# Specific and Nonspecific Effects of Glycosylation

Christopher R. Ellis, Buddhadev Maiti, and William G. Noid*

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

S *Supporting Information*

ABSTRACT: Glycosylation regulates vital cellular processes and dramatically influences protein folding and stability. In particular, experiments have demonstrated that asparagine (N)-linked disaccharides drive a "conformational switch" in a model peptide. The present work investigates this conformational switch via extensive atomically detailed replica exchange molecular dynamics simulations in explicit solvent. To distinguish the effects of specific and nonspecific interactions upon the peptide conformational ensemble, these simulations considered model peptides that were N-linked to a disaccharide and to a steric crowder of the same shape. The simulations are remarkably consistent with experiment and provide detailed insight into the peptide structure ensemble. They suggest that steric crowding by N-linked disaccharides excludes extended conformations, but does not significantly impact the tetrahedral structure of the surrounding solvent or otherwise alter the peptide free energy surface. However, the combination of steric crowding with specific hydrogen bonds and hydrophobic stacking interactions more dramatically impacts the peptide ensemble and stabilizes new structures.
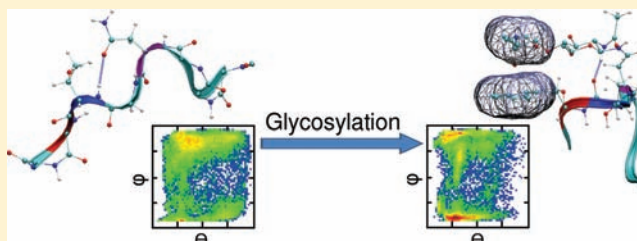
INTRODUCTION

N-linked glycosylation, which refers to the covalent linkage of a carbohydrate to an asparagine side chain,[1] dramatically impacts protein folding, function, and stability.[2−4] N-linked glycans direct intra- and extracellular trafficking,[5] mediate cell recognition,[6] hinder aggregation,[7] and regulate chaperone-assisted protein folding[8] and degradation.[9] Moreover, direct interactions with N-linked glycans also influence protein folding and stability. In the case of short peptides, carbohydrates can stabilize secondary structures[10−12] or compact conformations.[13−15] In the case of larger glycoproteins, glycans typically influence local conformational preferences and dampen fluctuations.[16] In some cases, carbohydrates influence protein folding without impacting stability.[17,18] In other cases, carbohydrates can either stabilize[19−24] or destabilize[25−27] folded proteins. Bioinformatic estimates suggest that more than half of eukaryotic proteins may be glycoproteins, the large majority of which are N-linked glycoproteins.[28]

These considerations have motivated tremendous interest in protein−carbohydrate interactions and, in particular, deciphering the "glycosylation code" that determines their effect.[26,29] A complete understanding of these interactions would provide considerable insight into protein biophysics and many cellular processes. Moreover, this glycosylation code holds vast practical importance for developing improved pharmaceuticals[30,31] and for engineering novel protein structures.[32,33]

However, these interactions are complex and may proceed through various mechanisms. For instance, carbohydrates may exert nonspecific forces upon proteins by disrupting the local solvation structure[34] or by entropically destabilizing the unfolded ensemble.[23,35,36] In particular, computational studies with low-resolution models[29,37] have indicated that steric

crowding by N-linked glycans stabilizes several glycoproteins.[23,24] Subsequent studies, though, indicated that additional interactions are necessary to rationalize the effects of glycosylation upon other proteins.[27]

Alternatively, specific interactions with N-linked glycans may directly stabilize the native state of glycoproteins.[38,39] For instance, specific hydrogen bonds and van der Waals interactions stabilize the adhesion domain of human CD2.[19] Bioinformatic[40] and experimental[41−43] studies have indicated that direct, specific hydrophobic interactions may impact protein folding and stability. Recent studies have also emphasized the significance of strong stacking interactions between carbohydrates and aromatic π systems in molecular recognition.[44−47]

Despite considerable progress in rationalizing[26] and even manipulating these interactions[32,33] in particular glycoproteins, carbohydrate−protein interactions remain incompletely understood and difficult to predict.[27] Motivated by these considerations, the present study investigates and distinguishes the effects of specific and nonspecific interactions upon the conformational ensemble of a short peptide. In particular, this work reports extensive, atomically detailed replica exchange molecular dynamics (MD) simulations in explicit solvent for an octapeptide, the corresponding N-linked glycopeptide, and also for a computational model of a virtual glycopeptide, in which the carbohydrate exerts only nonspecific steric forces upon the peptide.

Figure 1 presents the chemical structure for the simulated glycopeptide systems. The glycan corresponds to the core
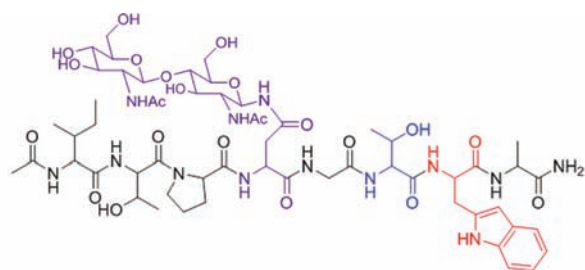
Ace1 – Ile2 – Thr3 – Pro4 – Asn5(GlcNAC)$_2$ – Gly6 – Thr7 – Trp8 – Ala9 – NH$_2$

Figure 1. Chemical structure and sequence of the glycopeptide.

disaccharide (chitobiose) of a larger carbohydrate (Glc$_3$Man$_9$GlcNac$_2$) that is covalently attached to N-linked glycoproteins as they emerge from membrane-bound ribosomes.[1,4] Several studies have indicated that this core disaccharide exerts the greatest influence upon protein folding and structure.[19,48,49] This model peptide reflects several common features of N-glycosylated sequences, including the strict consensus sequence (Asn-Xxx-Ser/Thr), a small, flexible residue (Gly6) immediately following the glycosylation site, and a bulky hydrophobic residue (Trp8) two residues later.[40] Furthermore, this peptide is a close analogue to one of the glycosylated sequences that adopts a β-turn in the influenza binding protein, hemagglutinin (HA).[50] In the absence of glycosylation, HA misfolds and demonstrates enhanced rates of aggregation.[8] Moreover, these N-linked glycans bind sialic acid receptors of the host cell and initiate pathogen entry.[51]

The work of Imperiali and co-workers provides an extensive body of experimental results for a model glycopeptide that differs from Figure 1 only in the addition of an ornithine residue before Ile2.[11,14,52,53] These landmark studies clearly demonstrated the dramatic influence of glycosylation upon short peptides. In particular, they demonstrated that an N-linked chitobiose glycan drove a "conformational switch" from an extended Asx-turn to a more compact β-turn.[11,53]

Our simulations are highly consistent with these experiments. In accord with experimental observations, the simulations demonstrate that the N-linked disaccharide destabilizes extended conformations, such as Asx-turns, in favor of more compact conformations, such as β-turns, via a sequence-dependent mechanism[14] that reflects contributions from the glycan N-acetyl groups, the distal glucosamine, and also hydrophobic forces.[52] By characterizing the conformational ensembles for the octapeptide, glycopeptide, and steric crowder models, our simulations distinguish the effects of nonspecific steric and solvation forces from the effects of direct attractive interactions with the glycan. Our simulations suggest that, although the excluded volume of the disaccharide destabilizes extended conformations, nonspecific steric forces are insufficient to stabilize any particular well-defined structure for the octapeptide. However, these nonspecific steric forces may couple with specific aromatic stacking and hydrogen-bonding interactions to channel unfolded peptides toward new regions of configuration space and to stabilize novel structures.

