEG) of molecular weight of either 550 Da or 5 kDa, referred to here as Gal3C-PEG 550 Da or Gal3C-PEG 5 kDa, respectively. Gal3C PEGylation was performed using maleimide-functionalized PEG, which reacted specifically with a single extrinsic cysteine at position 243 in Gal3C[T243C]. We refer to this modified protein as Gal3C throughout the text. Gal3C-PEG 5 kDa was prepared according to previously reported protocols[16], while Gal3C-PEG 550 Da was prepared using a similar protocol, substituting maleimide-PEG 5K with a functionalized maleimide-PEG starting material of 550 Da (Section S1).

Due to the inability to resolve unconjugated Gal3C and Gal3C-PEG 550 Da via gel filtration chromatography, PEGylation efficiency of Gal3C-PEG 550 Da was measured using fluorescence spectroscopy with the thiol-reactive fluorescent dye N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM). By measuring the fluorescence spectra of Gal3C-PEG 550 Da in the presence of CPM, we determined that >85% of Gal3C was successfully conjugated following the reaction protocol described in Section S1 (**Figure S1**). To test the activity of Gal3C-PEG 550 Da we recorded tryptophan fluorescence spectra to quantify the binding affinity of lactose to Gal3C-PEG 550 Da and compared this to binding affinity of lactose to Gal3C and Gal3C-PEG 5 kDa (**Figure S2**). The observed K_d

values for lactose to Gal3C, Gal3C-PEG 550 Da and Gal3C-PEG 5 kDa were essentially identical, confirming that Gal3C-PEG 550 Da retained native activity (**Figure S2**).

To assess the impact of PEG 550 Da conjugation on the thermal stability of Gal3C, we recorded circular dichroism (CD) spectra from 200 to 280 nm (**Figure S3**). CD spectra of Gal3C-PEG 550 Da at 30 ºC which were consistent with a well-folded protein containing mostly β-sheet secondary structure and are at most only marginally different from CD spectra of Gal3C and Gal3C-PEG 5 kDa recorded at the same temperature (**Figure S3**). Temperature-dependent CD spectra recorded with Gal3C-PEG 550 Da showed a single cooperative unfolding transition with an apparent T_m of ~58 ºC with a λ_{min} of 225 nm. The thermal unfolding behavior of Gal3C-PEG 550 Da was highly similar to that of Gal3C but contrasted starkly to the thermal unfolding behavior of Gal3C-PEG 5 kDa. Consistent with earlier findings,[16] thermal unfolding of Gal3C-PEG 5 kDa showed two separate unfolding transitions and the formation of an intermediate conformation, as evidenced by a shift in λ_{min} from 225 to 215 nm (**Figure S3**). This stark contrast in the thermal unfolding behavior between Gal3C-PEG 550 Da and Gal3C-PEG 5 kDa motivated a deeper investigation into the molecular mechanisms underlying these distinct unfolding pathways.

A 'shroud-like' protein-polymer conformation is observed for Gal3C PEGylated with PEG5K but not with PEG 550 Da.

The extent of interaction between the protein and polymer affects the diffusion behavior of the protein-polymer conjugate. For the dumbbell model, minimal interaction allows the protein and polymer to behave as independent molecules, and we would predict that the diffusion behavior of such a system would be nearly identical to the unconjugated protein. In contrast, for the shroud model, more extensive protein-polymer interactions would decrease the rate of diffusion of the system relative to the unconjugated protein. Thus, by measuring the diffusion properties of the conjugated protein, it is possible to distinguish between the two proposed models. NMR spectroscopy can be used to quantify the diffusion properties of conjugates

by measuring the rotational correlation time (τ_c) of the conjugated protein. We determined τ_c values by measuring the ^{15}N longitudinal (T_1) and transverse (T_2) relaxation times of the backbone. This approach has been used to quantitatively determine rotational correlation times for globular proteins up to ~25 kDa[39,40]. Previously, we determined the τ_c values for Gal3C and Gal3C–PEG 5 kDa (Section S3) and showed that PEGylation increased the rotational correlation time of Gal3C from 10.6 (±0.7) ns to 12.0 (±0.8) ns. This was consistent with observed increased line broadening in ^{15}N-edited ^{1}H NMR spectra of Gal3C–PEG 5 kDa and an increased transverse relaxation time. The increased τ_c value for Gal3C–PEG 5 kDa was interpreted as supporting a shroud-like conformation of the polymer within the protein-polymer conjugate. Here, we tested whether the shorter chain PEG in conjugated Gal3C-PEG 550 Da also formed a shroud-like conformation by measuring the τ_c value for Gal3C-PEG 550 Da. Unlike Gal3C–PEG 5 kDa, we determined a τ_c value for Gal3C-PEG 550 Da to be 10.6 (±1.0) ns, which is similar to the τ_c value measured for Gal3C and significantly shorter than the τ_c value measured for Gal3C–PEG 5 kDa (**Figure 1**, Section S3). The similar τ_c values for Gal3C and Gal3C-PEG 550 Da are consistent with similar linewidths observed in their respective ^{15}N-edited ^{1}H NMR spectra (**Figure S4**). Together, these data support a shroud-like model for Gal3C–PEG 5 kDa and a more dumbbell-like model for Gal3C-PEG 550Da.

