JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Combined Molecular Dynamics Simulations and Experimental Studies of the Structure and Dynamics of Poly-Amido-Saccharides

Stacy L. Chin,[†] Qing Lu,[§] Eric L. Dane,[†] Laura Dominguez,[†] Christopher J. McKnight,[∥] John E. Straub,*[,†,§] and Mark W. Grinstaff*[,†,‡,§,⊥]

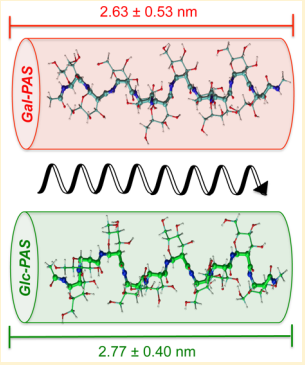
[†]Department of Chemistry, and [‡]Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, United States

[§]Division of Materials Science and Engineering, Boston University, Brookline, Massachusetts 02446, United States

[∥]Department of Physiology and Biophysics, and [⊥]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts 02118, United States

Supporting Information

ABSTRACT: Poly-amido-saccharides (PAS) are carbohydrate-based, enantiopure synthetic polymers in which sugar repeat units are joined by amide linkages. This unique and relatively rigid pyranose backbone contributes to their defined helical secondary structure and remarkable chemical properties. Glucose- (glc-) and galactose- (gal-) PAS 10-mer structures are synthesized and investigated with molecular dynamics (MD) simulations and experimental measurements. Quantum mechanical DFT energy minimization calculations, as well as experimental observables including circular dichroism, ^{1}H,^{13}C-HSQC, and ^{1}H,^{1}H-NOESY 2D-NMR studies, validated the all-atom simulation models produced using a modified CHARMM force field. Water radial distribution functions show distinct differences in the glc- and gal-PAS systems that correlate well with observed differences in solubility between gal-PASs and glc-PASs. The computational analysis and MD simulations are in good agreement with experimental results, validating the proposed models as reliable representations of novel glc- and gal-PASs.



INTRODUCTION

Synthetic biopolymers are responsible for many breakthroughs in the biotechnology, medical device, and pharmaceutical areas.[1] Mimetics of nucleic acids, such as peptide nucleic acid (PNA),[2−4] glycol nucleic acid (GNA),[5] locked nucleic acid (LNA),[6,7] and phosphorodiamidate morpholino oligomer (PMO),[4,8] represent successes with profound impacts on science and medicine. These novel biopolymers and their applications arose as a consequence of advances in synthetic methodology and computation. In contrast to nucleic acid structures, polysaccharides are remarkably diverse in stereochemistry, functionalization, linkage types, and degree of branching, and, thus, are challenging synthetic targets.[9,10] Notable structural and functional mimics of polysaccharides include glycopolymers[11−15] (those containing pendent sugars), nonether linked carbohydrate polymers,[16] and those that are structural analogues.[17] Synthetic polysaccharides have the potential to similarly impact society given the important biological roles they perform in structure, storage, and molecular recognition.[18] Therefore, it is important to develop new synthetic strategies and characterization tools for natural polysaccharides or synthetic polysaccharide mimetics.[19]

Poly-amido-saccharides (PASs) are well-defined, enantiopure carbohydrate polymers that share many important features with natural polysaccharides, including being stereochemically defined, hydrophilic, and possessing pyranose rings in the backbone.[19] With carbohydrate units joined by unnatural α-(1,2)-amide linkages, PASs exhibit interesting chemical properties and structures. PAS are synthesized by an anionic ring-opening polymerization (AROP) reaction of β-lactam sugar monomers that provide glucose-derived[19] (glc-) and galactose-derived[20] (gal-) PASs in high-yields with batch-to-batch consistency, defined molecular weights, and low polydispersity. This synthetic approach addresses many of the common challenges associated with synthesizing carbohydrate polymers such as maintaining the high density of repeated functional groups, rigid pyranose backbone, and control over stereochemistry at the glycosidic linkages. One of the unique advantages of synthesizing PAS structures using AROP is the ability to strictly control the monomer composition and subsequent polymerization. Our initial studies with PASs show that glc-PAS can be recognized by the lectin concanavalin A,[19] and that gal-PAS and glc-PAS are noncytotoxic to mammalian cell lines.[20] Continued investigations of these polymers require additional details and insight into their solid-state and solution structures, as well as their resultant properties.

Given the challenges of growing single crystals of a synthesized carbohydrate-based polymer, molecular dynamics (MD) simulations using all-atom (AA) models is an effective approach to relate structure and dynamics of these complex

Received: February 18, 2016
Published: April 27, 2016

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

macromolecules to their physical properties.[21] Molecular mechanic force fields have been developed to model the energetics of carbohydrate systems while taking into consideration the structural diversity found within polysaccharides.[22,23] Matthews and co-workers used three carbohydrate force fields, CHARMM35, GLYCAM06, and Gromos 45a4, to model and study conformation changes in hydrated 36-chain cellulose Iβ microfibrils.[24] Their simulation results derived from two force fields, CHARMM35 and GLYCAM06, were consistent with experimentally observed behaviors of cellulose microfibrils. Kuttel et al. developed a CHARMM Carbohydrate Solution Force Field (CSFF) that accurately reproduced conformational distributions of carbohydrate structures in solution.[25] Their analysis of the rotational frequency of the pyranose primary alcohol group were in agreement with experimental NMR results. More recently, Guvench et al. developed force field parameters for monosaccharide derivatives as an extension of the CHARMM all-atom additive biomolecular force field.[26] The parameters were validated against X-ray crystallographic data of corresponding monosaccharides, as well as against NMR data for larger systems. These findings suggest that recent advances in the quality and scope of carbohydrate force fields represent significant progress in the ability to use MD simulations to probe the structural and thermodynamic properties of carbohydrate systems.

Glc- and gal-PAS polymers are stereoisomers that differ only at the C_4 position (Figure 1); however, they display different chemical and physicals properties, for example, solubility in water. The optimized and validated computational models developed in this study allow visualization and quantification of the conformational preferences of PAS structures at multiple length scales. As such, these computed models and complementary experimental studies provide a fundamental understanding between PAS composition, structure, and properties, as well as a foundation for further understanding of how synthesized polysaccharide mimics can represent the diversity found within the natural polysaccharides.

Herein, we report the (1) synthesis and characterization of glc- and gal-PAS 10-mer samples; (2) development of a modified force field to include parameters for the peptide backbone of poly-amido-saccharides using CHARMM36 lipid and carbohydrate force fields;[22,26−29] (3) quantum mechanical (QM) calculations and potential energy scans of glc- and gal-PAS representative structures to understand the φ,ψ dihedral space within the PAS backbone; (4) experimental and calculated circular dichroism analysis of the helical PAS structures; (5) experimental 2D-NMR analysis and solubility studies of glc- and gal-PAS structures in conjunction with MD simulations and water-PAS structural radial distribution functions (RDF); and, finally, (6) end-to-end distance and radius of gyration based on MD simulations of glc- and gal-PAS structures to compare chain lengths and polymer rigidity.

EXPERIMENTAL SECTION

Complete details of the synthesis, characterization, and experimental approaches are found in the Supporting Information (SI).