METHODS

Simulations. We report extensive MD simulations for three systems described by Figure 1: (1) a nonglycosylated octapeptide; (2) the corresponding N-linked glycopeptide; and (3) a computational model of this glycopeptide in which all nonbonded peptide–carbohydrate interactions were modeled with purely repulsive potentials. In the following, we refer to these three simulations as

the octapeptide, glycopeptide, and excluded-volume simulations, respectively. Each simulation explicitly modeled solvent with the SPC/E water potential[54] and the peptide with the OPLS-AA force field.[55] In the glycopeptide simulation, the OPLS-AA force field[56] (with standard mixing rules) was used to describe all carbohydrate interactions, with the exception of the Asn5–carbohydrate linkage, which required an addition 17 bonded parameters that were determined from chemically similar interactions. For the excluded-volume simulation, we made two modifications to the carbohydrate nonbonded potential: (1) all peptide–carbohydrate van der Waals interactions were replaced with purely repulsive potentials that were optimized to describe the repulsive part of the corresponding Lennard-Jones potential; and (2) all carbohydrate charges were neutralized. This charge neutralization eliminated the solvent–carbohydrate electrostatic interactions and disrupted the water structure immediately solvating the carbohydrate in the excluded-volume simulation. We emphasize that the solvent–carbohydrate van der Waals interactions were treated identically in the glycopeptide and excluded-volume simulations.

All MD simulations were performed with the Gromacs 4.5.3 software package[57,58] following standard protocols.[59−63] For each system, an initial configuration was obtained from the PRODRG server.[64] Upon equilibration, this configuration was replicated 63 times and the replicas were annealed to target temperatures between 298 and 515 K.[65] During the ensuing simulations, each carbohydrate remained in a stable $^4C_1$ chair conformation, while the anomeric and glycosidic linkages both sampled expected conformations.[66,67]

Replica exchange MD (REMD) simulations[68] were then performed with 64 replicas of each system for a duration of 110 ns. The last 100 ns of these simulations were analyzed in the following calculations. Exchanges were attempted every 1 ps[69] and the corresponding acceptance probability was 38–42% over the entire temperature range. The transit time for a replica to move from the lowest temperature to the highest temperature and return was roughly 15 ns. These REMD simulations included approximately 7500 total atoms and required 350 000 CPU-hours to generate a combined total of 19.2 μs for analysis.

Calculations. Principal component analysis was performed for the (nonglycosylated) octapeptide simulation using backbone coordinates sampled by the room temperature replica.[70] The first two eigenvectors of the resulting covariance matrix corresponded to hinging (r^2 = 0.998) and twisting (r^2 = 0.967) motions, respectively, which were used to characterize the global free energy surface for each system. The weighted histogram analysis method[71−73] was employed to calculate probability distributions.

Hydrophobic solvent accessible surface area was calculated according to the double cubic lattice method.[74] NMR J-couplings were calculated according to the Karplus curve.[75] The tetrahedral order parameter, q_{tet}, was calculated to characterize the local structure of water.[76] The tetrahedral solvation structure for water molecule k is characterized by $q_k = 1 - 3/8\sum_i\sum_{j>i}(\cos \psi_{ikj} + 1/3)^2$, where the double sum is performed over the vertices of a tetrahedron defined by the nearest hydrophilic peptide atoms or water molecules and ψ_{ikj} is the angle defined by molecule k and the vertices i and j. The tetrahedral order parameter, q_{tet}, is defined by an average over water molecules. The first solvation shell was defined by water molecules with oxygen atoms within 0.35 nm of heavy peptide atoms. Hydrophobic and hydrophilic atoms were distinguished based upon a net charge of less than or greater than 0.3 e, respectively.

The Supporting Information section provides detailed information regarding the simulated potentials, equilibration procedures, convergence, and free energy calculations.

RESULTS

The present work reports explicit solvent REMD simulations for three systems: (1) an octapeptide; (2) the N-glycosylated variant of the octapeptide; and (3) a computational model for this glycopeptide with purely repulsive carbohydrate–peptide interactions. In the following, these systems will be referred to
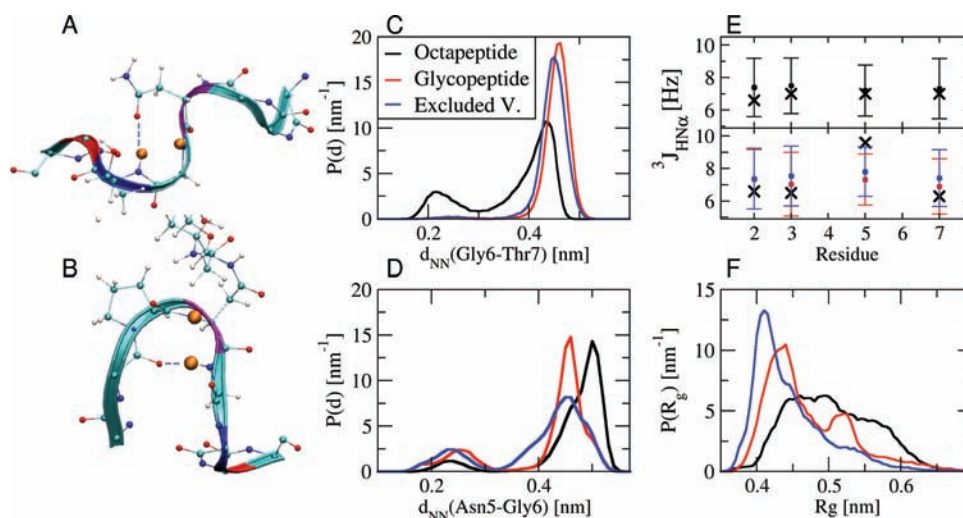
Figure 2. Comparison with experimental results. Panels A and B present Asx-turn and β-turn configurations sampled from simulations of the octapeptide and glycopeptide, respectively. Panels C and D present distance distributions for the Gly6-Thr7 amide proton pair and for the Asn5-Gly6 amide proton pair, respectively. Panel E presents experimental (X) and simulated J-couplings with standard deviations indicated by uncertainty bars. Panel F presents the simulated radius of gyration for the peptide backbone. In each panel, black, red, and blue curves indicate results for the octapeptide, glycopeptide, and excluded-volume simulations, respectively. This figure and Figures 4, 5, and 6 were made with VMD.[94]

as the octapeptide, glycopeptide, and excluded-volume systems, respectively. Figure 1 presents the sequence and chemical structure for the glycopeptide. In Figure 1 and in subsequent figures, Asn5, Thr7, and Trp8 are colored purple, blue, and red, respectively.

Comparison with Prior Observations. Figure 2 demonstrates that the present REMD simulations are remarkably consistent with the experimental observations of Imperiali and co-workers.[11,14,52,53] Panels A and B of Figure 2 present configurations sampled from REMD simulations of the octapeptide and glycopeptide, respectively. Figure 2A demonstrates that the simulated octapeptide did indeed adopt Asx-turn configurations that were stabilized by a hydrogen bond between the Asn5 side chain and the Thr7 carbonyl. Figure 2B demonstrates that the simulated glycopeptide did indeed sample β-turn configurations that were stabilized by a hydrogen bond between the Thr3 backbone carbonyl and the Gly6 amide proton. The octapeptide simulations never sampled a β-turn configuration, while the glycopeptide simulations sampled the Asx-turn in only 15 out of 100 000 configurations. These observations are all in perfect agreement with previous experimental conclusions.[52] However, as discussed further below, both the octapeptide and glycopeptide simulations sampled a diverse ensemble of conformations.

