

Intrinsic Conformational Dynamics of Alanine in Water/Ethanol Mixtures: An Experiment-Driven Molecular Dynamics Study

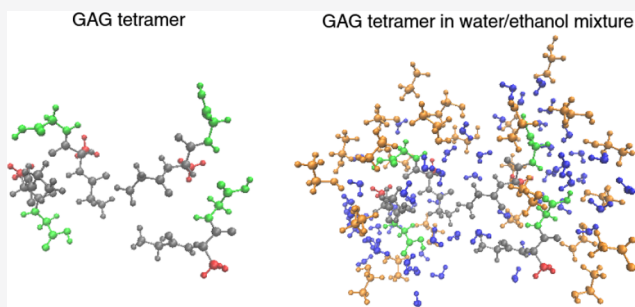
Shuting Zhang, Brian Andrews, Reinhard Schweitzer-Stenner, and Brigita Urbanc*

 Cite This: <https://dx.doi.org/10.1021/acs.jpcb.0c08245> Read Online

ACCESS |

 Metrics & More Article Recommendations Supporting Information

ABSTRACT: *In vitro*, cationic glycylalanylglycine (GAG) forms a hydrogel in binary mixtures of water and ethanol. In water, alanine residue is known for its high polyproline II (pPII) content. Spectroscopic data, including three J-coupling constants and amide I' profiles, indicate that addition of 42% ethanol to water significantly reduces the pPII content of alanine residue in GAG. Here, experiment-based Gaussian Ramachandran distributions of alanine in GAG at different ethanol fractions are examined and three MD force fields are evaluated with respect to their ability to capture these ethanol-induced conformational changes. MD simulations on monomeric GAG in eight different water/ethanol mixtures within Amber ff14SB, OPLS-AA/M, and CHARMM36m reveal that only Amber ff14SB partially captures the ethanol-induced conformational changes of alanine residue in monomeric GAG when 42% ethanol is added to water. MD simulations of 200 mM GAG ensembles in pure water and in the aqueous solution with 42% ethanol showcase the ability of CHARMM36m to capture the effect of ethanol on the average pPII content of alanine in GAG and provide a plausible explanation for this effect, which may stem from an increased propensity of GAG to form oligomers in the presence of ethanol.



INTRODUCTION

Short peptides are great model systems for exploring intrinsic conformational propensities of amino acid residues in water and other solvents.^{1–6} Substantial evidence accumulated over the past decade shows that intrinsic conformational ensembles of natural amino acid residues in water sample a significantly more restricted region of the Ramachandran space than would be expected based on sterical considerations. These findings are important to our understanding of unfolded as well as intrinsically disordered proteins.^{7–12}

Alanine-based peptides have attracted a lot of interest because several lines of experimental evidence indicate an extraordinary high polyproline II (pPII) propensity,^{13–18} which was previously associated only with trans-proline residues and their nearest glycine neighbors in polyproline peptides or in rather inflexible proteins like collagen.^{19–21} The discovery of this conformational disposition was also unexpected because dialanine peptide has been considered a canonical model system showcasing random sampling of the sterically allowed region of the Ramachandran space.^{1,22–24} Molecular dynamics (MD) studies of short peptides provide some evidence that water plays a major role in stabilizing pPII.^{25–30} While this notion has gained a lot of support (with noteworthy exceptions^{31,32}), the mechanism of how water stabilizes pPII has not been fully established. We recently examined a conformational ensemble of alanine in cationic GAG in water and assessed the capability of six MD force fields from three different series (Amber, CHARMM, and OPLS) to

reproduce experimentally obtained amide I' profiles and J-coupling constants.³³ Among the six employed force fields, Amber ff14SB³⁴ combined with a TIP3P water model³⁵ best reproduced experimental results despite some notable discrepancies.³³ Importantly, this study demonstrated that pPII is a mesostate that allows water to form the most hydrogen bonds (HBs) with functional groups of alanine residue, affirming the central role of water in pPII stabilization.³³ Moreover, a subsequent experiment-driven MD study on conformational dynamics of the central glycine residue in GGG in water and other solvents revealed that the high pPII content does not originate in the peculiar side chain properties of alanine but rather reflects the affinity of water to form hydrogen bonds with the functional backbone groups.³⁶

If water plays such a prominent role in the selection of the conformational ensemble of alanine residue, what happens when cosolvents are added to water? Milorey et al. investigated the effect of ethanol as a cosolvent on the conformational ensemble of the alanine residue in cationic GAG et al. by measuring the temperature dependence of the $^3J(H^N, H^C\alpha)$

Received: September 9, 2020

Revised: November 23, 2020

coupling constant for different water/ethanol mixtures with up to 48% ethanol.³⁷ Experiments on samples with higher ethanol content were not possible due to an unexpected sol $\rightarrow$ gel transition observed in ternary GAG/water/ethanol samples with the ethanol fraction of 55% and above a peptide concentration of ~ 200 M.^{37,38} The sol–gel boundary depends on the peptide concentration and ethanol fraction, as reported by DiGiuseppi and collaborators.³⁹ Below this transition, a systematic but nonmonotonic increase of the $^3J(H^N, H^{\alpha})$ coupling constant with increasing ethanol fraction was observed. Milorey et al. identified three regions of ethanol fractions, referred to as I, II, and III, associated with abrupt changes in the $^3J(H^N, H^{\alpha})$ value, which indicate highly nonideal water/ethanol mixtures.³⁷ The increase of this J-coupling constant was tentatively assigned to a relative stabilization of β -strand over pPII conformations.³⁷ In a follow-up study, DiGiuseppi et al. measured and analyzed amide I' profiles of IR, isotropic, and anisotropic Raman and vibrational circular dichroism (CD) spectra of cationic GAG in different water/ethanol mixtures associated with regions I, II, and III.⁴⁰ The application of the Gaussian model, in which each state in the Ramachandran distribution is assigned to a Gaussian subdistribution, followed by optimization of propensities, locations, and widths of subdistributions so as to find the best match to experimental data, showed that an increasing ethanol concentration stabilizes β -strand over pPII.⁴⁰ Further analysis of the effect of ethanol on amide I' wavenumbers, chemical shifts, and J-coupling constants of amide protons revealed a rather complicated picture, leading to the hypothesis that GAG accumulates at interfaces between water and ethanol clusters with the C-terminal half of the peptide immersed in ethanol and the N-terminal half exposed to water.⁴⁰