RESULTS AND DISCUSSION

β-Lactam Monomer Synthesis and PAS Polymerization. PAS polymerization reactions of the glc-derived and gal-derived β-lactam monomers afforded glc-PAS and gal-PAS 10-mer samples following previously published procedures as summarized in Scheme 1.[19,20] Briefly, glc- and gal-derived β-

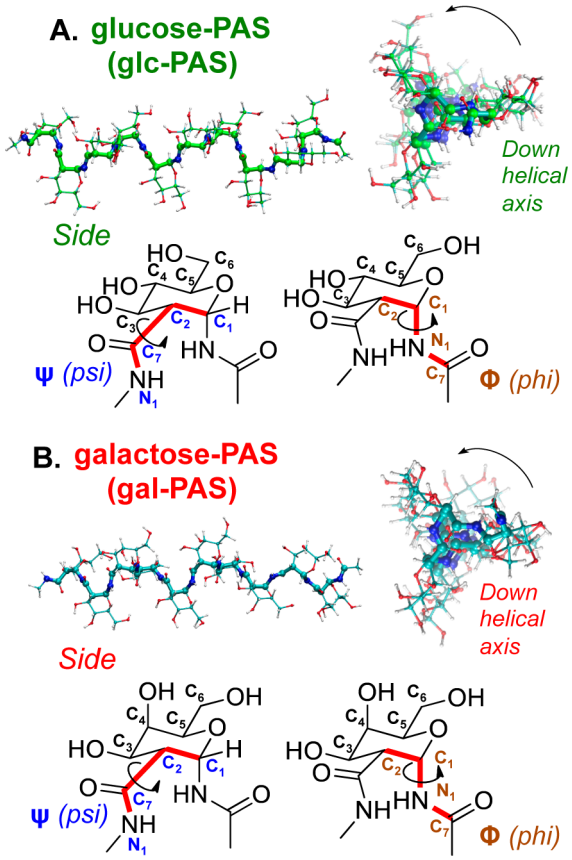
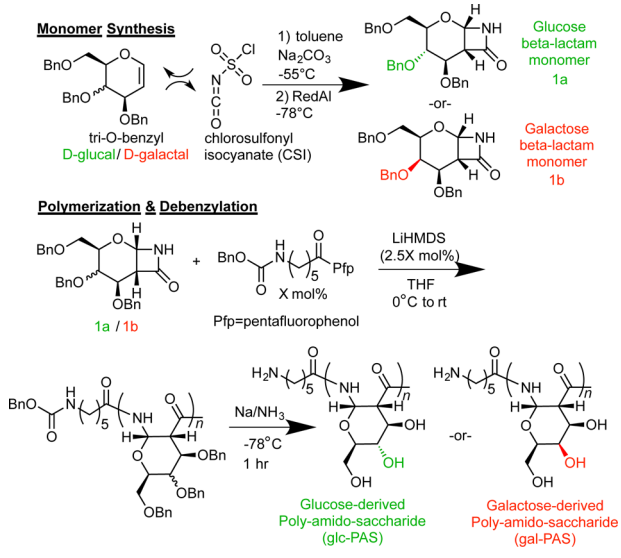


Figure 1. AA models of (A) glc-PAS and (B) gal-PAS 10-mer structures shown in side view (left) and down the helical axis (right). The helical PAS backbone is highlighted in bold in the side view, as well as down the axis. The Ψ (psi) and Φ (phi) dihedral angles within the PAS backbone are as defined for both glc- and gal-PAS structures. The dihedral angle (in red) centered on $C_7-N_1-C_1-C_2$ is defined as Φ, while the dihedral angle centered on $C_1-C_2-C_7-N_1$ as Ψ along the PAS backbone.

Scheme 1. β-Lactam Monomer Synthesis and Polymerization of Glucose-Derived PAS (glc-PAS) or Galactose-Derived PAS (gal-PAS)



DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

lactam monomers were obtained via the stereoselective cycloaddition of tri-*O*-benzyl D-glucal or D-galactal with chlorosulfonyl isocyanate (CSI). Both monomer samples were obtained in moderate yields (65−78%) and purified by column chromatography. The PAS polymerization was performed by using the pentafluorophenol ester of *Z*-9-amino-hexanoic acid as the initiator. The polymers were debenzylated with sodium metal and liquid ammonia, purified via dialysis, and lastly lyophilized to isolate the final products. All samples were obtained with good yields of 85−90%. Samples were characterized via gel permeation chromatography (Table S1), ^{1}H NMR, and IR spectroscopy (Figure S1).

Simulated Structural Ensembles of glc- and gal-PAS Suggests Rigid Helical Structures in Agreement with Circular Dichroism Spectroscopy. We carried out 200 ns MD simulations of the 10-mer glc- and gal-PAS polymers in explicit TIP3P solvent in the constant NPT (isothermal−isobaric) ensemble using the GROMACS package[30,31] and the CHARMM36 lipid and carbohydrate[22,26−29] force fields modified to include parameters for the PAS peptide backbone. The pressure and temperature were constrained (1 atm, 300 K) using the Parrinello−Rahman barostat and Berendsen thermostat. The overall system size was N_{PAS} = 1 and N_{H_2O} = 13 486. Analysis of the general PAS backbone structure revealed that there are four unique bonds that repeat along the backbone: C_7−N_1, N_1−C_1, C_1−C_2, and C_2−C_7, as highlighted in red in Figure 1. Rotation around these bonds represented major potential sources of disorder along the polymer backbone, with variations in bond angles and lengths playing a secondary role. Using the conventions developed for β-peptides,[32−34] we defined the dihedral angle centered on N_1−C_1 (C_7−N_1−C_1−C_2) as φ, and the dihedral angle centered on C_2−C_7 (C_1−C_2−C_7−N_1) as ψ. The generalized pseudo-Ramachandran scatter plots for the last 20 ns of the simulations are shown in Figure 2. No significant differences were observed in these contour plots between glc- and gal-PAS 10-mer structures as both plots showed tight clustered populations centered about $\varphi = -100°$ and $\psi = 80°$, indicating that the PAS backbone was relatively rigid with unimodal fluctuations. We performed additional calculations with the 12-mer and the 14-mer glc- and gal-PAS structures which showed similar clustered positions, suggesting that extending the chain length did not affect the φ,ψ space of the PAS backbone (Figure S3).

Previously, we reported that both glc- and gal-PAS polymers exhibit a strong CD signal with a positive peak near 190 nm and a negative peak near 220 nm (Figure 3A), suggestive of a helical secondary structure.[19,20] In our initial report we modeled a glc-PAS oligomer using MMFF94s in the gas phase, and observed a left-handed helical structure with 3-fold symmetry that contained inter-residue hydrogen bonds.[19] In order to investigate whether hydrogen bonds between PAS repeat units play an important role in stabilizing the secondary structure, we conducted additional CD and NMR experiments.