Panels C and D of Figure 2 present distributions of interproton distances corresponding to key NOE observations.[11] Figure 2C corresponds to the Gly6 and Thr7 amide protons (highlighted in Figure 2A), which were used to experimentally identify Asx-turn configurations. In agreement with the observed decrease in NOE intensity upon glycosylation,[11] this proton pair samples very short separations in the octapeptide simulations, but samples significantly larger distances in the glycopeptide simulations. Figure 2D corresponds to the Asn5 and Gly6 amide protons (highlighted in Figure 2B), which were used to experimentally identify β-turn configurations. In agreement with the observed increase in NOE intensity upon glycosylation, this proton pair samples shorter distances more frequently in the glycopeptide simulations.

The Supporting Information section presents results for the other two reported NOE signals.[11] The REMD simulations are consistent with NOE observations for the α carbon proton of Pro4 and the amide proton of Asn5. However, the simulations are somewhat inconsistent with NOE observations for the α carbon proton of Asn5 and the amide proton of Gly6. While experiments observed strong NOE signals for this pair that became weaker upon glycosylation, this proton pair remains in very close proximity (within 4 Å of one another) in all simulations. The Supporting Information section more quantitatively analyzes the simulated interproton distributions.

Significantly, the observed NOE changes do not distinguish between specific and nonspecific interactions. Both the glycopeptide and excluded-volume simulations sampled very similar distributions for each pair of backbone protons.

Imperiali and co-workers reported J-coupling constants and variable temperature (VT) coefficients for the octapeptide and glycopeptide systems.[11] The top panel of Figure 2E demonstrates that the REMD simulations reproduce the experimental J-couplings for the octapeptide within the relatively large standard deviation of the simulated values. The bottom panel of Figure 2E demonstrates that the REMD simulations of the glycopeptide reproduce the experimental J-couplings for most residues with reasonable accuracy. The significant exception is for Asn5, for which the experimentally observed J-coupling is considerably larger than the calculated J-coupling. The Supporting Information section demonstrates that glycosylation-induced changes in VT coefficients are also highly consistent with hydrogen bonding changes that are observed in simulations.

Finally, Figure 2F presents simulated distributions for the radius of gyration of the peptide backbone. In perfect agreement with fluorescence energy transfer (FET) studies, the N-linked disaccharide significantly reduces the population of extended conformations sampled by the glycopeptide.[14] Interestingly, the elimination of attractive carbohydrate-peptide interactions results in an even more compact ensemble.

Local Conformational Preferences. Figure 3 presents Ramachandran maps that demonstrate the impact of specific
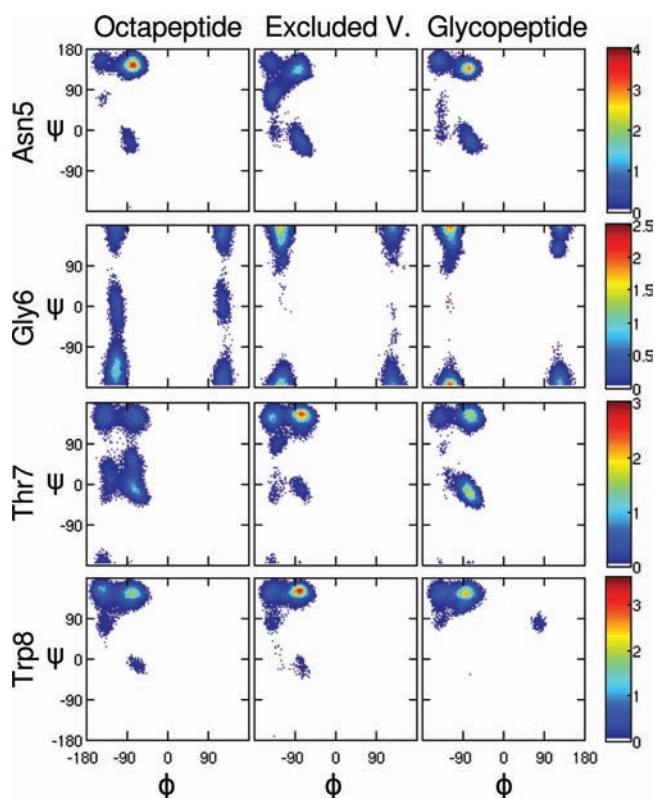
Figure 3. Ramachandran maps for Asn5, Gly6, Thr7, and Trp8. The left, center, and right columns correspond to the octapeptide, excluded-volume, and glycopeptide simulations. In this figure and Figure 5, the intensity of Ramachandran plots are presented in units of 10^{-3} deg^{-2}.

and nonspecific carbohydrate–peptide interactions upon the conformational preferences of Asn5, Gly6, Thr7, and Trp8. As expected, the effect of the carbohydrate is relatively local. The carbohydrate demonstrates little influence upon the conformational preferences of Ile2, Thr3, Pro4, or Ala9.

The carbohydrate generally exerts greater influence upon ψ than ϕ. In the case of Asn5, the backbone adopts predominantly extended conformations prior to glycosylation, but interactions with the carbohydrate slightly stabilize more helical configurations. The carbohydrate most dramatically impacts the conformation of Gly6. Prior to glycosylation, Gly6 sampled conformations corresponding to $\psi \approx -150$ and $0°$ with similar probability. However, the carbohydrate completely destabilizes the equilibrium state at $\psi \approx 0°$ and shifts the remaining equilibrium state to even more extended states for which $\psi \approx \pm 180°$. Notably, the most significant changes to Asn5 and Gly6 conformation result from nonspecific steric crowding.

The carbohydrate demonstrates more subtle effects upon the backbone of Thr7 and Trp8. In the case of Thr7, steric crowding stabilizes more extended backbone conformations and destabilizes helical conformations, but attractive interactions with the carbohydrate also stabilize helical conformations. In the case of Trp8, steric crowding again stabilizes more extended backbone conformations. However, attractive interactions with the carbohydrate stabilize a new left-handed helical conformation at $\phi \approx \psi \approx 90°$.

Global Free Energy Surfaces. To characterize global features of the ensemble, we performed principal component analysis for the octapeptide simulation. The first two principal

components closely correlated with the hinging ($r^2 = 0.998$) and twisting ($r^2 = 0.967$) motions illustrated in Figure 4, panels
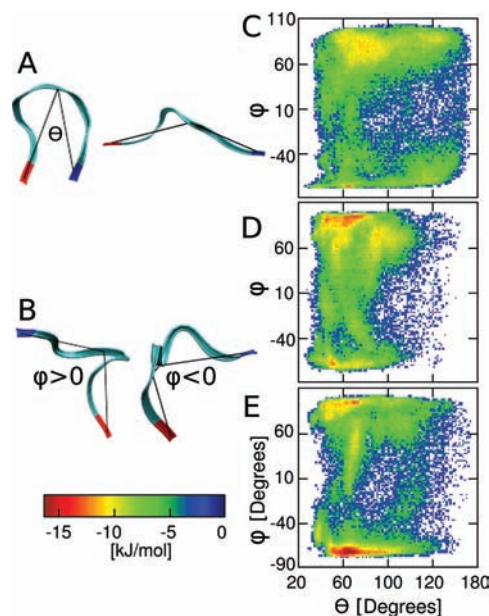

Figure 4. Global free energy surfaces. Panels A and B describe the hinging, θ, and twisting, φ, angles that correspond to the first two principal components of the octapeptide covariance matrix. (To assist in visualizing these angles, Ile2 and Ala9 are colored red and blue, respectively.) Panels C–E describe the free energy surfaces of the octapeptide (C), excluded-volume (D), and glycopeptide simulations (E), as a function of θ and φ.