Motivated by the above experimental findings,^{37,38,40} we here examine the effect of ethanol on intrinsic conformational dynamics of alanine in cationic GAG. On the basis of findings of our previous study,³³ we use Amber ff14SB, OPLS-AA/M,⁴¹ and CHARMM36m⁴² to explore conformational dynamics of alanine in GAG in various water/ethanol mixtures. The overarching aim of this study is to evaluate the three MD force fields (alongside their respective water and ethanol models) with respect to their ability to reproduce the experimental data on cationic GAG as well as capture the nonideality of water/ethanol mixtures and preferential binding of water and ethanol to the N- and C-terminal end of the peptide, respectively. We perform MD simulations on a single GAG peptide in several binary water/ethanol mixtures as well as simulations of 16 GAG peptides (corresponding to a peptide concentration of 200 mM), including those used in the above experimental studies.^{37,38,40} The MD-derived Ramachandran distributions of alanine in cationic GAG are used to calculate the J-coupling constants and amide I' profiles to facilitate a direct comparison to the experimental data and the respective Gaussian modeling. We also explore the interaction between solvent molecules and functional groups of the peptide backbone to probe the specific roles of water and ethanol in peptide solvation and self-assembly.

METHODS

Molecular Dynamics Simulations. MD simulations of a single GAG peptide in explicit solvent are performed within GROMACS 5.1.2^{43–49} using three MD force fields: Amber ff14SB,³⁴ OPLS-AA/M,⁴¹ and CHARMM36m.⁴² TIP3P water

model³⁵ is used in Amber ff14SB simulations. A modified version of TIP3P water model, which was developed alongside the force field parametrization, is used in CHARMM36m simulations. In OPLS-AA/M simulations, we use the TIP4P water model,³⁵ which was shown to be optimal for studies of conformational dynamics of alanine residue in GAG.³³ As ethanol is a small organic molecule, its force field parameters are included in the General Amber Force Field (GAFF),^{50,51} OPLS-AA general force field,⁵² and CHARMM General Force Field (CGenFF).^{53–57} The ability of these three general force fields to reproduce key ethanol properties, such as the density, enthalpy of vaporization, the surface tension, the heat capacity at constant volume and pressure, the isothermal compressibility, the volumetric expansion coefficient, and the static dielectric constant, have been shown to be comparable among the three general force fields, with CGenFF ranking somewhat above and GAFF ranking a bit below OPLS-AA.⁵⁸ The GAG peptide is capped with protonated N-terminus (NH^{3+}) and neutral C-terminus ($-\text{CO}-\text{NH}_2$ for Amber ff14SB and $-\text{COOH}$ for OPLS-AA/M and CHARMM36) to mimic the acidic condition of NMR experiments. For each simulation, a single GAG peptide is solvated in a cubic box of $40 \times 40 \times 40 \text{ \AA}^3$ (~ 26 mM). The numbers of water and ethanol molecules for each water/ethanol mixture used in simulations are listed in Table S1. The energy minimization involves steepest descent minimization for 100,000 steps, followed by a 20 ps-long pressure equilibration at 300 K and 1.0 bar. One Cl^- is added to each simulation box to make the net electrostatic charge zero. Each 100 ns long trajectory is acquired using the velocity rescale thermostat⁵⁹ and Berendsen barostat.⁶⁰

In addition to simulations of monomeric GAG, MD simulations of ensembles of 16 GAG in a $50 \times 50 \times 50 \text{ \AA}^3$, corresponding to ~ 220 mM peptide concentration, are conducted within each of the three force fields. For each MD force field, two 300 ns-long trajectories of 200 mM GAG ensemble are acquired, one in pure water and the other in the aqueous solution with 42% ethanol. The numbers of water and ethanol molecules for each type of 200 mM GAG simulation are listed in Table S1. Due to the larger size of the system, the pressure equilibration step is extended from 20 to 200 ps at 300 K and 1.0 bar. Sixteen (16) Cl^- ions are added to each simulation box to create an electrically neutral system. Other details are identical to those described in the monomeric simulations.