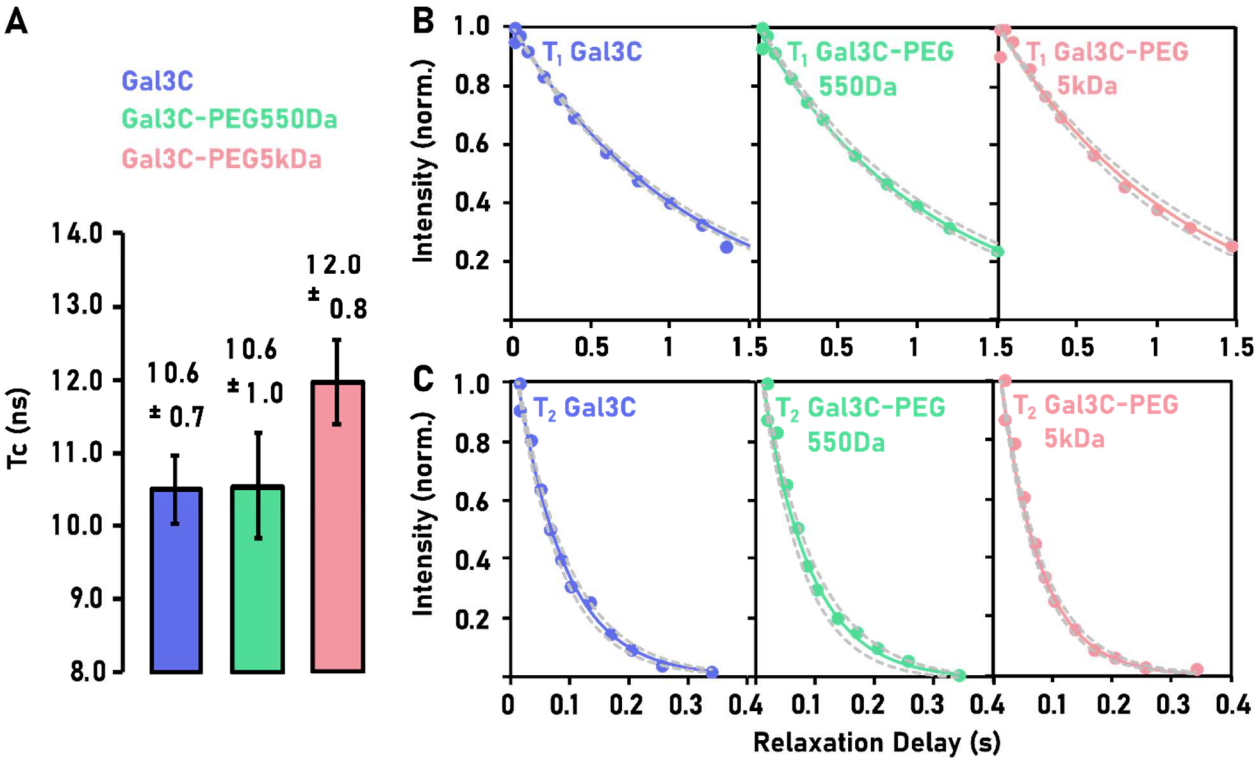

Figure 1. Rotational diffusion rates of Gal3C, Gal3C-PEG 5KDa, and Gal3C-PEG 550 Da. **(A)** Rotational correlation times (τ_c) of Gal3C (blue), Gal3C-PEG 550 Da (green) and Gal3C-PEG 5 kDa (pink) at 30 °C determined from ^{15}N relaxation. **(B)** Global ^{15}N longitudinal relaxation times (T_1) of amide groups in the structured core of Gal3C, Gal3C-PEG 550 Da and Gal3C-PEG 5 kDa. **(C)** Global ^{15}N transverse relaxation times (T_2) of amide groups in the structured core of Gal3C, Gal3C-PEG 550 Da and Gal3C-PEG 5 KDa. Solid circles represent normalized integral values obtained for each relaxation delay, and the solid lines are the non-linear regression used to quantify T_1 and T_2. Dotted lines indicate 95 % confidence intervals from the regression procedure. The T_1 and T_2 values are computed and presented in **Table S1**.

As shown in previous NMR studies of PEGylated proteins, 2D [^{15}N,^{1}H]-heteronuclear single quantum correlation (HSQC) spectra can reveal the impact of PEGylation on the protein backbone conformation and dynamics, as manifested by changes in chemical shifts or line broadening upon PEGylation. To observe the extent to which PEGylation with the shorter, 550 Da polymer impacted conjugated Gal3C structure and dynamics, we recorded 2D [^{15}N,^{1}H]-HSQC spectra of Gal3C–PEG 550 Da at 30 ºC (**Figure S4**). [^{15}N,^{1}H]-HSQC spectra of Gal3C–PEG 550 Da were well dispersed, consistent with a well-folded protein as expected from the CD measurements. Comparison of the [^{15}N,^{1}H]-HSQC spectra of [u-^{15}N] Gal3C–PEG 550 Da with [u-^{15}N] Gal3C and [u-^{15}N] Gal3C–PEG 5 kDa showed that the spectra

largely overlapped, indicating that conjugation with PEG 550 Da did not change the overall folding of Gal3C (**Figure S4**). We also compared these data with HSQC spectra of [u-^{15}N] Gal3C- PEG 550 Da to Gal3C conjugated with N-methylmaleimide without PEG, [u-^{15}N] Gal3C-NMM. These spectra were also largely overlapped. The high resolution and dispersion of the HSQC spectra facilitated a more detailed comparison of the chemical shifts and linewidths among [u-^{15}N] Gal3C, [u-^{15}N] Gal3C–PEG 550 Da and [u-^{15}N] Gal3C–PEG 5 kDa. By comparing chemical shift perturbations between Gal3C and Gal3C–PEG 550 Da, we observed that conjugation with PEG 550 Da resulted in minor perturbations, primarily localized to the conjugation site, indicating that conjugation with PEG 550 Da resulted in only minor perturbations to the Gal3C backbone (**Figure 2**). Except for one residue near position 220, we observed at most only minimal line broadening, typically less than 20 Hz, across all residues in spectra of Gal3C- PEG 550 Da.

In contrast to these observations, more pronounced differences were observed when comparing HSQC spectra of Gal3C and Gal3C–PEG 5 kDa. Conjugation of Gal3C with PEG 5 kDa resulted in larger chemical shift perturbations, particularly between residues 134–140, 193–200, 219, and 249. Additionally, Gal3C–PEG 5 kDa exhibited more extensive line broadening, with most residues showing increased line broadening of ≥ 20 Hz, and residues near the conjugation site showing greater line broadening of ≥ 40 Hz. Together, the HSQC data and rotational diffusion measurements for Gal3C–PEG 550 Da and Gal3C–PEG 5 kDa suggest that the longer PEG chain interacts more extensively with Gal3C, influencing both residue-level conformation and dynamics, as well as macromolecular diffusion properties. These findings support a model in which Gal3C–PEG 5 kDa adopts a shroud-like conformation, whereas Gal3C–PEG 550 Da assumes a more dumbbell-like structure.