A and B, respectively. The hinging angle, θ, which is defined by the α-carbons of Ile2, Gly6, and Ala9, also closely correlates with the peptide end-to-end distance and radius of gyration. The twisting angle, φ, is defined by a pseudodihedral angle formed by the α-carbons of Ile2, Pro4, Gly6, and Ala9.

Figure 4C presents a 2D intensity plot of the free energy surface (FES) for the (nonglycosylated) octapeptide at 298 K as a function of θ and φ. The octapeptide FES demonstrates two distinct basins with opposite peptide twists of $\varphi > +45°$ and $\varphi < -45°$, that are sampled by approximately 57% and 20% of simulated configurations, respectively. The dominant basin ($\varphi > 45°$) features a broad, shallow minima at $\theta \approx 80°$, $\varphi \approx 75°$, but also contains more compact, as well as highly extended, conformations. Consequently, the FES reveals that the simulated octapeptide samples a highly heterogeneous ensemble of disordered conformations. Within this ensemble, the peptide adopts an Asx-turn in approximately 2.2% of sampled configurations.

Panels D and E of Figure 4 present corresponding 2D intensity plots of the FES for the excluded-volume and glycopeptide systems, respectively. The carbohydrate significantly reduces the conformational space accessible to the peptide and eliminates extended conformations from the peptide ensemble. Prior to glycosylation, the octapeptide adopted configurations with a hinging angle $\theta \geq 125°$ in more than 22% of sampled configurations. However, these extended configurations are adopted in only 2% and 7% of configurations sampled by the excluded-volume and glycopeptide simulations, respectively. In contrast, more compact turn conformations (including i to $i + 3$, i to $i + 4$, and i to $i + 5$ turns) are adopted in more than 30% and 12% of
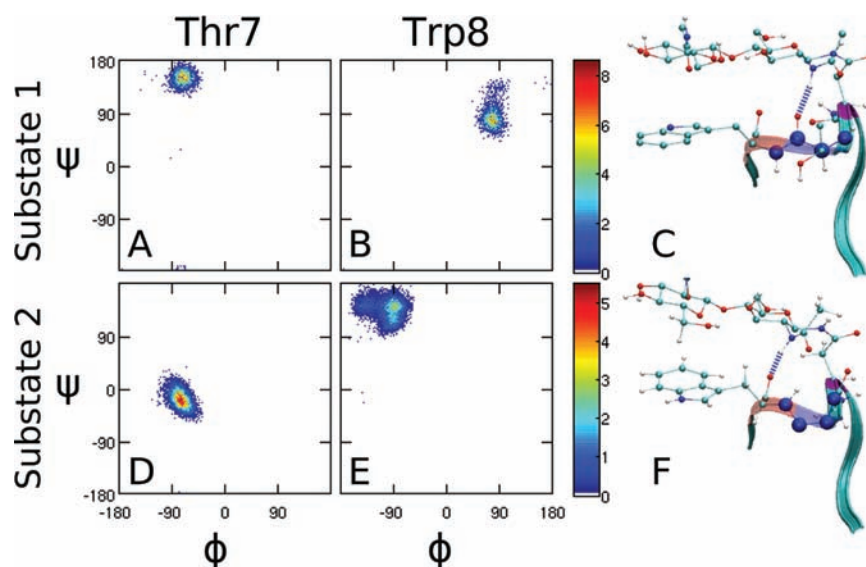
Figure 5. Characterization of the glycopeptide NSC substates 1 and 2 in rows 1 and 2, respectively. Columns 1 and 2 present the corresponding Ramachandran maps sampled by Thr7 and Trp8, respectively. Panels C and F present configurations from NSC substates 1 and 2, respectively with Asn5, Thr7, and Trp8 highlighted by purple, blue, and red, respectively. The blue spheres highlight, for Thr7, the switch from an extended to a helical conformation.

configurations sampled by the excluded-volume and glycopeptide simulations, respectively.

Panels C and D of Figure 4 also demonstrate more subtle differences between the FES for the octapeptide and excluded-volume simulations. In particular, the FES minima slightly deepens and extends toward more compact conformations in the excluded-volume simulation. In addition, the shift in Gly6 torsional preferences divides this minima into separate regions. Nevertheless, panels C and D of Figure 4 demonstrate that the octapeptide and excluded-volume simulations sample very similar regions of configuration space. The excluded-volume simulation sampled the $\varphi > +45°$ and $\varphi < -45°$ basins in 59% and 16% of configurations, respectively, which almost perfectly coincides with the sampling in the octapeptide simulations. Most significantly, the peptide remains disordered in the excluded-volume simulations and does not adopt any stable conformation.

Figure 4E demonstrates that attractive interactions with the carbohydrate much more significantly alter the glycopeptide FES. In particular, although the glycopeptide still samples a fairly diverse ensemble, attractive glycan−peptide interactions stabilize a new native state cluster (NSC). This NSC demonstrates a different twist ($\varphi < -45°$) relative to the dominant configurations sampled by the octapeptide or excluded-volume simulations ($\varphi > +45°$). This NSC corresponds to a much deeper and better defined minima in the glycopeptide FES that is centered at $\theta = 65°$, $\varphi = -71°$. While this region accounts for only 5% and 8% of configurations in the octapeptide and excluded-volume simulations, respectively, this minima accounts for 33% of glycopeptide configurations.

Hydrogen Bonding Stabilizes the NSC. The glycopeptide NSC identified in Figure 4 can be decomposed into three distinct substates that are distinguished by the backbone conformations of Asn5, Thr7, and Trp8. Figure 5 presents Ramachandran plots for Thr7 and Trp8, while using configurations sampled from substates 1 and 2 of the glycopeptide NSC. In substate 1, which accounts for 10% of the NSC, Thr7 is extended and Trp8 adopts a left-handed helix.

As demonstrated in Figure 3, Trp8 only adopts a left-handed helix in the presence of attractive interactions with the carbohydrate and almost all (94%) of these configurations correspond to NSC substate 1. In substate 2, which accounts for 62% of the NSC, Thr7 is helical and Trp8 is extended.

As illustrated by Figure 5C,F and quantified in the Supporting Information section, these backbone conformations promote specific hydrogen bonds that stabilize the NSC. Figure 5C demonstrates that, in NSC substate 1, the extended conformations of Asn5, Gly6, and Thr7 promote stable hydrogen bonding between the Thr7 backbone and the N-acetyl group of the proximal glucosamine. Simultaneously, the left-handed helical conformation of Trp8 promotes stacking interactions between the indole side chain and the distal glucosamine. Figure 5F demonstrates that, in NSC substate 2, the helical conformation of Thr7 promotes hydrogen bonding of the Gly6 and Trp8 backbone with the N-acetyl group of the proximal glucosamine. At the same time, the extended conformation of Trp8 promotes stacking interactions with the distal glucosamine. In both NSC substates 1 and 2, the glycopeptide backbone samples helical configurations to optimize specific attractive interactions with the carbohydrate. In contrast, in the excluded volume system, the carbohydrate and peptide cannot form favorable contacts and both Thr7 and Trp8 remain extended.

The third substate (not shown) is somewhat different and corresponds to outer regions of the NSC. In this substate, both Trp7 and Thr8 adopt extended backbone conformations that promote backbone hydrogen bonding with the N-acetyl group of the proximal glucosamine, but that interact less favorably with the glycan. It should be noted that the compact glycopeptide NSC also reflects hydrogen bonding involving the C-terminal amide group.