Analysis. Gaussian Model. Gaussian modeling, introduced by Schweitzer-Stenner,⁶¹ produces a 2D distribution of backbone dihedral angles ϕ and ψ , i.e., Ramachandran distribution, which is assumed to be a linear superposition of multiple Gaussian subdistributions:

$$P(\phi_i, \psi_j) = \sum_k \chi_k G_k(\phi_i, \psi_j) \quad (1)$$

where G_k is the Gaussian subdistribution of the k th state with the maximum $\phi_{k,\text{max}}$ and $\psi_{k,\text{max}}$, half-widths $\sigma_{k,\phi}$ and $\sigma_{k,\psi}$, and propensities χ_k . The locations of the maxima, half-widths, and propensities of Gaussian subdistributions are calibrated to provide the best fit to spectroscopic data.⁶¹

Definition of Mesostates. The following definitions of mesostates (see Figure 1) are used in the analysis: (a) polyproline II (pPII) ($-90^\circ < \phi < -42^\circ$ and $100^\circ < \psi < 180^\circ$), (b) antiparallel β -strand ($\alpha\beta$) ($-180^\circ < \phi < -130^\circ$ and $130^\circ < \psi < 180^\circ$), (c) the transition region between $\alpha\beta$ and pPII (βt) ($-130^\circ < \phi < -90^\circ$ and $130^\circ < \psi < 180^\circ$), (d)

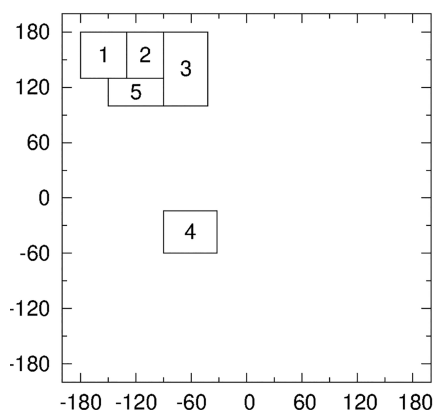


Figure 1. Definition of the five mesostates in the Ramachandran plot. The regions 1–5 correspond to $\alpha\beta$, $\beta\beta$, pPII, right-handed α -helix, and $\beta\beta$, respectively.

right-handed helix ($-90^\circ < \phi < -32^\circ$ and $-60^\circ < \psi < -14^\circ$), and (e) parallel β -strand ($\beta\beta$) ($-150^\circ < \phi < -90^\circ$ and $100^\circ < \psi < 130^\circ$).

MD-Derived Ramachandran Distributions. Ramachandran plots are constructed from dihedral angles (ϕ , ψ) within GROMACS 5.1.2 using MD trajectory between 50 and 100 ns with time frames 2 ps apart, resulting in 25000 time frames in total. Normalized two-dimensional (2D) distributions are calculated with a bin size of $2^\circ \times 2^\circ$, resulting in $180 \times 180 = 32400$ bins to facilitate the comparison to Ramachandran distributions of Gaussian models (a) and (b). The color scale in Ramachandran plots corresponds to a probability that alanine residue adopts dihedral (ϕ and ψ) angles within a bin, such that the sum of these per-bin probabilities over all bins is equal to 1.

Calculation of J-Coupling Constants and Amide I' Profiles from MD Data. As we did in the previous work on GAG and AAA in pure water,³³ the J-coupling constants and the VCD band profile, which are the most sensitive to the sampling of the upper left quadrant of the Ramachandran plot, are used to examine the ability to reproduce experimental results of the Gaussian models and MD force fields. MD-derived Ramachandran distributions are used to calculate the J-coupling constants and amide I' profiles for alanine residue in GAG in various water/ethanol mixtures. For any given dihedral angle distribution, $P(\phi_i, \psi_i)$, the J-coupling constants are calculated using Karplus equations with previously reported Karplus parameters.^{15,62} The Karplus parameters and corresponding uncertainty are the same as in our previous work.³³

Can the Karplus parameters used for proteins in water be also applied to GAG in water/ethanol mixtures? First, published work on solvent effects has focused mostly on fluorinated molecules, which are characterized by a high electron withdrawing capability and are thus highly polar and easily polarizable. J-coupling constants involving fluorine are huge (>100 Hz) and the corresponding solvent-induced changes in the order of $10^{-3}\%$, whereas the respective changes of the chemical shift are on the order of $10^{-2}\%$.⁶³ Peptide groups should have a much lower polarizability than any fluorinated molecule. Second, the chemical shift of the amide I proton of GAG in water/ethanol mixtures, more specifically, its N-terminal region, which is more hydrated than the C-terminal region, decreases with increasing ethanol fraction and reaches a minimum below a mole fraction of 0.1. This corresponds to

the microcluster region or ethanol and most likely reflects a slight change of the water structure.⁴⁰ Upon increasing the ethanol fraction further, the chemical shift moves downfield and reaches its water value around 0.4. The maximal change is $\sim 10^{-3}\%$, which is an order of magnitude smaller than the changes observed for fluorinated molecules. The ethanol dependence of the chemical shift does not correlate with the rather noncontinuous behavior of $^3J(H^N, H^{C\alpha})$, which varies up to 8%. Third, a change of the reaction field is known to cause a change of the amide I wavenumber.⁶⁴ However, as shown by Milorey et al.³⁷ and DiGiuseppi et al.,⁴⁰ the N-terminal wavenumber hardly changes upon addition of ethanol to water. These considerations suggest that the values of the Karplus parameters determined for proteins in water are applicable to GAG in water/ethanol mixtures as well.