Using CD, we experimentally observed that the secondary structure in both glc- and gal-PASs did not significantly change over a wide range of conditions for pH[36] (from 2 to 12), temperature (from 20 to 80 °C), and varying ionic salts concentration (i.e., 1 M NaCl, KCl, and LiCl) (Figure S4), as well as in the presence of common protein denaturants, such as 5 M of urea or guanidine HCl. If *inter*-residue hydrogen bonding played an important role in secondary structure, we would have expected to see significant changes in the CD

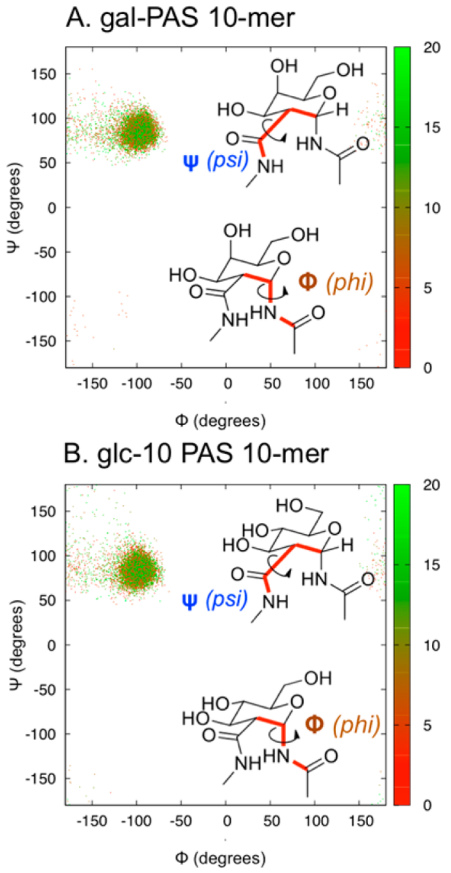


Figure 2. Generalized pseudo-Ramachandran scatter plot of φ and ψ dihedral angles of the last 20 ns of the 200 ns MD simulation results of (A) gal-PAS and (B) glc-PAS 10-mer structures performed in explicit water. Color bars indicate simulation time from 0 to 20 ns. Results are shown for the dihedral angles of the noncapping residues. Similar scatter plots were obtained for 12-mer and 14-mer glc- and gal-PAS structures.

spectrum when varying these conditions. We did observe a loss in the PAS CD signal following sodium periodate oxidation and tautomerization reactions to open the pyranose rings along the PAS backbone by oxidizing the vicinal diols of the C_3 and C_4 to aldehyde groups.[36] Taken as a whole, these observations suggested that the rigid secondary helical structure observed in the PAS polymers is a consequence of backbone steric constraints resulting from the conformationally restricted pyranose rings and that intramolecular interactions play a minor role.

We confirmed this observation by conducting experimental hydrogen−deuterium exchange[43,44] studies to examine the amide protons along the PAS backbone. Hydrogen−deuterium exchange for both glc- and gal-PAS 10-mer samples occurred rapidly, within 4 min, and amide hydrogens did not appear to be protected from rapid solvent exchange (Figure S5). This finding supported the conclusion that inter-residue hydrogen bonding interactions are not primarily responsible for the rigidity of the PAS backbone and their helical characteristics.

Next, we computed CD spectra using glc- and gal-PAS polymer conformations derived from 100 ps of MD simulations with the Dichro-Cal[39−42] software program to determine whether the newly developed AA models reliably represented the glc- and gal-PAS systems (Figure 3B). In agreement with our experimental data, the calculated spectra showed strong CD

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

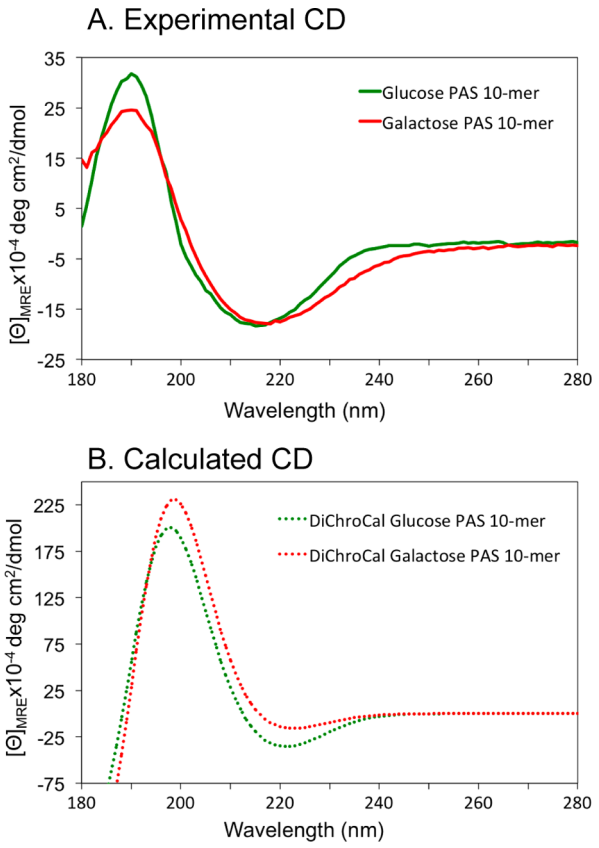


Figure 3. Circular dichroism (CD) spectra of glc-derived PAS and gal-derived PAS samples using experiment (A) in comparison to calculated CD analysis of 100 ps MD simulation of glc- and gal-PAS 10-mer structures from DichroCal.[39−42] Both spectra depict features indicative of helical conformations in both glc- and gal-PAS structures.

signals with a positive peak around 198 nm (compared to the 191 nm in experiment) and a negative peak at 220 nm (compared to the 219 nm in experiment). These results supported the validity of our AA-models and the structural ensembles derived from our MD simulations.

Interrogation of the simulated structures revealed that both PAS oligomers formed a left-handed helix with approximately 3-fold symmetry. The backbone amide groups orient such that the carbonyl and the nitrogen−hydrogen bonds point away from the helical axis and interact predominantly with solvent water molecules (Figure 1). The observation that the amide bonds are well-solvated is in agreement with our experimental results suggesting that they are not engaged in helix-stabilizing inter-residue hydrogen bonding.

Previously reported helical secondary structures observed in β-polypeptides,[32−34,37,38] such as the 14-helix, are stabilized by inter-residue hydrogen bonding. These alpha-PAS oligomers cannot form the 14-helix observed, for example, for *trans*-2-aminocyclohexane carboxylic acid (ACHC) peptides[33−35] as the backbone bonds are in a *cis*-geometry. However, based on initial preliminary gas-phase modeling, the β-isomers of the PASs, which possess a trans-geometry, may be able to access a 14-helix conformation, as there are no clear torsional or steric barriers to the formation of the required stabilizing inter-residue hydrogen bonds (see the SI). Figure S2 depicts a side-by-side comparison of the alpha-PAS, 14-beta-peptide (ACHC), and beta-PAS helices. Thus, the helical secondary structure of PAS oligomers arises due to factors other than *inter*-residue hydrogen bonding, such as the conformational constraints imposed at the level of the individual PAS repeat unit.