The hydrogen bonding interactions are considerably different in the (nonglycosylated) octapeptide, excluded-volume, and glycopeptide simulations. Prior to glycosylation, the two most stable hydrogen bonds correspond to weak interactions of the Gly6 backbone carbonyl with the Thr3 hydroxyl side chain and

with the Trp8 backbone amide, both of which form in approximately 10% of configurations. Neither of these hydrogen bonds forms in the presence of the carbohydrate, which is consistent with the dramatic changes in the Gly6 backbone conformation. (See Figure 3.) The compact conformations sampled by the excluded-volume system are partially stabilized by hydrogen bonds between Thr3 and Thr7. In the presence of attractive interactions with the glycan, these hydrogen bonds are destabilized and replaced with the hydrogen bonds identified for the glycopeptide NSC.

Carbohydrate−Aromatic Interactions Stabilize the NSC. Figures 5 and 6 demonstrate that attractive stacking
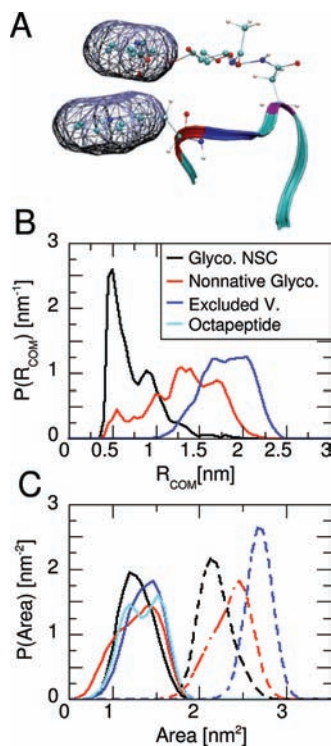

Figure 6. Stacking interactions between Trp8 and the distal saccharide. Panel A presents a configuration from glycopeptide NSC substate 1. Panel B presents distributions for the distance between the centers of mass (COM) for the distal glucosamine and Trp8 side chain indole ring. Panel C presents distributions of SASA for the carbohydrate (dashed curves) and for the Trp8 indole ring (solid curves). The black, red, and blue curves correspond to configurations sampled from the glycopeptide NSC, to glycopeptide nonnative configurations, and to configurations sampled from the excluded-volume simulations, respectively. The cyan curve corresponds to configurations sampled from the octapeptide simulation.

interactions between the distal glucosamine and the Trp8 indole side chain also contribute to stabilizing the glycopeptide NSC. Figure 6A indicates this interaction in a configuration sampled from NSC substate 1. Figure 6B presents distributions for the center of mass distance between the distal glucosamine and the indole ring of the Trp8 side chain. In the absence of attractive peptide−carbohydrate interactions, the glycan never approaches the Trp8 side chain. In the presence of attractive peptide−carbohydrate interactions, the two groups sample a large range of distances. In the glycopeptide NSC, however, the pair remains in close contact. Figure 6C demonstrates that this attractive interaction significantly reduces the hydrophobic solvent accessible surface area (SASA) of the carbohydrate

(dashed curves), which is most effectively buried in the glycopeptide NSC.

The effect upon the SASA for the Trp8 side chain (solid curves) is somewhat more subtle, though. The Trp8 indole ring is most exposed in the excluded-volume simulations and samples a wide range of environments in glycopeptide nonnative configurations. This indole ring samples a narrower and generally more buried set of environments in the glycopeptide NSC. The overlap between the distributions sampled in the glycopeptide NSC and nonnative configurations largely corresponds to configurations that are very near the NSC on the FES. The reduction in hydrophobic SASA for the simulated glycopeptide NSC is consistent with FET studies demonstrating that hydrophobic forces influence the glycopeptide collapse.[14]

Glycopeptide Solvation Structure. Previous studies have suggested that the presence of a bulky carbohydrate can indirectly impact peptide structure by disrupting the surrounding water structure.[34] The Supporting Information demonstrates that, despite altering the peptide structure ensemble, the disaccharide does not dramatically impact the density of water in the peptide's first solvation shell. Figure 7 presents
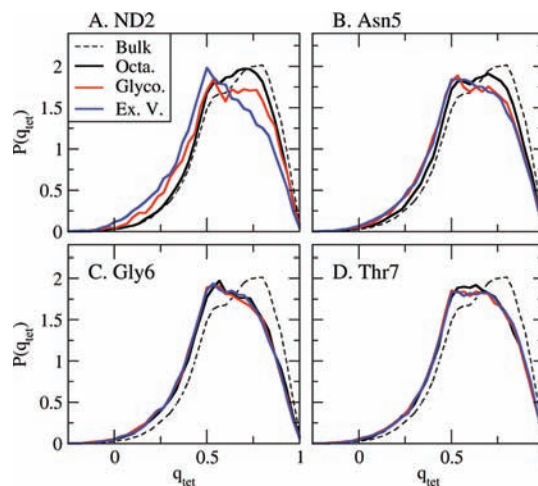

Figure 7. Distributions of the tetrahedral order parameter for water in the first solvation shell of the peptide. Panel A corresponds to the Asn5 side chain amide, while Panels B, C, and D correspond to the entire Asn5, Gly6, and Thr7 residues, respectively. The solid black, red, and blue curves correspond to simulations of the octapeptide, glycopeptide, and excluded-volume systems, respectively. The dashed black curve provides a comparison for bulk water.

distributions of the order parameter, q_{tet}, characterizing the tetrahedral structure of water in the first solvation shells of the Asn5 side chain amide (Figure 7A), as well as the entire Asn5 (Figure 7B), Gly6 (Figure 7C), and Thr7 (Figure 7D) residues. Prior to glycosylation, the tetrahedral structure of water solvating the Asn5 side chain amide is very similar to the structure of bulk water. The excluded volume of the peptide more significantly impacts the solvation structure of Gly6 and Thr7.

Figure 7 demonstrates that the disaccharide generates surprisingly little disruption to the tetrahedral structure of the peptide solvation shell. In the excluded-volume simulations, the glycan charge neutralization eliminates solvent−glycan hydrogen bonds, disrupting the structure of the first solvation shell for the glycan and for the Asn5 side chain amide, which is

immediately adjacent. However, the glycan−solvent electrostatic interactions restore the solvation structure around the Asn5 side chain to a state that is very similar to the solvation structure of Gly6 and Thr7 prior to glycosylation. Moreover, Figure 7B−D demonstrates that, in both the excluded-volume and glycopeptide simulations, the carbohydrate minimally impacts the solvation structure for the rest of Asn5 and has even less effect for Gly6, Thr7, or other residues.

■ DISCUSSION

Bioinformatic studies have estimated that more than half of all proteins are glycosylated and that most of these are N-linked glycoproteins.[28] N-linked glycosylation regulates many cellular processes[2,77] and impacts the efficacy of protein therapeutics.[30,31] N-linked glycans also directly impact protein folding, stability, and structure via both specific and nonspecific interactions. Low-resolution computational studies have clearly demonstrated the significance of nonspecific steric interactions for the folding and thermodynamics of model proteins.[37,78−81] However, excluded volume arguments cannot explain the effects of glycosylation upon other proteins.[27] Despite considerable success in rationalizing[26] and even manipulating[32] glycan−protein interactions in specific cases, it remains difficult to predict the effects of these interactions.[33]

Motivated by these considerations, we have investigated the impact of specific and nonspecific interactions with N-linked glycans upon the conformational ensemble of a short peptide. Our investigations employed extensive atomically detailed simulations in explicit solvent, while using enhanced sampling techniques[68] to adequately characterize the peptide free energy surface (FES). In particular, we simulated three systems: an octapeptide, the corresponding N-glycopeptide, and a virtual model of the same glycopeptide, in which the glycan exerts only nonspecific steric forces upon the peptide. The simulated glycan (chitobiose) corresponds to the core disaccharide that is covalently attached during cotranslational glycosylation and that exerts the greatest influence upon protein folding and stability.[19,48,49] This glycopeptide system closely corresponds to the system studied in the landmark experiments of Imperiali and co-workers[11,14,52,53] and reflects many statistically significant features of glycosylated sequences.[40]