For each MD-derived or Gaussian model (ϕ , ψ) distribution $P(\phi, \psi)$, if $f(\phi_i, \psi_i)$ describes the dependence of the amide I' profile or J-coupling constant on the two dihedral angles, then the average value is calculated as

$$\langle f \rangle = \sum_{i,j} f(\phi_i, \psi_j) P(\phi_i, \psi_j) \quad (2)$$

With regard to J-coupling constants, we directly compare the calculated and experimental values and further quantify this comparison by the χ^2 function:

$$\chi^2_J = \frac{1}{N} \sum_{i=1}^N \frac{(J_{i,\text{exp}} - J_{i,\text{calc}})^2}{s_i^2} \quad (3)$$

where $J_{i,\text{exp}}$ and $J_{i,\text{calc}}$ denote experimental and calculated J-coupling constants, respectively. N is the number of considered J-coupling constants, whereas s_i are the corresponding statistical errors. We calculate the latter as a combination of experimental errors¹⁷ and errors due to the reported uncertainties of the Karplus parameters^{62,65} using Gaussian error propagation. Statistical uncertainties s_i are ensemble averages of the following function $\bar{s}_i(\phi)$

$$\bar{s}_i(\phi) = \sqrt{s_{A_i}^2 \cos^4(\phi + \theta_i) + s_{B_i}^2 \cos^2(\phi + \theta_i) + s_{C_i}^2 + s_{J_i}^2} \quad (4)$$

where s_{A_i} , s_{B_i} , and s_{C_i} are the statistical uncertainties of the three Karplus parameters, θ_i is the phase associated with the J-coupling constant J_i , s_{J_i} is the corresponding experimental error (as previously reported for GAG¹⁷), and $s_i = \langle \bar{s}_i \rangle$ whereby the ensemble average $\langle \rangle$ is defined by eq 2. s_{A_i} , s_{B_i} , and s_{C_i} values are reported by Wang and Bax as we used in the previous work.⁶⁵ For $^1J(N, C_\alpha)$, the statistical errors of two Karplus parameters are not known, thus only the experimental errors are included in the calculation.

Similarly, the reproduction of the VCD band profile is assessed by

$$\chi^2_{\text{VCD}} = \frac{1}{N'} \sum_{k=1}^{N'} \frac{(\Delta\epsilon_{\text{exp},k} - \Delta\epsilon_{\text{calc},k})^2}{s_k^2} \quad (5)$$

where $\Delta\epsilon$ represents the molar dichroism and k labels the different data point. N' is the number of data points, and the standard deviation s_k is derived from an analysis of a spectral region dominated by noise.

Solvation Layer Properties around Alanine Residue. The numbers of ethanol and water molecules within the solvation layer around the GAG alanine backbone and the numbers of

HBs formed by GAG alanine are counted for all time frames within 50–100 ns that are 2 ps apart, resulting in 25000 frames in total. Considering the size of the ethanol molecule, 5 Å is used as the thickness of the solvation layer. An ethanol or a water molecule is considered to be within the solvation layer if its center of mass is within a distance of 5 Å from the peptide. Hydrogen bonding is defined by a cutoff distance of 3 Å between a donor and an acceptor, and an upper limit of 20° is used for the cutoff hydrogen–donor–acceptor angle. The numbers of HBs between the peptide backbone and the solvent are counted individually for amide and carbonyl groups as well as for each of the three amino acid residues in GAG. The number of HBs is averaged over all frames, and the corresponding standard error of the mean (SEM) values are calculated.

Radial Distribution Function. Two radial distribution functions of water and ethanol oxygen atoms to the three C_α atoms of GAG, $g_{C_\alpha-O}(r)$, are calculated within GROMACS 5.1.2. These distribution functions show the number density of water and ethanol oxygen atoms in dependence on the distance r from C_α atoms of GAG. Time frames between 50 and 100 ns that are 2 ps apart are used for this calculation.

Water and Ethanol Orientation Plots. Using Visual Molecular Dynamics (VMD) software package,⁶⁶ a solvent orientation plot is calculated for those water or ethanol molecules whose center of mass is within a distance of 5 Å from alanine backbone atoms. The orientation of each water molecule is defined by two angles (η , θ) as reported previously.^{30,33} The local normal to the solvent accessible surface (SAS) of the backbone atoms of alanine residue, $\hat{n}$, is derived for each water molecule using three proximate points on the triangulated SAS area of the side chain. For each water molecule, the water symmetry axis and the normal axis to the plane defined by the three atoms in the water molecule (water plane) are calculated. The two angles (η , θ) are defined as the angle between $\hat{n}$ and water symmetry axis, and the angle between $\hat{n}$ and the normal axis to the water plane, respectively. The orientation of the ethanol molecule follows the definition of water orientation, whereby the two axes are defined with respect to the three ethanol atoms, the central carbon and the two atoms in hydroxyl group. The 2D distribution of (η , θ) angles for water or ethanol molecules in the solvation layer adjacent to the backbone atoms of alanine residue in GAG, i.e., water or ethanol orientation plot, is calculated using $90 \times 90 = 8,100$ bins in the (η , θ) space.

Ethanol Cluster Size Distribution. Two ethanol molecules are considered to belong to the same cluster if the distance between any two atoms belonging to different ethanol molecules is within 2.5 Å. For each time frame between 50 and 100 ns of each trajectory (2 ps apart), the histogram of ethanol cluster masses is derived by recording the mass of each ethanol cluster and all per-frame histograms are averaged over all time frames. The ethanol cluster size distribution is then obtained from the averaged histogram by normalizing by the total mass of all ethanol molecules.

Size Distribution of GAG Oligomers. Two peptides are considered to belong to the same oligomer if the distance between any two atoms belonging to different GAG peptides is within 3.5 Å. For each time frame between 250 and 300 ns of each MD trajectory containing 16 GAG peptides, the histogram of GAG oligomer sizes is derived by recording the size of each GAG oligomer and counting the respective

occurrence numbers for each size. The resulting per-frame histograms are averaged over all time frames. The resulting GAG oligomer size distribution is calculated by normalizing the averaged histogram by the total number of peptides (16) so as to obtain the probability of a GAG peptide to be within an oligomer of a specific size.