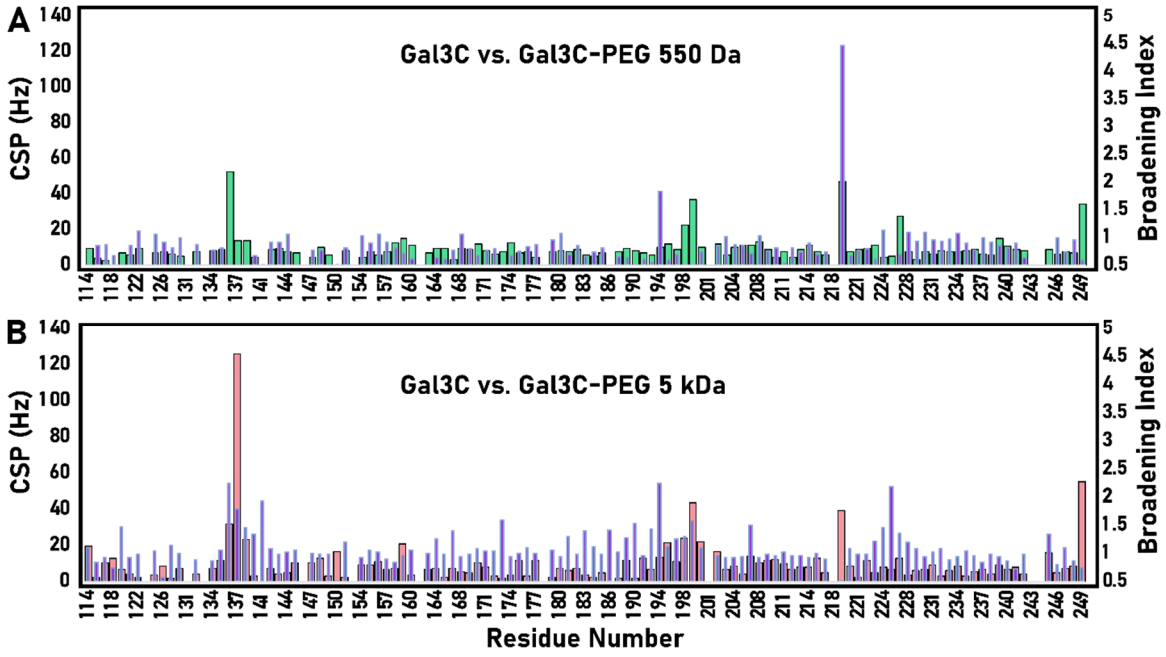

Figure 2. Histogram of the chemical shift perturbations (green and pink bars) and line broadening (purple bars) observed for **(A)** Gal3C relative to Gal3C-PEG 550 Da, **(B)** Gal3C relative to Gal3C-PEG 5 kDa as a function of the residue position.

2D IR reveals PEGylation-induced protein conformational and environmental changes on ultrafast timescales.

We investigated the impact of PEGylation on the conformation and environment around the Gal3C backbone on ultrafast timescales by recording 2D IR spectra of Gal3C, Gal3C-PEG 550 Da and Gal3C-PEG 5 kDa with two waiting times of 150 fs and 1400 fs shown in **Figure 3**. The amide I band, corresponding to the backbone amide I stretching mode, provides insights into changes to the global protein structure and environment, and is used to determine the impact of PEGylation. The 2D spectra of unconjugated Gal3C, Gal3C-PEG 550 Da and Gal3C-PEG 5 kDa show one pair of major peaks along the diagonal, where the red and blue colors represent positive or negative peaks, corresponding to transitions from ground to first excited, or first-to-second excited states respectively.[35,41] In all spectra, an elongated peak shape across the amide I stretching region is observed with two main peak centers near 1630 and 1680 cm^{-1}, corresponding

to the anti-parallel and parallel β-sheet vibrations, respectively[35] (**Figure 4**). Notably, the positive peak elongation is more pronounced for Gal3C-PEG 550 Da compared to unconjugated Gal3C, and this elongation is further enhanced for Gal3C-PEG 5 kDa. This trend is observed at both waiting times of 150 fs and 1400 fs, the peak elongation introduced by PEGylation could be related to three main factors: the increase of structural heterogeneity and/or the solvent exposure, energy transfer among the amide I modes. The more intense elongation in the positive peak for Gal3C-PEG 5 kDa implies that the long PEG chain would induce more significant conformational flexibility and/or solvent interaction compared to the small PEG chain. Additionally, at short waiting time (t_2) such as 150 fs (**Figure 3**, left), the peak is diagonally elongated, indicating strong correlation between excitation and detection frequencies. This correlation is lost at longer t_2 due to the effects mentioned above (**Figure 3, right**, **Figures S5-S7**).

Amide I absorption spectra report on backbone structure. Here we computed the pump slice amplitude (PSA) from the 2D IR spectra at 150 fs to extract features comparable to FTIR spectra. In brief, PSA analysis projects a 1D spectrum at each excitation frequency and computes the difference between the maxima and the minima of a series of pump-probe slices[42]. We selected a waiting time of 150 fs to avoid pulse-overlap artifacts while ensuring that peak shapes remained unaffected by energy dissipation, which occurs at longer waiting times. The PSAs of both Gal3C and Gal3C–PEG 550 Da exhibit a maximum around 1630 cm^{-1} (**Figure 4**A). Relative to Gal3C and Gal3C–PEG 550 Da, the maximum of Gal3C–PEG 5 kDa shows a 5 cm^{-1} redshift, likely indicating an increase in antiparallel β-sheet structures or changes in the backbone dihedral angles that lower the vibrational frequency. Both Gal3C–PEG 550 Da and Gal3C–PEG 5 kDa exhibit broader lineshapes from 1650 to 1680 cm^{-1} compared to Gal3C, but this is more pronounced for Gal3C–PEG 5 kDa. These findings indicate that PEGylation, particularly with the long

PEG chain (5 kDa), increases the structural heterogeneity, likely due to increased flexibility and solvent exposure in β-turn/loop regions of the backbone.[43]

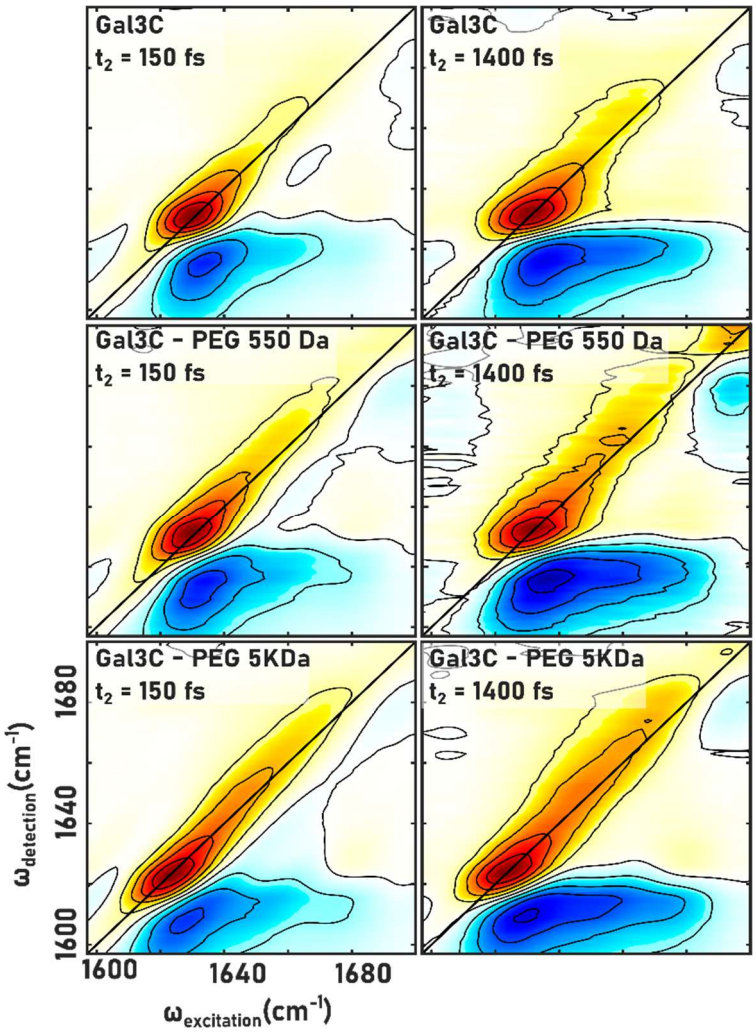

Figure 3. 2D IR spectra of Gal3C, Gal3C–PEG 550 Da and Gal3C–PEG 5 kDa at two selected waiting times, t_2 = 150 and t_2 = 1400 fs, as indicated in the panels.