Generalized Pseudo-Ramachandran Plots Show Multiple Potential Energy Minima in φ,ψ for PAS Monomers. We investigated the role of carbohydrate-derived repeating units in promoting the helical conformation along the PAS backbone using QM modeling with DFT calculations. To study how the repeat unit structure constrained the φ and ψ angles (Figure 1) within PASs, we used ab initio quantum chemical calculations to probe model structures representing glc- and gal-PAS, which we referred to as **1**-*Glc* and **1**-*Gal*, respectively (Figure 4). These model structures were composed

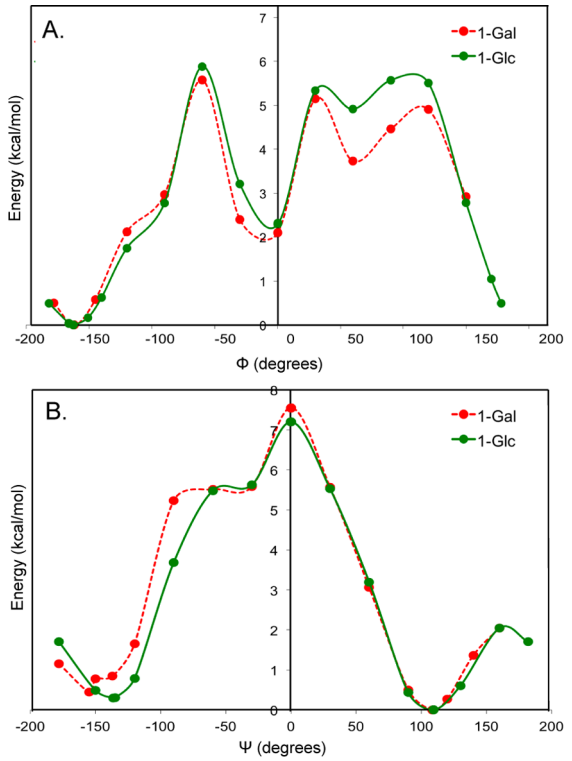


Figure 4. DFT calculations showing the relationship between the (A) φ dihedral angle and potential energy for **1**-*Glc*-PAS (solid green) and **1**-*Gal-PAS* (dashed red); and (B) ψ dihedral angle and potential energy for **1**-*Glc* (solid green) and **1**-*Gal* (dashed red).

of a single repeat unit of glc- and gal-PASs, with the glycosidic nitrogen (N_1) capped with an acetyl group and the carbonyl group at C_7 capped as an *N*-methyl amide. The structures were minimized and studied with DFT calculations (M06-2x/6-31G(d)) using a polarizable continuum model (PCM) to approximate aqueous solvation. The minimized structures observed had (φ,ψ) values of (−163°,109°) for **1**-*Glc* and (−162°,108°) for **1**-*Gal*. These consistent values between the two isomers suggested that a change in stereochemistry at C_4 has minimal impact on the dihedral angle preferences of the model structures.

To gain insight into the origin of these torsional preferences and to characterize the barriers to rotation, we constructed potential energy scans by independently varying the φ and ψ angles from −180° to 180°. For each model, we began with the minimized structure and manually adjusted φ to desired values through bond rotation. The resulting structure was then geometry optimized with φ constrained to the specified value

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

and the remaining degrees of freedom were optimized without constraints. The minimization was terminated and the relative energy recorded when the largest component of the energy gradient was less than 0.06 kcal/mol. This same approach was applied to the ψ angle of **1**-*Glc*, as well as to the φ and ψ angles of **1**-*Gal* to generate potential energy scans as shown in Figure S6.

The interrelationship between φ and ψ dihedral angles was examined by plotting the resultant ψ angle for each constrained φ angle that was sampled, as noted by the white circles, and the measured φ for each constrained ψ, as noted by the white triangles in Figure 5. The QM DFT model is also depicted in Figure S6. The scans suggested that the rotational freedom of N_1-C_1 was likely strongly correlated with the ability of the C_2-C_7 bond to sample ψ values within the energy basin containing the local minimum (lower left). However, the rotation of ψ through the sampled values caused only minor changes in the resultant φ values, and thus rotation of ψ was likely not strongly correlated with the ability of φ to sample values far from its global minimum.

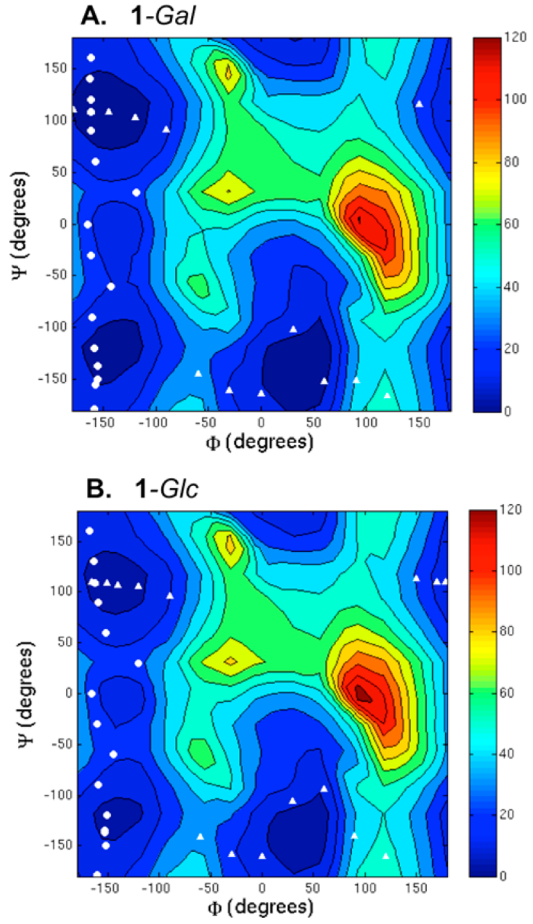


Figure 5. Potential energy scan of φ and ψ dihedral angles for (A) **1**-*Gal* and (B) **1**-*Glc* in implicit-solvent. The zero potential energy point is taken to be the lowest point on the potential energy scan. The color bars on the right side of the contour plot show the corresponding value of potential energy measured in kJ/mol. The superimposed QM data is noted by the white circles, which refer to the φ angle of the geometry optimized structure with constrained ψ angle, and the white triangles refer to the ψ angle plotted for geometry-optimized structure with constrained φ angle.

We validated this newly designed CHARMM force field against QM results obtained from DFT calculations by carrying out potential energy scans of φ and ψ dihedral angles for **1**-*Glc* and **1**-*Gal* from −180° to 180°. These scans were conducted in both implicit solvent (Figure 5) and vacuum (Figure S7). Only the scanned dihedral angles were constrained while all other degrees of freedom were minimized. Data points from the QM DFT model for **1**-*Glc* and **1**-*Gal* from Figure S6 were superimposed onto the potential energy contour plots in Figure 5 for comparison. Data points of QM energy minimization studies were mainly localized within the lower potential energy (dark blue) basins of the potential energy surface. This indicated that the energy minimization for the backbone dihedral angles using the modified CHARMM force field agrees with the data obtained from QM calculations, further validating the molecular mechanics force field.

To explore whether glc- and gal-PAS 10-mer structures had a minimum other than our reported values of $\varphi = -100°$ and $\psi = 80°$, we repeated the MD simulation at a higher temperature of 323 K. According to Figure 5, we expected to observe three different energy minima in the upper left quadrant, lower left quadrant, and lower middle in the generalized pseudo-Ramachandran plots based on these potential energy scans for the **1**-*Gal* and **1**-*Glc* monomeric models. The MD simulations of the PASs reported a localized dominant minimum at $\varphi = -100°$ and $\psi = 80°$ (Figure 2), as suggested by the energy minima in the upper left quadrant (Figure 5). When repeating the MD simulation at a higher temperature of 323 K (Figure S8), the glc- and gal-PASs became accessible to a second minimum around $\varphi = -100°$ and $\psi = -100°$, which was suggested by the minima noted in the lower left quadrant in Figure 5. However, a third potential energy minimum for the PAS was not observed in the bottom right quadrant. We hypothesized that configuring the PAS dihedral angles around $\varphi = 40°$ and $\psi = -150°$ will result in a steric clash between adjacent monomeric units, resulting in a low probability for the PAS to be in this configuration.