RESULTS AND DISCUSSION

As Farrell et al. reported, cationic GAG in an aqueous solution with 55 mol % ethanol can form a sample-spanning network of extraordinary long crystalline fibrils with yet unknown structure.³⁸ Motivated by this puzzling behavior and availability of spectroscopic data, i.e., three J-coupling constants $^3J(H^N, H^{C_\alpha})$, $^3J(H^N, C')$, and $^1J(N, C')$ as well as amide I' profiles, on GAG in water/ethanol mixtures at molar ethanol fractions of 3%, 12%, and 42%, we here explore the conformational ensemble of alanine residue in a cationic GAG peptide immersed in water/ethanol mixtures by MD simulations. Hereafter, the ethanol fraction refers to the molar fraction of ethanol in the aqueous solution.

This section is organized as follows. We first revisit the Gaussian model-based analysis of the amide I' band and J-coupling constants of alanine residue in GAG for three different water/ethanol mixtures in order to better reproduce one of the J-coupling constants ($^1J(N, C')$), which was overestimated in the initial analysis by DiGuiseppi and collaborators.⁴⁰ Following Milorey et al., we hereafter refer to the ethanol fractions regions 0%–3%, 12%–18%, and 32%–43%, which show abrupt changes in $^3J(H^N, H^{C_\alpha})$, as regions I, II, and III, respectively. We then describe the results of MD simulations of cationic GAG in water/ethanol mixtures for three experimentally studied as well as higher ethanol fractions and pure ethanol. MD simulations are performed in three commonly used MD force fields: Amber ff19SB,³⁴ OPLS-AA/M,⁴¹ and CHARMM36m⁴² and their respective water and ethanol models, as described in **Methods**. We assess these force fields with respect to their ability to capture the available spectroscopic data by using MD-derived Ramachandran distributions to calculate amide I' profiles and the three J-coupling constants to facilitate a direct comparison to experimental data. The combined use of the first two J-coupling constants is a good test of the Ramachandran distribution with respect to the ϕ -coordinate because the two corresponding Karplus curves are nearly out-of-phase and have distinct amplitudes, which eliminates the possibility of multiple Gaussian model solutions. The J-coupling constant $^1J(N, C')$ is important because its value solely depends on the ψ -coordinate.⁶⁷ The corresponding Karplus curve is less complex than that of ϕ -dependent J-coupling constants and thus less amenable to ambiguities. We further describe the analysis of solvent-peptide interactions that can be inferred from MD simulations to provide molecular-level insights into the effect of ethanol on conformational dynamics of alanine residue in cationic GAG.

Gaussian Modeling of Alanine Distributions in Water/Ethanol Mixtures. Amide I' profiles (Figure S1) and the three J-coupling constants (Tables S2–S4) of alanine residue in cationic GAG as experimentally determined for three ethanol fractions, 3% (region I), 12% (region II) and 42% (region III),^{37,40} are used as input parameters for Gaussian modeling, which aims to derive the experiment-based Ramachandran distributions, which best reproduce the spectroscopic data. The Gaussian model-based analysis of the

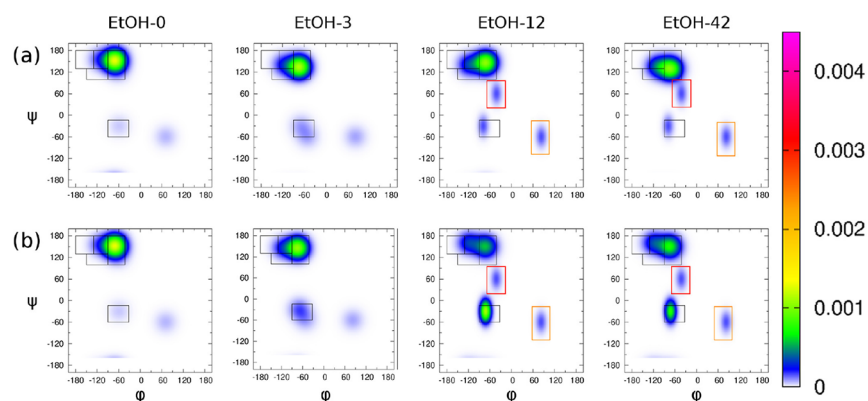


Figure 2. Gaussian modeling-derived Ramachandran plots for alanine residue in cationic GAG. Panels (a) and (b) correspond to Gaussian models (a) and (b), respectively, for GAG in pure water and three aqueous solutions with 3, 12, and 42% ethanol. The black frames correspond to the major five mesostates defined in [Methods](#). The red and orange frames correspond to the inverse and classical γ -turn mesostates, respectively.