The studies of Imperiali and co-workers provide an extensive body of experimental results. NOE signals corresponding to the amide protons of Gly6 and Thr7 indicated that, prior to glycosylation, the peptide adopted an Asx-turn conformation.[11] Fluorescence studies indicated that the N-linked glycan destabilized extended conformations in favor of more compact conformations.[14] Moreover, NMR studies for the glycopeptide did not observe an NOE signal for the Gly6-Thr7 amide protons, but instead observed (1) an enhanced NOE signal for the Asn5-Gly6 amide protons, (2) a very high $^3J_{HN\alpha}$-coupling for the Asn5 amide, and (3) a reduced variable temperature coefficient for Gly6.[11] These observations suggested that the glycopeptide no longer sampled Asx-turn conformations, but instead adopted a β-turn conformation.[11] Short MD simulations provided further evidence for the stability of this β-turn.[53]

These studies also provided important insight into the mechanism underlying these conformational changes. Fluorescence studies revealed that the glycan-induced collapse was sensitive to peptide sequence and was mitigated in a less polar water/acetonitrile solvent, which suggested an important role for hydrophobic effects.[14] However, fluorescence energy transfer (FET) studies indicated that the glycan did not significantly impact the observed fluorescence decay for Trp8,[14] while NMR studies observed only three weak NOE signals for interactions between the Trp8 side chain and the glycan.[11] These observations suggested that Trp8 did not form strong specific interactions with the carbohydrate. Subsequent NMR studies[52] indicated that the observed conformational changes were sensitive to (1) the presence of both glucosamines in the disaccharide; (2) the presence of N-acetyl groups on both glucosamines; and (3) the structure of the distal sugar. The lack of specific observables relating the glycan and peptide structures motivated the conclusion that nonspecific effects and, especially, steric crowding dominated the peptide−carbohydrate interactions.[52]

Our MD simulations are remarkably consistent with these experimental observations. With the exception of a relatively low simulated value for the J-coupling of the glycosylated Asn5, the MD simulations are consistent with all of the preceding quantitative observations and with the large majority of additional reported results. In accord with experimental FET studies,[14] our simulations indicate that the carbohydrate destabilizes extended conformations in favor of compact conformations that are stabilized by sequence-dependent interactions and that are sensitive to hydrophobic effects. The MD simulations are highly consistent with the reported glycan-induced changes in NOE intensities and, in particular, are in perfect accord with the changes noted above. The simulations are consistent with the experimental observation that the proximal N-acetyl group and the distal glucosamine both significantly impact peptide conformation. The simulations are consistent with the peptide sampling an Asx-turn prior to, but not after, glycosylation. Furthermore, as proposed by Imperiali and co-workers, our simulations indicate that turn configurations and, in particular, β-turns become more stable upon glycosylation.

The present simulations provide exquisite insight into the structure ensemble sampled by the peptide model before and after glycosylation. Prior to glycosylation, the simulated octapeptide samples a highly heterogeneous distribution of disordered structures, including Asx-turns. However, the most stable region of the octapeptide FES is a broad, shallow basin that corresponds to a diverse range of somewhat compact conformations. This is perhaps not surprising, as peptides that demonstrate stable secondary structure, for example, the tryptophan zipper, are considerably longer.[82]

It is generally accepted that carbohydrates rigidify protein structure and alter the torsional preferences of neighboring residues.[16] Our simulations indicate that these effects primarily reflect steric crowding and, for the simulated sequence, most dramatically impact Gly6. Prior to glycosylation, Gly6 samples helical and extended conformations with similar probability. Upon glycosylation, Gly6 no longer samples this helical conformation and the extended conformation becomes even more extended. The carbohydrate excluded volume slightly perturbs the Asn5 and Thr7 backbone dihedrals and has minimal impact upon the other residues. It is intriguing that this steric effect is most pronounced for the most flexible residue, Gly6, and that flexible residues are commonly found immediately following glycosylated Asn residues.[40] These considerations suggest that steric forces may guide folding by limiting the conformational space of specific flexible residues in the glycosylation sequon.[29]

Nonspecific interactions with the N-linked disaccharide also eliminate extended structures from the simulated peptide ensemble, significantly alter the peptide hydrogen-bonding pattern, and destabilize Asx-turns, while stabilizing the proposed β-turn.[11] Nevertheless, nonspecific interactions with the carbohydrate do not dramatically alter the ensemble of compact conformations sampled by the peptide and certainly do not stabilize any single well-defined structure.

While nonspecific interactions are insufficient to stabilize any particular configuration, the combination of nonspecific interactions with specific attractive interactions guides the simulated glycopeptide to new regions of configuration space and stabilizes a NSC with an altered twist. In this case, direct attractive peptide–glycan interactions reinforce sterically induced changes in local backbone torsional preferences. Particular hydrogen bonds between the peptide backbone and the N-acetyl group of the proximal glucosamine contribute to stabilizing this NSC. These hydrogen bonds are consistent with experimental observations that the peptide conformation is remarkably sensitive to the presence of this N-acetyl group.[52] Simultaneously, attractive stacking interactions between the Trp8 aromatic side chain and the distal carbohydrate bury significant hydrophobic surface area and further stabilize the NSC. This stacking interaction is consistent with numerous recent studies that have quantified strong aromatic–glycan interactions,[47,83–85] demonstrated their significance for stabilizing novel protein structures,[32,33] and revealed their biological function in molecular recognition[46,86] and enzymatic catalysis.[87,88] Moreover, this interaction is also consistent with experimental studies,[41–43] including the FET studies of Imperiali and co-workers,[14] demonstrating that carbohydrate–protein interactions reflect hydrophobic forces. However, the NSC is only metastable, which may explain why NMR studies observed only weak NOE signals associated with Trp8–carbohydrate interactions and why these studies concluded that specific interactions did not significantly stabilize any particular peptide structure.[11,52]

Prior studies have suggested that carbohydrates may indirectly impact peptide conformation by reorganizing the surrounding solvent structure and, thus, enhancing the hydrophobic driving force for collapse.[14,34,52] Our simulations suggest that disaccharides may have relatively little effect upon the tetrahedral structure of water solvating unfolded peptides. In fact, the disaccharide appeared to have less impact upon the solvation structure than the peptide itself. However, previous simulation studies[34] suggest that larger carbohydrates, for example, the large glycan that is attached during N-linked glycosylation, may have a greater effect upon the hydrophobic collapse of larger proteins.

Prior studies have also suggested that carbohydrates predominantly impact protein structure by entropically destabilizing the unfolded ensemble of extended conformations.[35,37] Our simulations suggest that, although carbohydrates eliminate extended conformations, excluded volume alone does not significantly alter the ensemble of compact conformations sampled by a short disordered peptide. Moreover, an estimate based upon the calculated FES suggests that this compaction corresponds to a relatively modest entropic cost of approximately 1 kJ/mol. Of course, these effects may be more important for larger proteins and glycans.