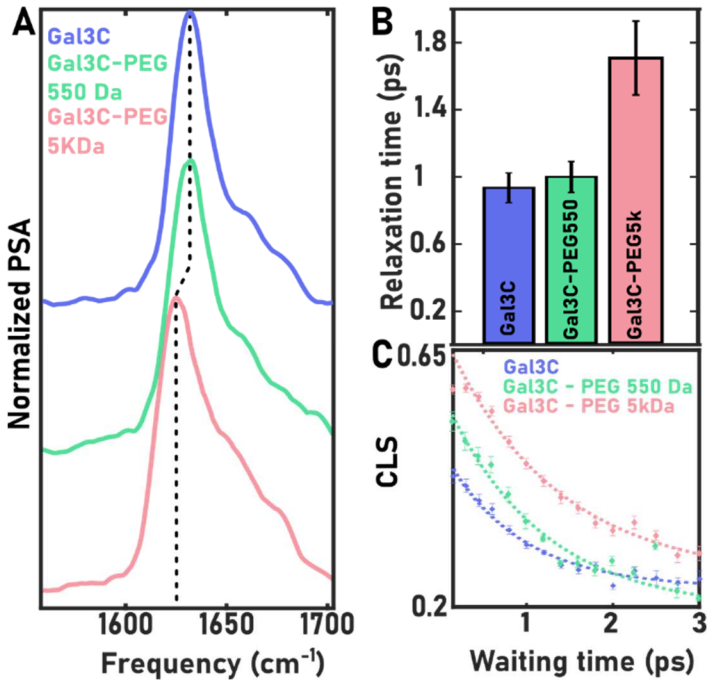

Figure 4. 2D IR spectral analysis for protein structure and dynamics. **(A)** PSA at waiting time of 150 fs for Gal3C[T243C], Gal3C[T243C] – PEG 550 Da and Gal3C[T243C] – PEG 5KDa. **(B)** Relaxation time constants extracted from the exponential fits of CLS analysis. Fitting parameters were evaluated through a bootstrapping technique to estimate the standard error of the fits. The error bars are the standard errors of the time constants estimated from the standard deviation of the fitted curves. **(C)** CLS decay as a function of the waiting time. The CLS curves fit to a monoexponential function (dashed lines). See Section S4 for details of CLS analysis.

Conjugation with longer PEG alters the Gal3C surface solvation network.

The 2D IR lineshape evolution for the three systems under investigation were analyzed using center line slope (CLS) analysis (**Figure 4B** and **C**). The CLS is obtained by identifying the detection frequencies (vertical axis) with the ridge of maximum intensity along the excitation axis (horizontal axis) and then fitting a linear slope through these points (Section S4) [44,45]. The CLS decays fit well to a monoexponential function (**Figure 4C**), serving as an experimental measure of the frequency–frequency correlation function (FFCF). More specifically, FFCF reflects the decorrelation between excitation and detection frequencies with waiting time, capturing the environmental fluctuations along the backbone[44,46]. The CLS analyses indicate similar decay time constants for Gal3C and Gal3C-PEG 550 Da, but Gal3C–PEG 5 kDa shows a

decay almost twice as slow as that of the other two proteins. These results suggest that the long PEG chain significantly slows the local dynamics, while the short PEG chain has minor influence on dynamics.

The picosecond frequency fluctuations extracted from 2D IR can report on the hydrogen bond (H-bond) dynamics between the amide group and surrounding water molecules[47–49]. The origins of these effects on PEGylated protein dynamics could be due to multiple reasons including the more restricted fluctuations of the residues upon polymer contact, backbone dehydration, and less mobile water around the backbone. The different outcomes of short and long PEG chains appear to indicate different configuration models. It is most likely that the long PEG adopts a "shroud" model where the contact of PEG chains with certain protein residues can disrupt H-bond rearrangements around the protein. Hence, the reorientation of water molecules necessary for H-bond exchange may be restricted by the PEG chain. This disruption can prolong C=O–water H-bond lifetimes, potentially causing slower H-bond reorganization around the protein. However, based solely on the experimental results, it is difficult to pinpoint and visualize the main factor for the slower dynamics in Gal3C-PEG 5 kDa.

MD simulations help elucidate and visualize the structure and environment of Gal3C proteins.

To test the "shroud" model hypothesis with respect to the altered local environment along the protein backbone, molecular dynamics (MD) simulations were performed for three systems: Gal3C, Gal3C-PEG 600 Da and Gal3C-PEG 5 kDa. The root mean squared displacement (RMSD) of protein residues was computed to assess the equilibrium of the simulated proteins, which shows that the structure reaches a stable state for the last 30 ns out of the 100 ns trajectory (**Figure S8**). As a result, all structural analyses were conducted using the final 30 ns of the trajectory. To visualize the effects of the short and long PEG chain, the modeled structures of the PEGylated Gal3C are shown in **Figure 5A**. More views of these modeled structures are shown in **Figure S9**. In the case of the short PEG, only a small region around the conjugation site of residue 243 is covered, resembling the protein wearing a "hat". In contrast, PEG 5 kDa is highly

flexible and long enough to cover a substantial portion of the protein surface, particularly the back and sides, giving the appearance of a "backpack". Consequently, Gal3C-PEG 5 kDa has more "shrouding" features, with a greater potential for global effects on the protein backbone, for example, dehydration, protection, and shielding from the external environment.