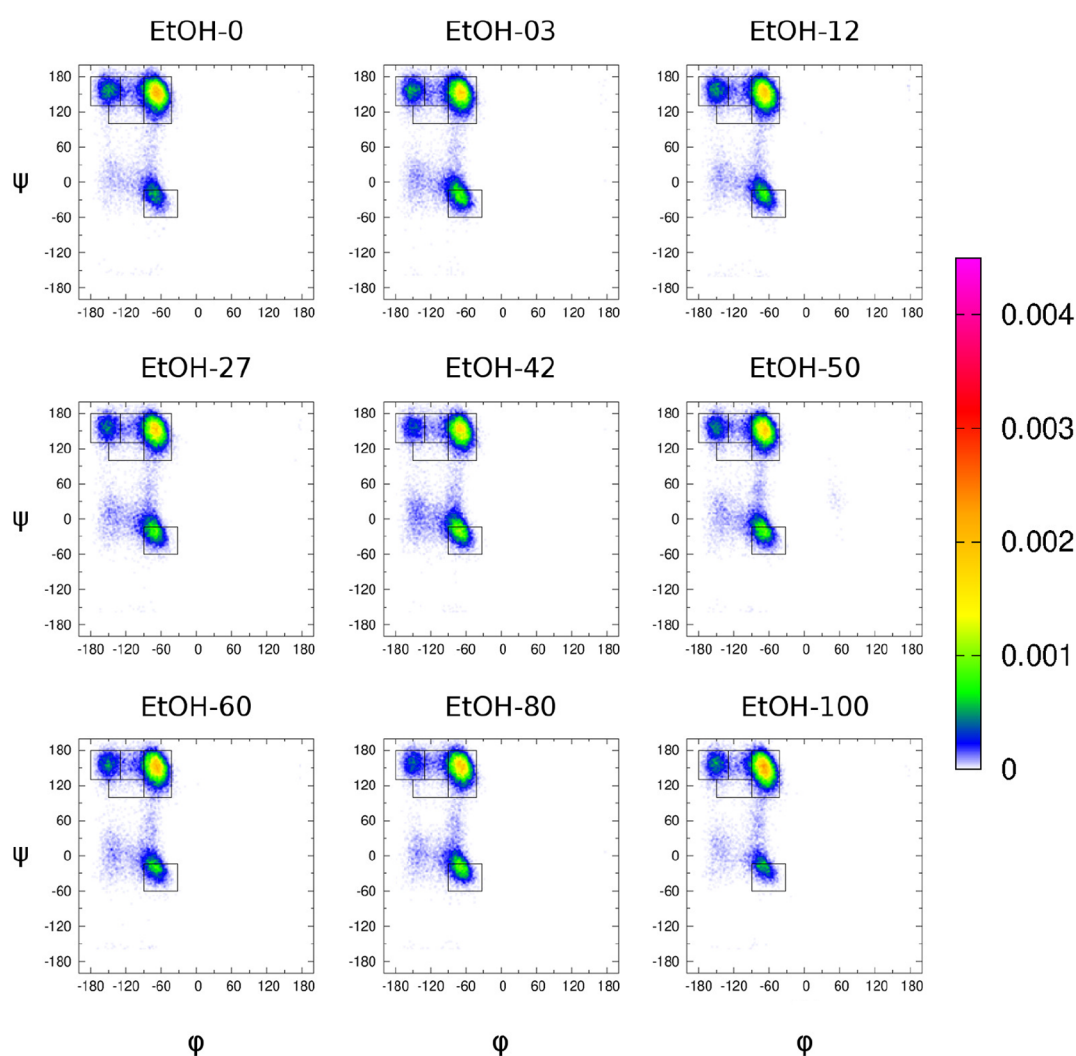


Figure 3. MD-derived Ramachandran plots of alanine residue in GAG: Amber ff14SB. Ramachandran plots of alanine in pure water, several water/ethanol mixtures, and pure ethanol obtained from MD simulations with Amber ff14SB combined with TIP3P water and GAFF ethanol models.

IR, Raman, and VCD profiles conducted by DiGiuseppi et al. revealed that the Gaussian Ramachandran distribution in regions I and II only slightly differs from the respective distribution of alanine residue in GAG in pure water. Ethanol induces a slight decrease in the pPII population alongside a shift in the pPII basin to more negative ϕ -values.⁴⁰ A stronger

effect was observed in region III, where spectroscopic results suggest a stabilization of β -strand at the expense of pPII, whereby the pPII propensity was shown to decrease from 0.72 to 0.55.⁴⁰ Experimental values for the coupling constant $^1J(N,C')$ in water/ethanol mixtures ([Table S4](#)) are systematically lower than the value reported for GAG in water (11.28