Agreement between Simulated Structural Ensembles for PAS 10-mers and 2D-NMR Data. We conducted ^{1}H,^{13}C-HSQC and ^{1}H,^{1}H-NOESY 2D-NMR studies with gal-PAS and glc-PAS 10-mer samples to compare the solution structures of the two polymers (Supporting Information). ^{1}H and ^{13}C peaks were assigned based on previously reported ^{1}H-NMR and ^{13}C-NMR analysis of glc-PAS[19] and gal-PAS[20] structures.

Differences in the ^{1}H and ^{13}C resonance and chemical shifts were compared between the ^{1}H,^{13}C-HSQC spectra of glc- (green) and gal-PAS (red) samples (see Figure S9). Both spectra demonstrated similar ^{1}H−^{13}C resonance and chemical shifts for the C_1-H_1 position, located in the PAS backbone, as well as along the aliphatic carbon chain of the initiator at the C_a-H_a, C_b-H_b, C_c-H_c, C_d-H_d, and C_e-H_e positions. These observations suggested that glc- and gal-PAS polymers possess similar structural features and chemical environments at these particular sites. However, the resonance correlating to C_2-H_2, C_3-H_3, C_4-H_4, C_5-H_5, and C_6-H_6 positions differ in chemical shifts and *J*-coupling patterns, which indicated structural differences between the glc- and gal-PAS within the pyranose rings rather than the backbone. These results were as expected since both polymers possess similar chemical structures and backbone, but only differ at the C_4 conformation.

In addition, we compared ^{1}H,^{1}H-NOESY spectra for glc- and gal-PAS 10-mer samples at a mixing time of 100 ms, which

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

revealed several key differences in structure. NOESY studies were repeated at mixing times of 200 ms and 50 ms, which revealed similar results. NOE signals were observed corresponding to protons located within the pyranose ring, specifically the H_2, H_3, H_4, H_5, and H_6, for glc- (green) and gal-PAS (red) structures (Figure 6). The strength of the NOE

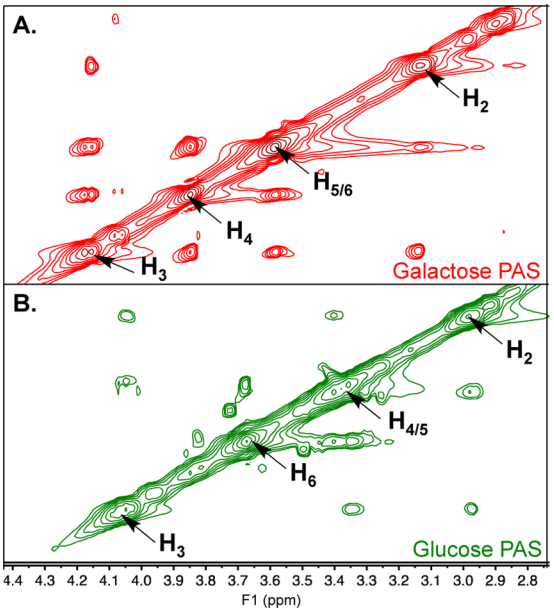


Figure 6. $^1H,^1H$-NOESY spectra of the H_3, $H_{4/5}$, and H_6 protons of (A) gal-PAS (red) and (B) glc-PAS (green) conducted in D_2O at a mixing time of 100 ms.

signals were conservatively classified as "strong" or "weak" in Table 1, which were related to approximate proton distances of less than 3.0 Å and greater than 3.0 Å (Figure S10), respectively.[45,46] Next, NOE data were compared to the average distances and variance between proton pairs within the 10 sugar units throughout the glc-and gal-PAS structure during the last 20 ns of our MD simulations (Table 1).

The majority of the structural measurements derived from the MD simulated PAS systems agreed with the NOE data. Measured proton distances that correlated to strong NOE signals corresponded to relative proton distances less than 3.0 Å, while those correlated to weak NOE signals corresponded to relative distances greater than 3.0 Å. We found similar proton measurements and NOE signal intensities between the glc- and gal-PAS samples for the H_1 to other protons located along the pyranose ring, suggesting that both PASs have comparable spatial relationships around the H_1 sites. The measured distances between the H_2−H_4 protons showed the most appreciable difference between PAS structures, with 2.49 ± 0.15 Å and 3.71 ± 0.08 Å for the glc- and gal-PAS, respectively. These differences were also reflected in our NOESY spectra as glc-PAS showed a stronger NOE signal than gal-PAS. A stronger NOE signal was also observed between the H_4 and H_6 protons in the gal-PAS sample than expected when compared to the calculated proton distance of 3.26 ± 0.35 Å from the MD simulation. This may result from *inter*-residual NOE signal contribution from neighboring sugar units. Two discrepancies were found in which the proton distance was not correlated with the strength of the NOE signal: the H_2−H_5 proton distance for the glc-PAS and the H_3−H_6 proton distance for the gal-PAS, which were not reported in Table 1. These NOE

Table 1. Comparison of Proton Distances within AA Models of Glc-PAS and Gal-PAS 10-mer Structures within the Pyranose Ring[a]

	galactose-PAS		glucose-PAS	
proton pairs	NOE signal	distance (Å)	NOE signal	distance (Å)
H_1−H_2	strong	2.34 ± 0.09	strong	2.37 ± 0.10
H_1−H_3	weak	3.78 ± 0.09	weak	3.78 ± 0.09
H_1−H_4	weak	4.84 ± 0.06	weak	4.02 ± 0.14
H_1−H_5	weak	3.67 ± 0.08	weak	3.67 ± 0.08
H_1−H_6	weak	4.64 ± 0.29	weak	4.82 ± 0.41
H_2−H_3	strong	3.03 ± 0.05	strong	3.03 ± 0.04
H_2−H_4	weak	3.71 ± 0.08	strong	2.49 ± 0.15
H_2−H_5	weak	3.86 ± 0.09	[b]	3.88 ± 0.10
H_2−H_6	weak	4.68 ± 0.38	weak	4.37 ± 0.42
H_3−H_4	strong	2.36 ± 0.09	strong	3.02 ± 0.05
H_3−H_5	strong	2.43 ± 0.14	strong	2.60 ± 0.18
H_3−H_6	[b]	4.59 ± 0.18	weak	4.58 ± 0.20
H_4−H_5	strong	2.39 ± 0.09	[b]	2.39 ± 0.09
H_4−H_6	strong	3.26 ± 0.35	strong	2.89 ± 0.42
H_5−H_6	[b]	2.74 ± 0.29	strong	2.69 ± 0.25

[a]Proton distances were based on the average distances between protons from MD simulations of glc- and gal-PAS, which were correlated to experimental NMR NOESY data. [b]NOE signals correlating to H_3−H_6 and H_5−H_6 for gal-PAS, as well as the H_2−H_5 and H_4−H_5 peaks for glc-PAS were not analyzed due to overlapping H_5 and H_6 NOE peaks in the gal-PAS spectra and H_4 and H_5 NOE peaks in the glc-PAS spectra

signals are complicated by the overlapping proton peaks with chemical shifts (Figure 6) contributing from the NOE signal for the $H_{5/6}$ peaks for the gal-PAS sample and from the $H_{4/5}$ peaks for the glc-PAS sample. Overall, we found the NOESY data to be in good agreement with the MD simulation analysis, further validating the AA simulations.