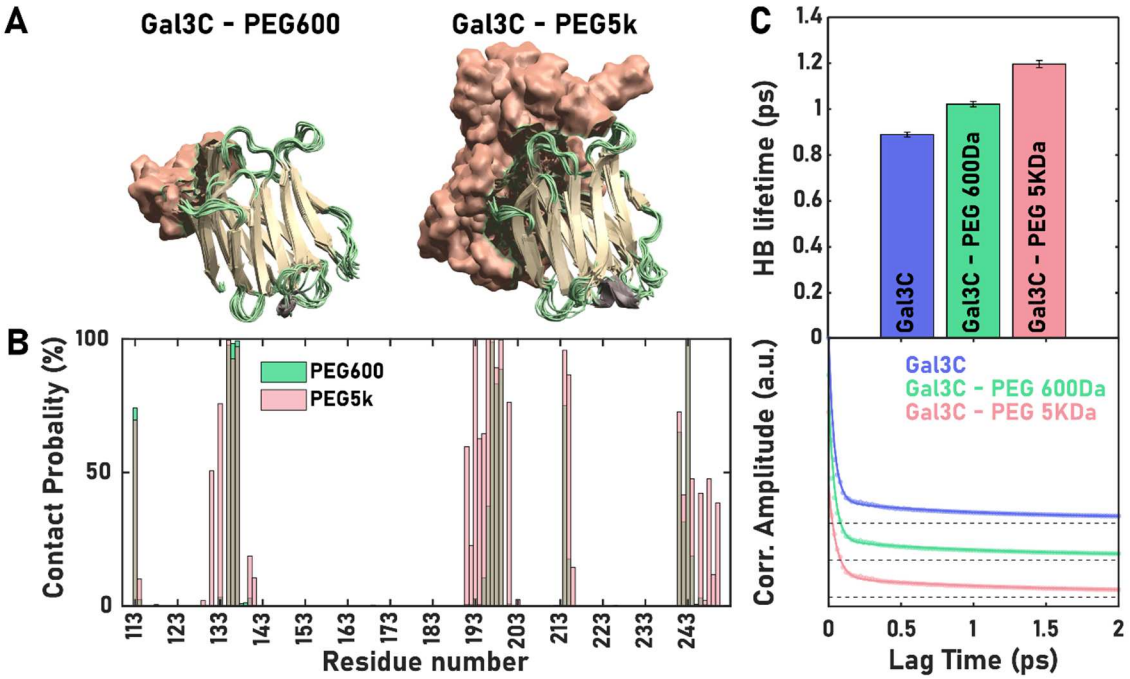

Figure 5. A summary of important results from MD analysis. **(A)** Ten snapshots taken from the MD trajectories every 10 ns for both Gal3C–PEG 550 Da and Gal3C–PEG 5 kDa were superimposed. PEG chains are shown in the pink surface representation, while the yellow, purple and green colors represent the β-sheet, random coil/turn and α-helical secondary structure of Gal3C, respectively (ribbon representation). **(B)** Contact probability (%) between PEG and each Gal3C residue was computed with a cutoff distance of 0.35 nm. **(C)** HB lifetime analysis and associated fittings for the correlation functions of hydrogen bond number. Details of the MD analysis are described in Section S5.

To gain further insights into the contacts between the protein and PEG chains, the minimum distance from the PEG chain to each protein residue was computed. The cutoff distance is set to be 0.35 nm, meaning when the PEG-residue distance is below 0.35 nm, it is considered as direct contact. The contact probability for each residue is then calculated as the number of MD frames when it is in contact with PEG over the total number of frames (**Figure 5B**). This analysis shows that atoms in the long PEG chain exhibit more sustained interactions with protein residues than those in the short PEG chain. These differences are evident

for the residues between 131-141, 191-202, 214-216 and 241-250 (**Figure 5B**), this is in alignment with the chemical shift perturbations measured by NMR, where the strongest perturbations were expected around residues 134-140, 193-200, 219 and 249 (**Figure 2**) In addition, the long PEG chain not only has a higher probability of staying in contact with the protein but also maintains these interactions for a longer duration compared to the short PEG chain (**Figure S10**). Notably, the contacts between the long PEG chain and protein are especially prolonged for residues 131-141 and 191-202 (**Figure S10**).

The root mean squared fluctuations (RMSF) of the protein residues was computed and compared between Gal3C before and after PEGylation to test the theory of PEGylation effects on protein mobility (**Figure S11**). For both PEG 600Da and 5KDa, most of the protein residues display lower fluctuations, meaning PEGylation introduces a global constraint across the whole protein. To visualize the effects of PEGylation on the protein fluctuations, the color-coded modeled structures of Gal3C-PEG 600 Da and -PEG 5 kDa are shown in **Figure S12**. The results show that the front (sugar-binding surface) and back of the protein are minimally affected by PEGylation, while parts of the side β-sheets and most of the random coils experience more restricted mobility due to the presence of the PEG chain. Although PEG 5 kDa shows more pronounced restriction effects than PEG 600 Da, the affected residues are largely overlapping between the two structures. Therefore, the difference between short and long PEG in the results of NMR and 2D IR is unlikely due to the mobility restrictions imposed by PEGylation.

Conjugation with longer PEG does not dehydrate the protein surface but alters solvent dynamics and the lifetimes of protein-solvent interactions.

To test the impact of PEGylation on solvent exposure of the Gal3C backbone, we computed the solvent accessible surface area (SASA) for the three systems (**Figure S13**). Overall, the effects of PEGylation on protein hydration are quite diverse depending on the position of the residues. In agreement with NMR data[16], the protein sugar-binding pocket surface was minimally affected by the PEG chain.

Though the changes in residue hydration are more prominent in PEG 5 kDa, the altered residues largely overlap between the two PEGylated proteins. We quantified the total number of hydrogen bonds between water and the backbone. Interestingly, the H-bond number of Gal3C, Gal3C-PEG 600 Da and Gal3C-PEG 5 kDa were all similar at 77.1, 76.5, and 76.4, respectively. This indicates that conjugation with a longer PEG does not dehydrate the protein surface and supports the notion that the differences observed in the behavior of Gal3C-PEG 5 kDa, compared to Gal3C-PEG 600 Da and unconjugated Gal3C in 2D IR and NMR experiments, are not attributable to replacement of water molecules. [39,4036–38]

The H-bond lifetime, averaged among all the H-bonds formed between protein amide groups and the nearby water molecules (**Figure 5C**), quantifies effects of PEGylation on water dynamics around protein backbone. The H-bond autocorrelation exhibits a decay constant on the picosecond scale, which is associated with the process of H-bond formation and dissociation[50,51]. Short PEG chain causes a ~10% slowdown in the lifetime while the long PEG chain leads to a ~30% slowdown, as compared to unconjugated Gal3C (**Table S3**). In the dynamics analysis of 2D IR above, the frequency fluctuations in the amide I vibrational range can be interpreted as related to the rapid H-bond fluctuations between protein amide groups and nearby water molecules[47–49]. The computed H-bond lifetime confirms our measurements of perturbed H-bond fluctuations upon PEGylation, as evidenced by the extended H-bond lifetimes computed from MD simulations. Moreover, the time scales of computed H-bond lifetimes are comparable to the relaxation time extracted from CLS analysis of 2D IR, where the slowdown in H-bond lifetime qualitatively agrees with the slower dynamics observed in the 2D IR experiments for Gal3C-PEG 5 kDa (**Figure 4B**).

CONCLUSIONS

The 2D IR measurements and MD simulations support a model in which PEGylation stabilizes Gal3C by reorganizing solvent dynamics within the surrounding protein solvent shell without any measurable changes to the protein secondary structure. Neither 2D IR nor MD simulations detected

evidence for PEGylation-driven desolvation of the protein surface or significant differences in the SASA among Gal3C, Gal3C-PEG 550 and Gal3C-PEG 5k. While the SASA was investigated in the present study at picosecond to nanosecond timescales, our findings align with earlier observations from NMR spectroscopy data that showed PEGylation did not alter the SASA of Gal3C over longer timescales.[52] The present conclusions appear to differ with those from earlier computational studies[38] and hydrogen-deuterium exchange experiments[53] that reported PEGylation reduced the conjugated protein's SASA. Instead, we observed that PEGylation slows solvent dynamics surrounding the surface of Gal3C, with the magnitude of this effect correlating with the length of the conjugated polymer. PEGylation-driven slowing of solvent dynamics correlated with the formation of a shroud-like protein-PEG conformation and with the increased thermal stability and redirection of the Gal3C unfolding pathway. By integrating experimental and computational approaches across multiple timescales, 2D IR and MD simulations for fast solvent dynamics, and NMR spectroscopy for slower protein-polymer interactions, we not only validate the accuracy of the simulations but also reveal a mechanistic link between polymer length and conjugate stability. Our findings suggest that a minimum polymer length is required to sufficiently slow protein-solvent dynamics and enhance stability, providing criteria for the intentional design of stabilized conjugates.