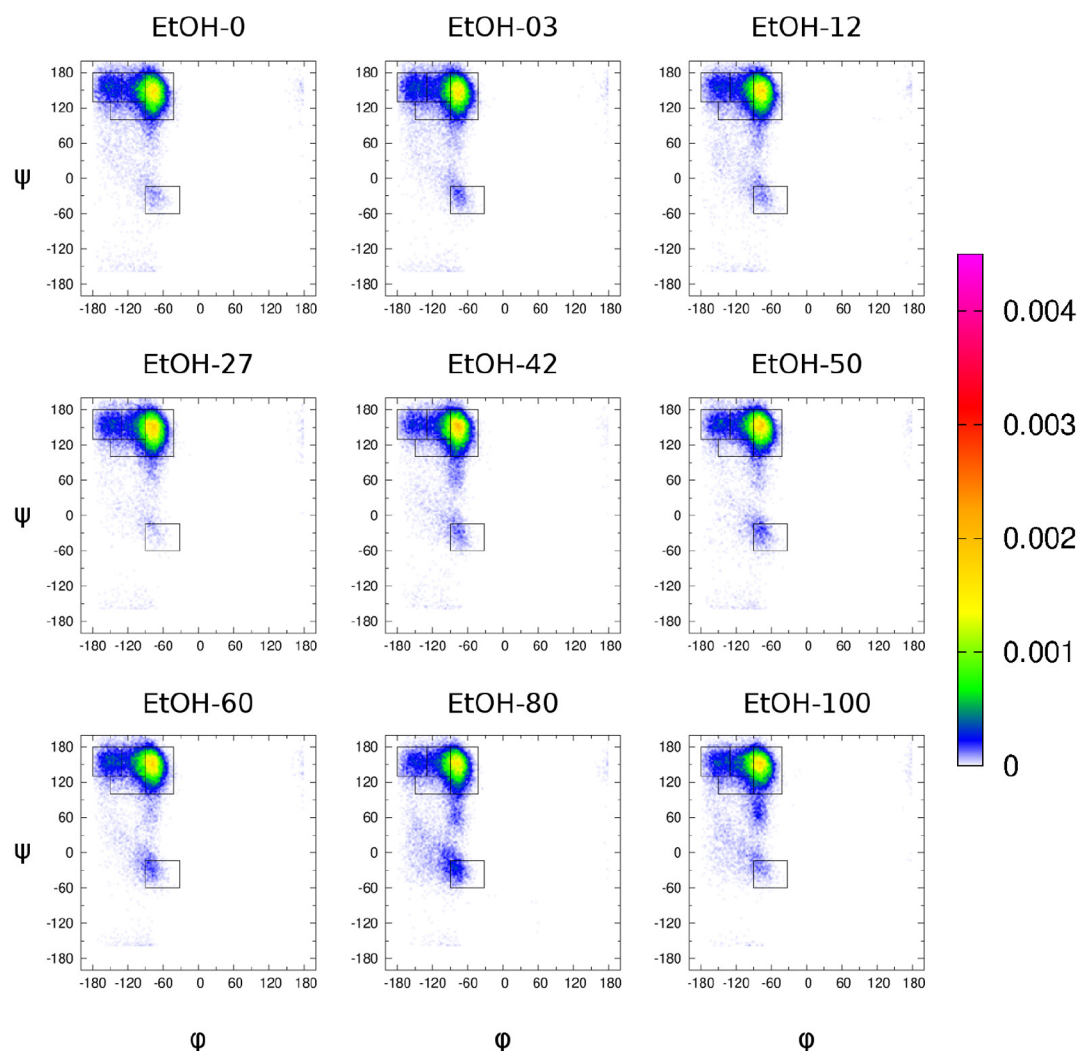


Figure 4. MD-derived Ramachandran plots of alanine residue in GAG: OPLS-AA/M. Ramachandran plots of alanine in pure water, several water/ethanol mixtures, and pure ethanol obtained from MD simulations with OPLS-AA/M combined with TIP4P water and OPLS-AA general force field ethanol models.

Hz). The Gaussian model analysis by DiGuseppi et al. overestimated this J-coupling constant for all three regions I–III.⁴⁰ There are two ways to bring this J-coupling constant closer to its experimental value within the Gaussian model. In option (a), one can decrease the ψ -values of both the pPII and the β -strand subdistributions. In option (b), one can increase the right-handed helical content and/or the inverse γ content of the Ramachandran distribution. We hereafter refer to options (a) and (b) as Gaussian models (a) and (b), respectively. These two models are defined in Table S5, which lists the centers, standard deviations, and propensities of Gaussian subdistributions corresponding to specific mesostates of alanine residue for pure water and aqueous solutions with three different ethanol fractions. Both options (a) and (b) reduce the rotational strength of the amide I' VCD, thereby putting a constraint on the extent to which these modifications can be implemented. With regard to the J-coupling constants, models (a) and (b) perform comparably for alanine residue in aqueous solutions with 3% and 42% ethanol fractions, while model (a) outperforms model (b) for 12% ethanol. Figure S1 shows that the amide I' VCD profiles of model (a) slightly better reproduce the experimental data than the corresponding values produced by model (b). The Ramachandran distribu-

tions associated with Gaussian models (a) and (b) are shown in Figure 2a,b, respectively.

The results of Gaussian model (a) suggest that the addition of ethanol causes a redistribution of the conformational ensemble of alanine residue in cationic GAG from pPII to β -strand. Considering our recently reported Gaussian model analysis for alanine in pure water,³³ this effect is slightly more pronounced than reported earlier.⁴⁰ The pPII propensity drops from 80% in water to 60% in the aqueous solution with 42% ethanol (Table S5). This change is concomitant with the increase in the β -strand population from 10% in pure water to 25% in 42% ethanol-containing solution. The β -strand basins in water/ethanol mixtures remain in the β region while exhibiting lower ψ -values than the pPII basins. Within Gaussian model (b), we find a substantial redistribution from the pPII to the right-handed helical basin and to a lesser extent to the β -strand basin. The pPII propensity drops from 80% in pure water to 30% in the aqueous solution with 42% ethanol. Because the results of UV circular dichroism (CD) spectroscopy are difficult to quantitatively compare to MD data, we did not include them in the previous study, which assessed the MD force fields with respect to conformational dynamics of alanine residue in water.³³ In this work, however, the CD data can help