Difference in Water Solubility of glc- and gal-PAS Polymers Correlates with Radial Distribution Functions of the Degree of Hydrogen Bonding of Polymer Structure to Solvent Molecules. We previously reported that gal-PAS samples possessed high water solubility at all chain lengths (ranging from 10- to 100-mers) at concentrations up to 100 mg/mL. In contrast, glc-PAS samples tended to precipitate out of solution at ∼25 mg/mL under the same conditions, especially at greater degrees of polymerization (i.e., 50- and 100-mers).[20,19] These substantial differences in solubility were surprising given the relatively minor structural differences between polymers in which the C_4 hydroxyl groups of the pyranose rings are situated axial (gal-PASs) or equatorial (glc-PASs). To gain insight as to why this particular stereochemical difference drastically affects the solubility of these PAS structures and their hydration properties, we calculated the number of hydrogen-bonding sites within our AA-models of the PAS structures, as well as with surrounding solvent molecules, and analyzed water-PAS structural correlations through computed radial distribution functions (RDFs).

We investigated the number of potential *intra*molecular hydrogen-bonding interactions within the PAS 10-mer structures. A hydrogen bond is formed between an atom with a hydrogen atom bonded to it (the donor, D) and another atom (the acceptor, A) provided that the distance D−A is less than the cutoff distance (3.5 Å) and the angle D-H-A is less than the cutoff angle (30°). The number of hydrogen bonds was then calculated using GROMACS under these criteria. Gal-

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

PAS was calculated to have an average of 3.01 hydrogen-bonding interactions per PAS unit, which is less than glc-PAS that had an average of 4.88. These findings correlated well with our experimental findings suggesting that less *intra*molecular hydrogen-bonding interactions will enable more *inter*molecular hydrogen bonds to form with solvent molecules. We further investigated correlations between neighboring water molecules and key hydrophilic groups along the PAS structures, specifically to the oxygen atom within the −OH groups, at the C_3, C_4, and C_6 positions along the sugar ring, and the nitrogen atom within the amide linkages. A total of 87 potential *inter*molecular interactions between nearby water molecules to these key hydrophilic groups were calculated for both glc- and gal-PAS, indicating that both PAS structures have the same number of potential hydrogen-bonding interactions with surrounding water molecules. Although this finding was expected, it cannot explain the observation that glc- and gal-PAS samples have drastically different solubility properties. As such, other factors may be responsible for these experimental differences rather than the quantity of *intra*molecular or *inter*molecular hydrogen-bonding interactions.

To further probe the nature of the water−PAS interactions, RDFs were computed to compare the probability of forming solvent hydrogen bonds with key hydrophilic sites along the PAS. As depicted in Figure 7, all three RDFs indicated a higher probability for nearby water molecules to hydrogen-bond with −OH groups at the C_3, C_4, and C_6 positions along the pyranose rings in gal-PAS (red) compared to those in glc-PAS (green). Particularly, the RDFs in Figure 7A, depicting the probability of forming solvent hydrogen bonds to −OH groups at the C_3 position, showed the most appreciable difference. A similar trend was also noted when computing RDFs between nearby water molecules and the nitrogen atom within the amide linkages (Figure S11).

Our RDF calculations revealed a greater probability for water molecules to be in closer contact to key hydrophilic groups within the sugar moieties for gal-PAS structures compared to glc-PAS. These findings suggested that orienting the −OH group at the C_4 position to equatorial hinders the ability for glc-PASs to be in closer contact with neighboring water molecules as oppose to gal-PAS structures in which the C_4−OH group is axial and situated out of the plane. While this analysis lacks evaluation of differences in stability of the glc- and gal-PAS solid states, these findings correlated with the experimental observations that gal-PASs are substantially more water-soluble.

Glc- and Gal-PAS Structures Are Similar in Polymer Length. Finally, we investigated how the orientation of the −OH group at the C_4 position affected the overall structural conformation and length of the PAS structures by comparing the end-to-end distances and radius of gyration of our MD simulation models. The end-to-end distance and radius of gyration were calculated based on the last 100 ns of the 200 ns MD simulations of glc- (green) and gal-PAS (red) 10-mer, 12-mer, and 14-mer structures (Figures S12 and S13), with their mean values and deviations reported in Table 2. No statistically significant difference was observed in the mean values for the glc- and gal-PAS 10-mer, 12-mer, and 14-mer structures. These findings suggest that the sole structural difference in the glc- and gal-PAS structures, the orientation of the −OH group at the C_4 position, does not have a major impact on polymer length or conformational fluctuations.

The dependence of persistence length on polymer length was further explored by analyzing the end-to-end distance using the predictions of the worm-like-chain (WLC) model. The WLC model is appropriate for describing polymers that form flexible, rod-like structures as we have observed for the case of PAS polymers.[47−49] According to the WLC model, the relationship

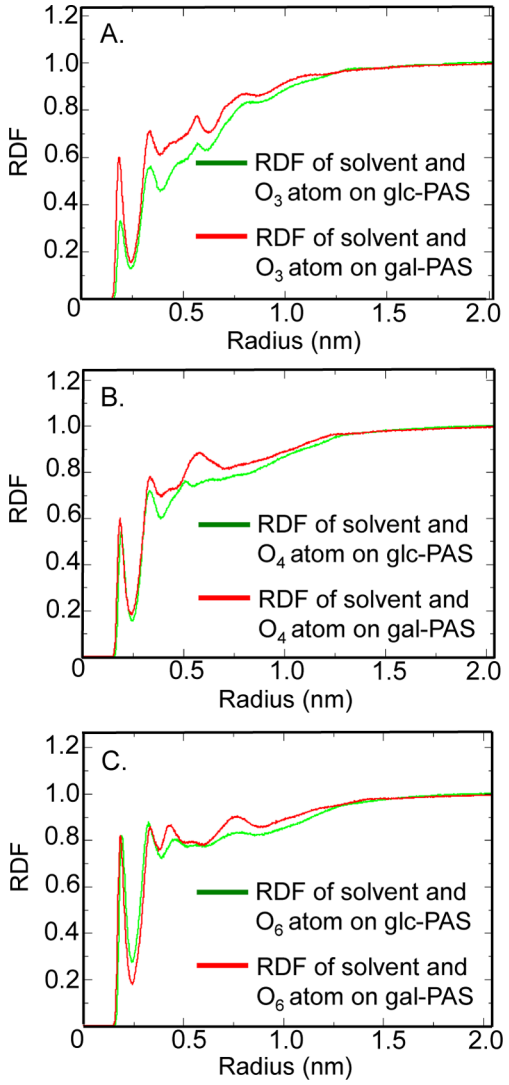


Figure 7. RDF calculations of water molecules to oxygen atoms from the −OH sites located at the (A) C_3, (B) C_4, and (C) C_6 positions along the pyranose ring within gal-PAS (red) and glc-PAS (green) 10-mer structures. All RDFs for gal-PAS show greater distributions than that of glc-PAS. The RDFs were computed based on the averages of all residues in the PAS 10-mer structures.