Data Availability Statement: The data supporting this article have been included as part of the Supplementary Information.

Supporting Information: The Supporting Information contains protein expression and characterization protocols including circular dichroism, and NMR spectroscopy. The SI also contains experimental details of the 2D IR measurements along with measured 2D IR spectra and analyses. The SI also contains details of the MD simulations and analysis of the trajectories.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health (R35GM133359), and the Welch Foundation (F-1891). Simulations were run on the Lonestar6 cluster at the Texas Advanced Computing

Center (TACC). M.T.E. and E.M. were supported by an NSF CAREER award 2339330 through the Division of Materials Research. A portion of this work was supported by the McKnight Brain Institute at the National High Magnetic Field Laboratory's AMRIS Facility, which is funded by the National Science Foundation Cooperative Agreement No. DMR-1644779.

REFERENCES

(1) Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front Microbiol* **2019**, *10* (JUN). https://doi.org/10.3389/fmicb.2019.01404.

(2) Anselmo, A. C.; Gokarn, Y.; Mitragotri, S. Non-Invasive Delivery Strategies for Biologics. *Nature Reviews Drug Discovery*. Nature Publishing Group December 28, 2018, pp 19–40. https://doi.org/10.1038/nrd.2018.183.

(3) Scott, E. C.; Baines, A. C.; Gong, Y.; Moore, R.; Pamuk, G. E.; Saber, H.; Subedee, A.; Thompson, M. D.; Xiao, W.; Pazdur, R.; Rao, V. A.; Schneider, J.; Beaver, J. A. Trends in the Approval of Cancer Therapies by the FDA in the Twenty-First Century. *Nature Reviews Drug Discovery*. 2023. https://doi.org/10.1038/s41573-023-00723-4.

(4) Pasut, G.; Veronese, F. M. State of the Art in PEGylation: The Great Versatility Achieved after Forty Years of Research. *Journal of Controlled Release*. July 20, 2012, pp 461–472. https://doi.org/10.1016/j.jconrel.2011.10.037.

(5) Grace, M. J.; Lee, S.; Bradshaw, S.; Chapman, J.; Spond, J.; Cox, S.; DeLorenzo, M.; Brassard, D.; Wylie, D.; Cannon-Carlson, S.; Cullen, C.; Indelicato, S.; Voloch, M.; Bordens, R. Site of Pegylation and Polyethylene Glycol Molecule Size Attenuate Interferon-α Antiviral and Antiproliferative

Activities through the JAK/STAT Signaling Pathway. *Journal of Biological Chemistry* **2005**, *280* (8), 6327–6336. https://doi.org/10.1074/jbc.M412134200.

(6) Milton Harris, J.; Chess, R. B. Effect of Pegylation on Pharmaceuticals. *Nature Reviews Drug Discovery*. March 2003, pp 214–221. https://doi.org/10.1038/nrd1033.

(7) Krall, N.; Da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. Site-Selective Protein-Modification Chemistry for Basic Biology and Drug Development. *Nature Chemistry*. Nature Publishing Group February 1, 2016, pp 103–113. https://doi.org/10.1038/nchem.2393.

(8) Shaunak, S.; Godwin, A.; Choi, J. W.; Balan, S.; Pedone, E.; Vijayarangam, D.; Heidelberger, S.; Teo, I.; Zloh, M.; Brocchini, S. Site-Specific PEGylation of Native Disulfide Bonds in Therapeutic Proteins. *Nat Chem Biol* **2006**, *2* (6), 312–313. https://doi.org/10.1038/nchembio786.

(9) Gupta, S. K.; Pittenger, A. L.; Swan, S. K.; Marbury, T. C.; Tobillo, E.; Batra, V.; Sack, M.; Glue, P.; Jacobs, S.; Affrime, M. Single-Dose Pharmacokinetics and Safety of Pegylated Interferon-α2b in Patients with Chronic Renal Dysfunction. *The Journal of Clinical Pharmacology* **2002**, *42* (10), 1109–1115. https://doi.org/10.1177/009127002237996.

(10) Zhang, B.; Sun, J.; Wang, Y.; Ji, D.; Yuan, Y.; Li, S.; Sun, Y.; Hou, Y.; Li, P.; Zhao, L.; Yu, F.; Ma, W.; Cheng, B.; Wu, L.; Hu, J.; Wang, M.; Song, W.; Li, X.; Li, H.; Fei, Y.; Chen, H.; Zhang, L.; Tsokos, G. C.; Zhou, D.; Zhang, X. Site-Specific PEGylation of Interleukin-2 Enhances Immunosuppression via the Sustained Activation of Regulatory T Cells. *Nat Biomed Eng* **2021**, *5* (11), 1288–1305. https://doi.org/10.1038/s41551-021-00797-8.

(11) Grigoletto, A.; Mero, A.; Zanusso, I.; Schiavon, O.; Pasut, G. Chemical and Enzymatic Site Specific PEGylation of HGH: The Stability and in Vivo Activity of PEG-N-Terminal-HGH and PEG-

Gln141-HGH Conjugates. *Macromol Biosci* **2016**, *16* (1), 50–56. https://doi.org/10.1002/mabi.201500282.

(12) Zaghmi, A.; Mendez-Villuendas, E.; Greschner, A. A.; Liu, J. Y.; de Haan, H. W.; Gauthier, M. A. Mechanisms of Activity Loss for a Multi-PEGylated Protein by Experiment and Simulation. *Mater Today Chem* **2019**, *12*, 121–131. https://doi.org/10.1016/j.mtchem.2018.12.007.

(13) Baker, S. L.; Munasinghe, A.; Murata, H.; Lin, P.; Matyjaszewski, K.; Colina, C. M.; Russell, A. J. Intramolecular Interactions of Conjugated Polymers Mimic Molecular Chaperones to Stabilize Protein-Polymer Conjugates. *Biomacromolecules* **2018**, *19* (9), 3798–3813. https://doi.org/10.1021/acs.biomac.8b00927.

(14) Chang, P. K.; Prestidge, C. A.; Barnes, T. J.; Bremmell, K. E. Impact of PEGylation and Non-Ionic Surfactants on the Physical Stability of the Therapeutic Protein Filgrastim (G-CSF). *RSC Adv* **2016**, *6* (82), 78970–78978. https://doi.org/10.1039/c6ra16254a.