We note that the present simulation studies, like any other classical MD studies, may be subject to systematic errors from inadequate sampling and from inaccuracies in empirical classical force fields. We emphasize that the present studies employed extensive REMD simulations in order to minimize sampling errors. Although it is difficult to prove complete sampling of the configuration space, the calculated free energy surfaces suggest that each simulation was adequately converged and accessed the relevant configuration space. On the other hand, we employed standard empirical force fields to describe the carbohydrate, peptide (OPLS-AA),[55,56] and solvent (SPC/E).[54] These models are certainly more detailed, more realistic, and also likely more accurate than empirical coarse-grained models[27,35,37] or implicit solvent models[26] that have been previously employed to investigate carbohydrate–peptide interactions. At the same time, the development of improved force fields for proteins, water, and especially carbohydrates[89–91] remains a very active field of current research. We note that the carbohydrate and Asn-linkage both remained in expected conformations throughout the simulations. Much more importantly, we emphasize that the current results are remarkably consistent with an extensive body of experimental observations. However, the empirical classical force field may possibly provide a poor model for particular interactions,[92,93] such as aromatic-glycan stacking, and thus incorrectly stabilize particular regions of configuration space, for example, Asx-turns in the octapeptide simulations or β-turns in the glycopeptide simulations.

Nevertheless, we suggest that the primary observations of the present work are robust and fundamentally correct. In particular, we suggest the following model, which is consistent with a wide array of experimental observations and almost all observations for this particular system: (1) Short peptides sample a heterogeneous ensemble both before and after glycosylation. (2) Steric interactions with N-linked disaccharides exclude highly extended structures from this ensemble and also alter local secondary structure preferences, especially for neighboring residues that are small and flexible. (3) Steric interactions alone, though, may not dramatically impact either the distribution of compact states within this ensemble or the tetrahedral structure of water solvating the peptide. (4) However, the combination of nonspecific steric interactions with specific context-dependent interactions, such as hydrogen-bonding and aromatic stacking, may provide considerable stabilization to conformations that would otherwise be only very rarely sampled. Because glycosylation occurs cotranslationally for incompletely folded proteins, these considerations may significantly impact folding by guiding glycoproteins to new regions of configuration space.

■ CONCLUSION

In summary, our simulations provide direct quantitative insight into the impact of specific and nonspecific glycan–peptide interactions upon the conformational tendencies of a disordered octapeptide. These simulations provide further evidence that the excluded volume of an N-linked carbohydrate reduces the accessible conformational space, alters torsional preferences, and stabilizes compact conformations. However, nonspecific steric effects only modestly impacted the configurational entropy of the unfolded ensemble and did not provide sufficient driving force to either stabilize specific structures or significantly alter the underlying FES for the simulated glycopeptide. Moreover, the simulations suggest that N-linked disaccharides, and, in particular, the core chitobiose glycan minimally disrupt the tetrahedral solvation structure for glycopeptides.

On the other hand, the combination of excluded volume and specific, sequence-dependent interactions did more dramatically alter the simulated FES. In particular, specific hydrogen bonds and carbohydrate−aromatic stacking interactions, when coupled with nonspecific steric forces, reduced the hydrophobic surface area and partially stabilized a new NSC that was not sampled prior to glycosylation. This NSC corresponds not only to an overall compaction, but also to an altered twist of the peptide backbone.

We emphasize that our simulations considered a short octapeptide and a small disaccharide. Although interactions with the disaccharide dramatically altered the peptide FES, these interactions were insufficient to completely stabilize a single well-defined conformation for such a short sequence. Steric and solvation forces may more significantly alter the FES in the case of a larger protein or glycan. Clearly, our results motivate future simulation studies that systematically investigate carbohydrate−protein interactions as a function of peptide sequence and glycan size. Nevertheless, our results support the hypothesis that specific and nonspecific interactions may act cooperatively to channel cotranslationally modified glycoproteins toward new regions of configuration space and stabilize specific intermediates on the path to productive and efficient folding.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information provides additional information and analysis for the reported simulations regarding the simulated force field, equilibration procedures, convergence, experimental comparisons, solvation structure, and hydrogen-bonding interactions. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
wnoid@chem.psu.edu
Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Taylor, M. E.; Drickamer, K. *Introduction to GlycoBiology*; Oxford Press: New York, NY, 1987.
(2) Rademacher, T.; Parekh, R.; Dwek, R. *Annu. Rev. Biochem.* **1988**, *57*, 785−838.
(3) Mitra, N.; Sinha, S.; Ramya, T.; Surolia, A. *Trends Biochem. Sci.* **2006**, *31*, 156−163.
(4) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364−2369.
(5) Drickamer, K.; Taylor, M. E. *Genome Biol.* **2002**, *3*, 1034.
(6) Hart, G. *J. Biol. Chem.* **1982**, *257*, 151−158.
(7) Bosques, C.; Imperiali, B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7593−7598.
(8) Marquardt, T.; Helenius, A. *J. Cell Biol.* **1992**, *117*, 505−513.
(9) Molinari, M. *Nat. Chem. Biol.* **2007**, *3*, 313−320.
(10) Laczko, I.; Hollosi, M.; Urge, L.; Ugen, K.; Weiner, D.; Mantsch, H.; Thurin, J.; Otvos, L. *Biochemistry* **1992**, *31*, 4282−4288.
(11) O'Connor, S.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 2295−2296.
(12) Slynko, V.; Schubert, M.; Numao, S.; Kowarik, M.; Aebi, M.; Allain, F. H. T. *J. Am. Chem. Soc.* **2009**, *131*, 1274−1281.
(13) Andreotti, A.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352−3353.
(14) Imperiali, B.; Rickert, K. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 97−101.
(15) Live, D.; Kumar, R.; Beebe, X.; Danishefsky, S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759−12761.
(16) Wormald, M.; Dwek, R. *Struct. Folding Des.* **1999**, *7*, R155−R160.
(17) Nishimura, I.; Uchida, M.; Inohana, Y.; Setoh, K.; Daba, K.; Nishimura, S.; Yamaguchi, H. *J. Biochem.* **1998**, *123*, 516−520.
(18) Yamaguchi, H.; Uchida, M. *J. Biochem.* **1996**, *120*, 474−477.
(19) Wyss, D.; Choi, J.; Li, J.; Knoppers, M.; Willis, K.; Arulanandam, A.; Smolyar, A.; Reinherz, E.; Wagner, G. *Science* **1995**, *269*, 1273−1278.
(20) Narhi, L.; Arakawa, T.; Aoki, K.; Elmore, R.; Rohde, M.; Boone, T.; Strickland, T. *J. Biol. Chem.* **1991**, *266*, 23022−23026.
(21) Riederer, M.; Hinnen, A. *J. Bacteriol.* **1991**, *173*, 3539−3546.
(22) Kern, G.; Schulke, N.; Schmid, F.; Jaenicke, R. *Protein Sci.* **1992**, *1*, 120−131.
(23) DeKoster, G.; Robertson, A. *Biochemistry* **1997**, *36*, 2323−2331.
(24) Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299−7307.
(25) Elliott, S.; Chang, D.; Delorme, E.; Eris, T.; Lorenzini, T. *J. Biol. Chem.* **2004**, *279*, 16854−16862.
(26) Chen, M. M.; Bartlett, A. I.; Nerenberg, P. S.; Friel, C. T.; Hackenberger, C. P. R.; Stultz, C. M.; Radford, S. E.; Imperiali, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22528−22533.
(27) 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.
(28) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Act.* **1999**, *1473*, 4−8.
(29) Shental-Bechor, D.; Levy, Y. *Curr. Opin. Struc. Biol.* **2009**, *19*, 524−533.
(30) Walsh, G.; Jefferis, R. *Nat. Biotechnol.* **2006**, *24*, 1241−1252.
(31) Sola, R. J.; Griebenow, K. *J. Pharm. Sci.* **2009**, *98*, 1223−1245.
(32) 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.
(33) Price, J. L.; Powers, D. L.; Powers, E. T.; Kelly, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 14127−14132.
(34) Cheng, S.; Edwards, S. A.; Jiang, Y.; Gräter, F. *ChemPhysChem* **2010**, *11*, 2367−2374.
(35) Hoffmann, D.; Florke, H. *Folding Des.* **1998**, *3*, 337−343.
(36) Imperiali, B.; O'Connor, S. *Curr. Opin. Chem. Biol.* **1999**, *3*, 643−649.
(37) Shental-Bechor, D.; Levy, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8256−8261.
(38) Sinha, S.; Surolia, A. *Biophys. J.* **2007**, *92*, 208−216.
(39) Kwon, K.; Yu, M. *Biochim. Biophys. Acta* **1997**, *1335*, 265−272.