Table 2. End-to-End Distance and Radius of Gyration Calculations of Glc- and Gal-PAS 10-mer, 12-mer, and 14-mer Polymer Structures[a]

		radius of gyration (nm)	end-to-end distance (nm)
10-mer PAS	gal	0.93 ± 0.07	2.63 ± 0.53
	glc	0.93 ± 0.07	2.77 ± 0.40
12-mer PAS	gal	1.01 ± 0.08	2.86 ± 0.62
	glc	1.03 ± 0.09	3.03 ± 0.59
14-mer PAS	gal	1.23 ± 0.08	3.66 ± 0.61
	glc	1.23 ± 0.08	3.64 ± 0.64

[a]The distribution and calculated measurements are based on the last 100 ns of the 200 ns MD trajectory.

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

between the mean square end-to-end distance of the polymer and its persistence length can be defined as

$$\langle R^2 \rangle = 2Pl\left[1 - \frac{P}{l}(1 = e^{(-l/P)})\right] \quad (1)$$

where l is the length of the polymer chain and P is the persistence length. The mean square end-to-end, distance, $\langle R^2 \rangle$, was calculated from the last 100 ns of the 200 ns MD trajectory. The chain length, l, was calculated by multiplying the average length of one unit, 0.35 nm for both glc- and gal-PAS, by the number of repeating units. Based on the WLC model, the resulting persistence length was calculated to be 2.23 nm for glc-PAS and 2.02 nm for gal-PAS structures. This suggests similar polymer stiffness characteristics between glc- and gal-PAS structures, with glc-PAS polymers being slightly stiffer than gal-PAS.

CONCLUSION

We report a combined experimental and computational study to analyze and compare glc- and gal-PAS solution structures. Our results suggest that the glc- and gal-PAS backbone structures are relatively rigid and display similar unimodal fluctuations in φ,ψ dihedral space around values consistent with the formation of a helical structure. Both glc- and gal-PAS samples show a strong circular dichroism (CD) signal indicative of a helical conformation, which is stable and unperturbed under a wide-range of conditions (i.e., high temperatures, pH, salt concentrations, and denaturants concentrations). Quantum mechanical density functional theory (DFT) calculations and generalized pseudo-Ramachandran scatter plots, showing an energy minima at particular φ,ψ dihedral angles, reveal that the backbone configuration of PASs is in good agreement with calculated potential energy scans. Calculations of radial distribution functions, which correlate water molecules to key hydrophilic groups along the PAS structure indicate that there is a measurably closer contact between solvent molecules and hydrophilic groups in sugar moieties for gal-PAS compared to glc-PAS. This simulation result is consistent with our experimental observations that gal-PASs are substantially more water-soluble than glc-PASs at all chain lengths. Results from $^{1}H,^{13}C$-HSQC and $^{1}H,^{1}H$-NOESY 2D-NMR studies of glc-PAS and gal-PAS 10-mer samples correlated with the MD simulations, further supporting the validity of our AA-models. Finally, end-to-end distance and radius of gyration measurements suggest that glc- and gal-PAS structures possess similar chain lengths and conformational fluctuations.

Overall, these results highlight the importance of the PAS repeat unit structural rigidity in promoting a stable helix that does not depend on *inter*-residue hydrogen bonding along the backbone. Additionally, the minor stereochemical change of the hydroxyl group at the C_4 position is responsible for substantial differences between glc- and gal-PAS water solubility properties based on the greater ability of gal-PASs to interact with solvent water molecules. Results from these studies demonstrate that our computational analysis and experimental observations are in agreement and validate our AA-models as reliable representations of the glc- and gal-PAS systems.

The outcome of these studies will contribute to a more thorough and insightful understanding of the relationship between composition, structure, and chemical properties of PASs. This modified CHARMM force field and MD simulation provide a versatile tool that can be applied toward other PAS systems consisting of different pyranose rings, different degrees of polymerization, structural modifications (i.e., with amphiphilic lipid groups), and, more importantly, to other synthetic polysaccharides. Additionally, these findings emphasize the importance of complementing theoretical findings with experimental results as a method to verify results, as well as the advantages of using both approaches to obtain a more comprehensive understanding of macromolecular systems.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b01837.

Experimental and instrument details, supplemental figures and tables, NMR spectra, and IR spectra (PDF)

AUTHOR INFORMATION

Corresponding Authors

*straub@bu.edu

*mgrin@bu.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Ruiqing Xiao for insightful discussions on the NMR studies of the PASs, Mr. George Pantelopulos for help preparing the initial structures for MD simulations, and Dr. Afra Panahi for discussion with the beta-PAS structure. C.J.M. would like to acknowledge the NIH Shared Instrument Grant S10 OD011941 "A Console Upgrade and Cryogenic Probe for a 500 MHz NMR System for Biomedical Research." E.L.D. acknowledges receipt of an NIH/NIGMS Postdoctoral Fellowship (1F32GM097781). J.E.S. is grateful to the National Science Foundation (Grant No. CHE-1362524) for support. M.W.G. is grateful for support from Boston University.

REFERENCES

(1) (a) Iha, R. K.; Wooley, K. L.; Nyström, A. M.; Burke, D. J.; Kade, M. J.; Hawker, C. J. *Chem. Rev.* **2009**, *109*, 5620. (b) Nyström, A. M.; Wooley, K. L. *Acc. Chem. Res.* **2011**, *44* (10), 969. (c) Jaeger, K. E.; Eggert, T. *Curr. Opin. Biotechnol.* **2002**, *13* (4), 390. (d) Oh, J. K.; Drumright, R.; Siegwart, D. J.; Matyjaszeqski, K. *Prog. Polym. Sci.* **2008**, *33* (4), 448. (e) Kim, B. S.; Mooney, D. J. *Trends Biotechnol.* **1998**, *16* (5), 224. (f) Van Vlierberghe, S.; Dubruel, P.; Schacht, E. *Biomacromolecules* **2011**, *12* (5), 1387. (g) Sokolsky-Papkov, M.; Agashi, K.; Olaye, A.; Shakesheff, K.; Domb, A. J. *Adv. Drug Delivery Rev.* **2007**, *59* (4−5), 187. (h) Martina, M.; Hutmacher, D. W. *Polym. Int.* **2007**, *56* (2), 145. (i) Chilkoti, A.; Dreher, M. R.; Meyer, D. E. *Adv. Drug Delivery Rev.* **2002**, *54* (8), 1093. (j) Vroman, I.; Tighzert, L. *Materials* **2009**, *2* (2), 307. (k) Rehm, B. H. A. *Nat. Rev. Microbiol.* **2010**, *8* (8), 578. (l) Ghobril, C.; Grinstaff, M. W. *Chem. Soc. Rev.* **2015**, *44* (7), 1820. (m) Cameron, N.; Deming, T. *Macromol. Biosci.* **2015**, *15* (1), 7. (n) Ricapito, N. G.; Ghobril, C.; Zhang, H.; Grinstaff, M. W.; Putnam, D. *Chem. Rev.* **2016**, *116* (4), 2664.