(15) Hauptstein, N.; Pouyan, P.; Kehrein, J.; Dirauf, M.; Driessen, M. D.; Raschig, M.; Licha, K.; Gottschaldt, M.; Schubert, U. S.; Haag, R.; Meinel, L.; Sotriffer, C.; Lühmann, T. Molecular Insights into Site-Specific Interferon-A2a Bioconjugates Originated from PEG, LPG, and PEtOx. *Biomacromolecules* **2021**, *22* (11), 4521–4534. https://doi.org/10.1021/acs.biomac.1c00775.

(16) Pritzlaff, A.; Ferré, G.; Mulry, E.; Lin, L.; Gopal Pour, N.; Savin, D. A.; Harris, M. E.; Eddy, M. T. Atomic-Scale View of Protein-PEG Interactions That Redirect the Thermal Unfolding Pathway of

PEGylated Human Galectin-3. *Angewandte Chemie - International Edition* **2022**, *61* (40). https://doi.org/10.1002/anie.202203784.

(17) Cheung, M. S.; García, A. E.; Onuchic, J. N. Protein Folding Mediated by Solvation: Water Expulsion and Formation of the Hydrophobic Core Occur after the Structural Collapse. *Proc Natl Acad Sci U S A* **2002**, *99* (2). https://doi.org/10.1073/pnas.022387699.

(18) Eisenberg, D.; Mclachlan, A. D. Solvation Energy in Protein Folding and Binding. *Nature* **1986**, *319* (6050). https://doi.org/10.1038/319199a0.

(19) Roder, H.; Colón, W. Kinetic Role of Early Intermediates in Protein Folding. *Curr Opin Struct Biol* **1997**, *7* (1). https://doi.org/10.1016/S0959-440X(97)80004-8.

(20) Rhee, Y. M.; Sorin, E. J.; Jayachandran, G.; Lindahl, E.; Pande, V. S. Simulations of the Role of Water in the Protein-Folding Mechanism. *Proc Natl Acad Sci U S A* **2004**, *101* (17). https://doi.org/10.1073/pnas.0307898101.

(21) Fernández-Escamilla, A. M.; Cheung, M. S.; Vega, M. C.; Wilmanns, M.; Onuchic, J. N.; Serrano, L. Solvation in Protein Folding Analysis: Combination of Theoretical and Experimental Approaches. *Proc Natl Acad Sci U S A* **2004**, *101* (9). https://doi.org/10.1073/pnas.0304180101.

(22) Lawrence, P. B.; Price, J. L. How PEGylation Influences Protein Conformational Stability. *Current Opinion in Chemical Biology*. 2016. https://doi.org/10.1016/j.cbpa.2016.08.006.

(23) Hamed, E.; Xu, T.; Keten, S. Poly(Ethylene Glycol) Conjugation Stabilizes the Secondary Structure of α-Helices by Reducing Peptide Solvent Accessible Surface Area. *Biomacromolecules* **2013**, *14* (11). https://doi.org/10.1021/bm401164t.

(24) Meng, W.; Guo, X.; Qin, M.; Pan, H.; Cao, Y.; Wang, W. Mechanistic Insights into the Stabilization of SrcSH3 by PEGylation. *Langmuir* **2012**, *28* (46). https://doi.org/10.1021/la303466w.

(25) Lawrence, P. B.; Gavrilov, Y.; Matthews, S. S.; Langlois, M. I.; Shental-Bechor, D.; Greenblatt, H. M.; Pandey, B. K.; Smith, M. S.; Paxman, R.; Torgerson, C. D.; Merrell, J. P.; Ritz, C. C.; Prigozhin, M. B.; Levy, Y.; Price, J. L. Criteria for Selecting PEGylation Sites on Proteins for Higher Thermodynamic and Proteolytic Stability. *J Am Chem Soc* **2014**, *136* (50). https://doi.org/10.1021/ja5095183.

(26) Pai, S. S.; Hammouda, B.; Hong, K.; Pozzo, D. C.; Przybycien, T. M.; Tilton, R. D. The Conformation of the Poly(Ethylene Glycol) Chain in Mono-PEGylated Lysozyme and Mono-PEGylated Human Growth Hormone. *Bioconjug Chem* **2011**, *22* (11), 2317–2323. https://doi.org/10.1021/bc2003583.

(27) Le Cœur, C.; Combet, S.; Carrot, G.; Busch, P.; Teixeira, J.; Longeville, S. Conformation of the Poly(Ethylene Glycol) Chains in DiPEGylated Hemoglobin Specifically Probed by SANS: Correlation with PEG Length and in Vivo Efficiency. *Langmuir* **2015**, *31* (30), 8402–8410. https://doi.org/10.1021/acs.langmuir.5b01121.

(28) Shu, J. Y.; Lund, R.; Xu, T. Solution Structural Characterization of Coiled-Coil Peptide-Polymer Side-Conjugates. *Biomacromolecules* **2012**, *13* (6), 1945–1955. https://doi.org/10.1021/bm300561y.

(29) He, L.; Wang, H.; Garamus, V. M.; Hanley, T.; Lensch, M.; Gabius, H. J.; Fee, C. J.; Middelberg, A. Analysis of MonoPEGylated Human Galectin-2 by Small-Angle X-Ray and Neutron Scattering: Concentration Dependence of PEG Conformation in the Conjugate. *Biomacromolecules* **2010**, *11* (12), 3504–3510. https://doi.org/10.1021/bm100999a.

(30) Ferebee, R.; Hakem, I. F.; Koch, A.; Chen, M.; Wu, Y.; Loh, D.; Wilson, D. C.; Poole, J. L.; Walker, J. P.; Fytas, G.; Bockstaller, M. R. Light Scattering Analysis of Mono- and Multi-PEGylated Bovine

Serum Albumin in Solution: Role of Composition on Structure and Interactions. *Journal of Physical Chemistry B* **2016**, *120* (20), 4591–4599. https://doi.org/10.1021/acs.jpcb.6b03097.

(31) Dhalluin, C.; Ross, A.; Leuthold, L.-A.; Foser, S.; Gsell, B.; Mü, F.; Senn, H.; Hoffmann-La, F. ARTICLES Structural and Biophysical Characterization of the 40 KDa PEG-Interferon-r 2a and Its Individual Positional Isomers. **2005**. https://doi.org/10.1021/bc049781.

(32) Farhadi, S. A.; Bracho-Sanchez, E.; Fettis, M. M.; Seroski, D. T.; Freeman, S. L.; Restuccia, A.; Keselowsky, B. G.; Hudalla, G. A. Locally Anchoring Enzymes to Tissues via Extracellular Glycan Recognition. *Nat Commun* **2018**, *9* (1). https://doi.org/10.1038/s41467-018-07129-6.