(40) Petrescu, A.; Milac, A.; Petrescu, S.; Dwek, R.; Wormald, M. *Glycobiology* **2004**, *14*, 103−114.

(41) Jitsuhara, Y.; Toyoda, T.; Itai, T.; Yamaguchi, H. *J. Biochem.* **2002**, *132*, 803−811.

(42) Toyoda, T.; Arakawa, T.; Yamaguchi, H. *J. Biochem.* **2002**, *131*, 511−515.

(43) Barb, A. W.; Borgert, A. J.; Liu, M.; Barany, G.; Live, D. Intramolecular Glycan-Protein Interactions in Glycoproteins. In *Methods in Enzymology*; Fukuda, M., Ed.; Elsevier Academic Press, Inc.: San Diego, CA, 2010; Vol. *478*: Glycomics; pp 365−388.

(44) Mizushima, T.; Hirao, T.; Yoshida, Y.; Lee, S.; Chiba, T.; Iwai, K.; Yamaguchi, Y.; Kato, K.; Tsukihara, T.; Tanaka, K. *Nat. Struct. Mol. Biol.* **2004**, *11*, 365−370.

(45) Vyas, N. K. *Curr. Opin. Struc. Biol.* **1991**, *1*, 732−740.

(46) Fernandez-Alonso, M.; Canada, F.; Jimenez-Barbero, J.; Cuevas, G. *J. Am. Chem. Soc.* **2005**, *127*, 7379−7386.

(47) Chavez, M.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J.; Asensio, G.; Muraki, M.; Canada, F.; Jimenez-Barbero, J. *J. Chem. Eur. J.* **2005**, *11*, 7060−7074.

(48) Hanson, S. R.; Culyba, E. K.; Hsu, T.-L.; Wong, C.-H.; Kelly, J. W.; Powers, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3131−3136.

(49) O'Connor, S.; Pohlmann, J.; Imperiali, B.; Saskiawan, I.; Yamamoto, K. *J. Am. Chem. Soc.* **2001**, *123*, 6187−6188.

(50) Wilson, I.; Skehel, J.; Wiley, D. *Nature* **1981**, *289*, 366−373.

(51) Gamblin, S. J.; Skehel, J. J. *J. Biol. Chem.* **2010**, *285*, 28403−28409.

(52) O'Connor, S.; Imperiali, B. *Chem. Biol.* **1998**, *5*, 427−437.

(53) Bosques, C.; Tschampel, S.; Woods, R.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421−8425.

(54) Berendsen, H.; Grigera, J.; Straatsma, T. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(55) Jorgensen, W.; Maxwell, D.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(56) Kony, D.; Damm, W.; Stoll, S.; van Gunsteren, W. *J. Comput. Chem.* **2002**, *23*, 1416−1429.

(57) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.; Berendsen, H. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

(58) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theor. Comp.* **2008**, *4*, 435−447.

(59) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

(60) Parrinello, M.; Rahman, A. *J. Chem. Phys.* **1982**, *76*, 2662−2666.

(61) York, D.; Darden, T.; Pedersen, L. *J. Chem. Phys.* **1993**, *99*, 8345−8348.

(62) Hess, B. *J. Chem. Theor. Comp.* **2008**, *4*, 116−122.

(63) Allen, M. P.; Tildesley, D. P. *Computer Simulation of Liquids*; Oxford Press: New York, NY, 1987.

(64) Schuttelkopf, A.; van Aalten, D. *Acta Crystallogr.* **2004**, *D60*, 1355−1363.

(65) Patriksson, A.; van der Spoel, D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073−2077.

(66) Wormald, M.; Petrescu, A.; Pao, Y.; Glithero, A.; Elliott, T.; Dwek, R. *Chem. Rev.* **2002**, *102*, 371−386.

(67) Salisburg, A. M.; Deline, A. L.; Lexa, K. W.; Shields, G. C.; Kirschner, K. N. *J. Comput. Chem.* **2009**, *30*, 910−921.

(68) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(69) Sindhikara, D. J.; Emerson, D. J.; Roitberg, A. E. *J. Chem. Theor. Comp.* **2010**, *6*, 2804−2808.

(70) Amadei, A.; Linssen, A.; Berendsen, H. *Proteins* **1993**, *17*, 412−425.

(71) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(72) Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40−57.

(73) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195−1198.

(74) Eisenhaber, F.; Lijnzaad, P.; Argos, P.; Sander, C.; Scharf, M. *J. Comput. Chem.* **1995**, *16*, 273−284.

(75) Vuister, G. W.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 7772−7777.

(76) Kumar, P.; Buldyrev, S. V.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 22130−22134.

(77) Varki, A. *Glycobiology* **1993**, *3*, 97−130.

(78) Friedel, M.; Baumketner, A.; Shea, J.-E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8396−8401.

(79) Friedel, M.; Baumketner, A.; Shea, J.-E. *J. Chem. Phys.* **2007**, *126*, 095101.

(80) Cheung, M.; Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4753−4758.

(81) Pincus, D. L.; Thirumalai, D. *J. Phys. Chem. B* **2009**, *113*, 359−368.

(82) Cochran, A.; Skelton, N.; Starovasnik, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578−5583.

(83) Screen, J.; Stanca-Kaposta, E.; Gamblin, D.; Liu, B.; Macleod, N.; Snoek, L.; Davis, B.; Simons, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 3644−3648.

(84) Terraneo, G.; Potenza, D.; Canales, A.; Jiménez-Barbero, J.; Baldridge, K. K.; Bernardi, A. *J. Am. Chem. Soc.* **2007**, *129*, 2890−2900.

(85) Laughrey, Z. R.; Kiehna, S. E.; Riemen, A. J.; Waters, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 14625−14633.

(86) Aboitiz, N.; Vila-Perelló, M.; Groves, P.; Asensio, J. L.; Andreu, D.; Cañada, F. J.; Jiménez-Barbero, J. *ChemBioChem* **2004**, *5*, 1245−1255.

(87) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004**, *382*, 769−81.

(88) van Bueren, A. L.; Boraston, A. B. *J. Mol. Biol.* **2007**, *365*, 555−560.

(89) Fadda, E.; Woods, R. J. *Drug Discovery Today* **2010**, *15*, 596−609.

(90) Raman, E. P.; Guvench, O.; MacKerell, A. D. *J. Phys. Chem. B* **2010**, *114*, 12981−12994.

(91) Hansen, H. S.; Hünenberger, P. H. *J. Comput. Chem.* **2011**, *32*, 998−1032.

(92) Spiwok, V.; Lipovova, P.; Skalova, T.; Vondrackova, E.; Dohnalek, J.; Hasek, J.; Kralova, B. *J. Comp.-Aided Mol. Des.* **2006**, *19*, 887−901.

(93) Woods, R. J.; Tessier, M. B. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575−83.

(94) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.