(2) Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4* (1), 5.

(3) Ray, A.; Norden, B. *FASEB J.* **2000**, *14* (9), 1041.

(4) Abes, R.; Arzumanov, A. A.; Saleh, A. F.; Said Hassane, F.; Gait, M. J.; Lebleu, B. *Methods Mol. Biol.* **2011**, *683*, 307.

(5) Seo, Y. J.; Lim, J.; Lee, E. H.; Ok, T.; Yoon, J.; Lee, J. H.; Lee, H. S. *Nucleic Acids Res.* **2011**, *39* (16), 7329.

(6) Koshkin, A. A.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Singh, S. K.; Wengel, J. *J. Am. Chem. Soc.* **1998**, *120* (50), 13252.

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540

(7) Condon, D. E.; Yildirim, I.; Kennedy, S. D.; Mort, B. C.; Kierzek, R.; Turner, D. H. *J. Phys. Chem. B* **2014**, *118* (5), 1216.
(8) Abes, S.; Ivanova, G. D.; Abes, R.; Arzumanov, A. A.; Williams, D.; Owen, D.; Lebleu, B.; Gait, M. J. *Methods Mol. Biol.* **2009**, *480*, 85.
(9) Stern, R.; Jedrzejas, M. J. *Chem. Rev.* **2008**, *108* (12), 5061.
(10) Dwek, R. A. *Chem. Rev.* **1996**, *96* (2), 683.
(11) Nishimura, S.; Matsuoka, K.; Furuike, T.; Ishii, S.; Kurita, K.; Nishimura, K. M. *Macromolecules* **1991**, *24* (15), 4236.
(12) Kiessling, L. L.; Grim, J. C. *Chem. Soc. Rev.* **2013**, *42* (10), 4476.
(13) Kiessling, L. L.; Splain, R. A. *Annu. Rev. Biochem.* **2010**, *79*, 619.
(14) Rabuka, D.; Forstner, M. B.; Groves, J. T.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130* (18), 5947.
(15) Godula, K.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132* (29), 9963.
(16) Metzke, M.; Guan, Z. *Biomacromolecules* **2008**, *9* (1), 208.
(17) (a) Wathier, M.; Stoddart, S. S.; Sheehy, M. J.; Grinstaff, M. W. *J. Am. Chem. Soc.* **2010**, *132* (45), 15887. (b) Mikami, K.; Lonnecker, A. T.; Gustafson, T. P.; Zinnel, N. F.; Pai, P. K.; Russell, D. H.; Wooley, K. L. *J. Am. Chem. Soc.* **2013**, *135*, 6826.
(18) Ernst, B.; Magnani, J. L. *Nat. Rev. Drug Discovery* **2009**, *8* (8), 661.
(19) Dane, E. L.; Grinstaff, M. W. *J. Am. Chem. Soc.* **2012**, *134* (39), 16255.
(20) Dane, E. L.; Chin, S. L.; Grinstaff, M. W. *ACS Macro Lett.* **2013**, *2* (10), 887.
(21) López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5* (12), 3195.
(22) Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5* (9), 2353.
(23) (a) Reiling, S.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem.* **1996**, *17* (4), 450. (b) Melberg, S.; Rasmussen, K. *Carbohydr. Res.* **1979**, *69* (1), 27. (c) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27* (12), 1927. (d) Grootenhuis, P. D. J.; Haasnoot, C. A. G. *Mol. Simul.* **1993**, *10* (2−6), 75. (e) Glennon, T. M.; Zheng, Y.-J.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15* (9), 1019.
(24) Matthews, J. F.; Beckham, G. T.; Bergenstråhle-Wohlert, M.; Brady, J. W.; Himmel, M. E.; Crowley, M. F. *J. Chem. Theory Comput.* **2012**, *8* (2), 735.
(25) Kuttel, M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2002**, *23* (13), 1236.
(26) (a) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2011**, *7* (10), 3162. (b) Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comput.* **2009**, *5*, 2353.
(27) Mallajosyula, S. S.; Guvench, O.; Hatcher, E.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8* (2), 759.
(28) (a) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2008**, *29* (15), 2543.
(29) Hatcher, E.; Guvench, O.; MacKerell, A. D. *J. Phys. Chem. B* **2009**, *113* (37), 12466.
(30) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. *J. Comput. Chem.* **2005**, *26* (16), 1701.
(31) Lindahl, E.; Hess, B.; Van Der Spoel, D. *J. Mol. Model.* **2001**, *7* (8), 306−317.
(32) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1996**, *79* (4), 913.
(33) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101* (10), 3219.
(34) Choi, S. H.; Guzei, I. A.; Spencer, L. C.; Gellman, S. H. *J. Am. Chem. Soc.* **2010**, *132* (43), 15456.
(35) Kubelka, J.; Huang, R.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109* (16), 8231.
(36) Stidham, S. E.; Chin, S. L.; Dane, E. L.; Grinstaff, M. W. *J. Am. Chem. Soc.* **2014**, *136* (27), 9544.
(37) Mandity, I. M.; Fulop, L.; Vass, E.; Toth, G. K.; Martinek, T. A.; Fulop, F. *Org. Lett.* **2010**, *12* (23), 5584.
(38) Gardiner, J.; Mathad, R.; Jaun, B.; Schreiber, J. r.; Flögel, O.; Seebach, D. *Helv. Chim. Acta* **2009**, *92* (12), 2698.
(39) Bulheller, B. M.; Hirst, J. D. *Bioinformatics* **2009**, *25* (4), 539.
(40) Cerutti, N.; Killick, M.; Jugnarain, V.; Papathanasopoulos, M.; Capovilla, A. *J. Biol. Chem.* **2014**, *289* (15), 10455.
(41) Majumdar, R.; Railkar, R.; Dighe, R. R. *J. Biol. Chem.* **2012**, *287* (41), 34514.
(42) Xiong, M.; Lee, M. W.; Mansbach, R. A.; Song, Z.; Bao, Y.; Peek, R. M., Jr.; Yao, C.; Chen, L. F.; Ferguson, A. L.; Wong, G. C.; Cheng, J. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (43), 13155.
(43) Huyghues-Despointes, B. M. P.; Pace, C. N.; Englander, S. W.; Scholtz, J. M. In *Protein Structure, Stability, and Folding*, Murphy, K. P., Ed.; Humana Press Inc: Totowa, NJ, 2001; Vol. *168*.
(44) Fisher, B. F.; Guo, L.; Dolinar, B. S.; Guzei, I. A.; Gellman, S. H. *J. Am. Chem. Soc.* **2015**, *137* (20), 6484.
(45) Massiah, M. A.; Ko, Y. H.; Pedersen, P. L.; Mildvan, A. S. *Biochemistry* **1999**, *38* (23), 7453.
(46) Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D.; et al. *Helv. Chim. Acta* **1996**, *79*, 2043.
(47) Doi, M.; Edwards, S. F. *The Theory of Poymer Dynamic*; Clarendon Press: Oxford, 1986.
(48) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press Inc.: New York, 2003.
(49) Kirby, B. J. *Micro- and Nanoscale Fluid Mechanics: Transport in Microfluidic Devices*; Cambridge University Press; New York, 2010.

DOI: 10.1021/jacs.6b01837
J. Am. Chem. Soc. 2016, 138, 6532−6540