(33) John, C. M.; Leffler, H.; Kahl-Knutsson, B.; Svensson, I.; Jarvis, G. A. Truncated Galectin-3 Inhibits Tumor Growth and Metastasis in Orthotopic Nude Mouse Model of Human Breast Cancer. *Clinical Cancer Research* **2003**, *9* (6).

(34) Mirandola, L.; Yu, Y.; Chui, K.; Jenkins, M. R.; Cobos, E.; John, C. M.; Chiriva-Internati, M. Galectin-3C Inhibits Tumor Growth and Increases the Anticancer Activity of Bortezomib in a

Murine Model of Human Multiple Myeloma. *PLoS One* **2011**, *6* (7). https://doi.org/10.1371/journal.pone.0021811.

(35) Baiz, C. R.; Peng, C. S.; Reppert, M. E.; Jones, K. C.; Tokmakoff, A. Coherent Two-Dimensional Infrared Spectroscopy: Quantitative Analysis of Protein Secondary Structure in Solution. *Analyst* **2012**, *137* (8), 1793–1799. https://doi.org/10.1039/c2an16031e.

(36) Munasinghe, A.; Mathavan, A.; Mathavan, A.; Lin, P.; Colina, C. M. Molecular Insight into the ProteinPolymer Interactions in N-Terminal PEGylated Bovine Serum Albumin. *Journal of Physical Chemistry B* **2019**, *123* (25), 5196–5205. https://doi.org/10.1021/acs.jpcb.8b12268.

(37) Munasinghe, A.; Mathavan, A.; Mathavan, A.; Lin, P.; Colina, C. M. PEGylation within a Confined Hydrophobic Cavity of a Protein. *Physical Chemistry Chemical Physics* **2019**, *21* (46), 25584–25596. https://doi.org/10.1039/c9cp04387j.

(38) Yang, C.; Lu, D.; Liu, Z. How PEGylation Enhances the Stability and Potency of Insulin: A Molecular Dynamics Simulation. *Biochemistry* **2011**, *50* (13), 2585–2593. https://doi.org/10.1021/bi101926u.

(39) Aramini, J. M.; Tubbs, J. L.; Kanugula, S.; Rossi, P.; Ertekin, A.; Maglaqui, M.; Hamilton, K.; Ciccosanti, C. T.; Jiang, M.; Xiao, R.; Soong, T. T.; Rost, B.; Acton, T. B.; Everett, J. K.; Pegg, A. E.; Tainer, J. A.; Montelione, G. T. Structural Basis of O6-Alkylguanine Recognition by a Bacterial Alkyltransferase-like DNA Repair Protein. *Journal of Biological Chemistry* **2010**, *285* (18), 13736–13741. https://doi.org/10.1074/jbc.M109.093591.

(40) Rossi, P.; Swapna, G. V. T.; Huang, Y. J.; Aramini, J. M.; Anklin, C.; Conover, K.; Hamilton, K.; Xiao, R.; Acton, T. B.; Ertekin, A.; Everett, J. K.; Montelione, G. T. A Microscale Protein NMR

Sample Screening Pipeline. *Journal of Biomolecular NMR*. January 2010, pp 11–22. https://doi.org/10.1007/s10858-009-9386-z.

(41) Ghosh, A.; Ostrander, J. S.; Zanni, M. T. Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chemical Reviews*. 2017. https://doi.org/10.1021/acs.chemrev.6b00582.

(42) Valentine, M. L.; Al-Mualem, Z. A.; Baiz, C. R. Pump Slice Amplitudes: A Simple and Robust Method for Connecting Two-Dimensional Infrared and Fourier Transform Infrared Spectra. *Journal of Physical Chemistry A* **2021**, *125* (29). https://doi.org/10.1021/acs.jpca.1c04558.

(43) Chen, X.; Roeters, S. J.; Cavanna, F.; Alvarado, J.; Baiz, C. R. Crowding Alters F-Actin Secondary Structure and Hydration. *Commun Biol* **2023**, *6* (1). https://doi.org/10.1038/s42003-023-05274-3.

(44) Kwak, K.; Rosenfeld, D. E.; Fayer, M. D. Taking Apart the Two-Dimensional Infrared Vibrational Echo Spectra: More Information and Elimination of Distortions. *Journal of Chemical Physics* **2008**, *128* (20). https://doi.org/10.1063/1.2927906.

(45) Fayer, M. D. Dynamics of Liquids, Molecules, and Proteins Measured with Ultrafast 2D IR Vibrational Echo Chemical Exchange Spectroscopy. *Annual Review of Physical Chemistry*. Annual Reviews Inc. 2009, pp 21–38. https://doi.org/10.1146/annurev-physchem-073108-112712.

(46) Park, S.; Kwak, K.; Fayer, M. D. Ultrafast 2D-IR Vibrational Echo Spectroscopy: A Probe of Molecular Dynamics. *Laser Physics Letters*. October 2007, pp 704–718. https://doi.org/10.1002/lapl.200710046.

(47) La Cour Jansen, T.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. Modeling the Amide I Bands of Small Peptides. *Journal of Chemical Physics* **2006**, *125* (4). https://doi.org/10.1063/1.2218516.

(48) Hamm, P.; Woutersen, S. *Coupling of the Amide Modes of the Glycine Dipeptide Coupling of the Amide Modes P. Hamm et Al*; 2002; Vol. 75.

(49) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. Ultrafast Vibrational Spectroscopy of Water and Aqueous N-Methylacetamide: Comparison of Different Electronic Structure/Molecular Dynamics Approaches. *Journal of Chemical Physics* **2004**, *121* (18), 8887–8896. https://doi.org/10.1063/1.1791632.

(50) Eaves, J. D.; Tokmakoff, A.; Geissler, P. L. Electric Field Fluctuations Drive Vibrational Dephasing in Water. *Journal of Physical Chemistry A* **2005**, *109* (42), 9424–9436. https://doi.org/10.1021/jp051364m.

(51) Stevenson, P.; Tokmakoff, A. Ultrafast Fluctuations of High Amplitude Electric Fields in Lipid Membranes. *J Am Chem Soc* **2017**, *139* (13), 4743–4752. https://doi.org/10.1021/jacs.6b12412.

(52) Pritzlaff, A.; Ferré, G.; Mulry, E.; Lin, L.; Gopal Pour, N.; Savin, D. A.; Harris, M. E.; Eddy, M. T. Atomic-Scale View of Protein-PEG Interactions That Redirect the Thermal Unfolding Pathway of PEGylated Human Galectin-3. *Angewandte Chemie - International Edition* **2022**, *61* (40). https://doi.org/10.1002/anie.202203784.

(53) Rodríguez-Martínez, J. A.; Solá, R. J.; Castillo, B.; Cintrón-Colón, H. R.; Rivera-Rivera, I.; Barletta, G.; Griebenow, K. Stabilization of α-Chymotrypsin upon PEGylation Correlates with Reduced Structural Dynamics. *Biotechnol Bioeng* **2008**, *101* (6). https://doi.org/10.1002/bit.22014.