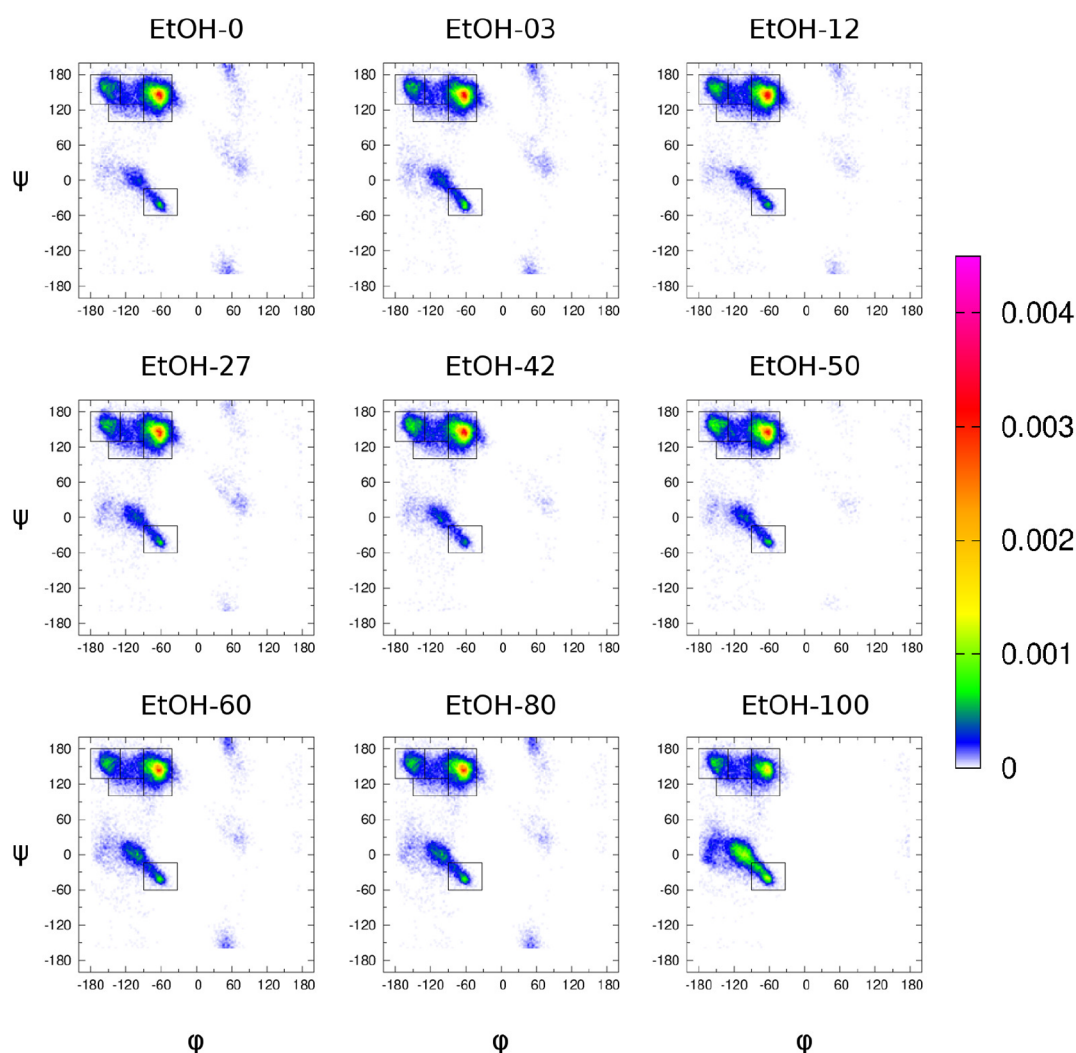


Figure 5. MD-derived Ramachandran plots of alanine residue in GAG: CHARMM36m. Ramachandran plots of alanine in pure water, several water/ethanol mixtures, and pure ethanol obtained from MD simulations with CHARMM36m combined with TIP3P water and CGenFF ethanol models.

assess the two Gaussian models with respect to their ability to capture ethanol-induced changes in the conformational ensemble of alanine. Experimental UV-CD spectroscopy results indicate that the conformational ensemble of alanine undergoes ethanol-induced conformational change between pPII and β -strand mesostates,³⁷ which is more consistent with Gaussian model (a) than with model (b).

MD-Derived Ramachandran Distributions of Alanine Residue in GAG. Assuming monomeric GAG, MD simulations are performed on a single GAG peptide using Amber ff14SB, OPLS-AA/M, and CHARMM36m in different water/ethanol mixtures (see [Methods](#) for details on the force field-specific water and ethanol models). To facilitate a comparison of MD-derived results to experimental data and Gaussian models (a) and (b), MD trajectories of GAG immersed in several solvents: pure water, 97% water/3% ethanol, 88% water/12% ethanol, and 58% water/42% ethanol are acquired. Additional simulations of cationic GAG are performed in 73% water/27% ethanol mixture (between 88% water/12% ethanol and 58% water/42% ethanol mixtures, for which experimental data exist) and in solvents with higher ethanol fractions: 50% water/50% ethanol, 40% water/60% ethanol, 20% water/80% ethanol, and pure ethanol. In

experiments with ethanol fractions above $\sim 50\%$ ethanol, cationic GAG at concentrations of 200 mM and above forms a hydrogel,⁶⁸ precluding a meaningful comparison of MD results to experimental data. The Ramachandran plots of alanine residue in GAG from Amber ff14SB, OPLS-AA/M, and CHARMM36 in different water/ethanol fractions are displayed in [Figures 3, 4, and 5](#), respectively. The rectangular frames correspond to five mesostates defined in [Methods](#) ([Figure 1](#)), which allows us to visually compare the results of MD simulations and the two Gaussian models. Surprisingly, all three MD force fields produce Ramachandran distributions that are not strongly affected by ethanol. The conformations of alanine residue in GAG mainly populate the pPII mesostate, followed by $a\beta$, and right-handed helix regions and the respective distributions obtained in solvents with different ethanol fractions are remarkably similar. The addition of ethanol only slightly modifies the mesostate populations without any obvious ethanol-induced trends.

Of the three MD force fields, only CHARMM36m displays a significant decrease in the pPII population when the ethanol fraction increases from 80% to 100% ([Figure 5](#)). The locations of the respective mesostate basins are also not affected by the presence of ethanol. In OPLS-AA/M, the pPII basins of

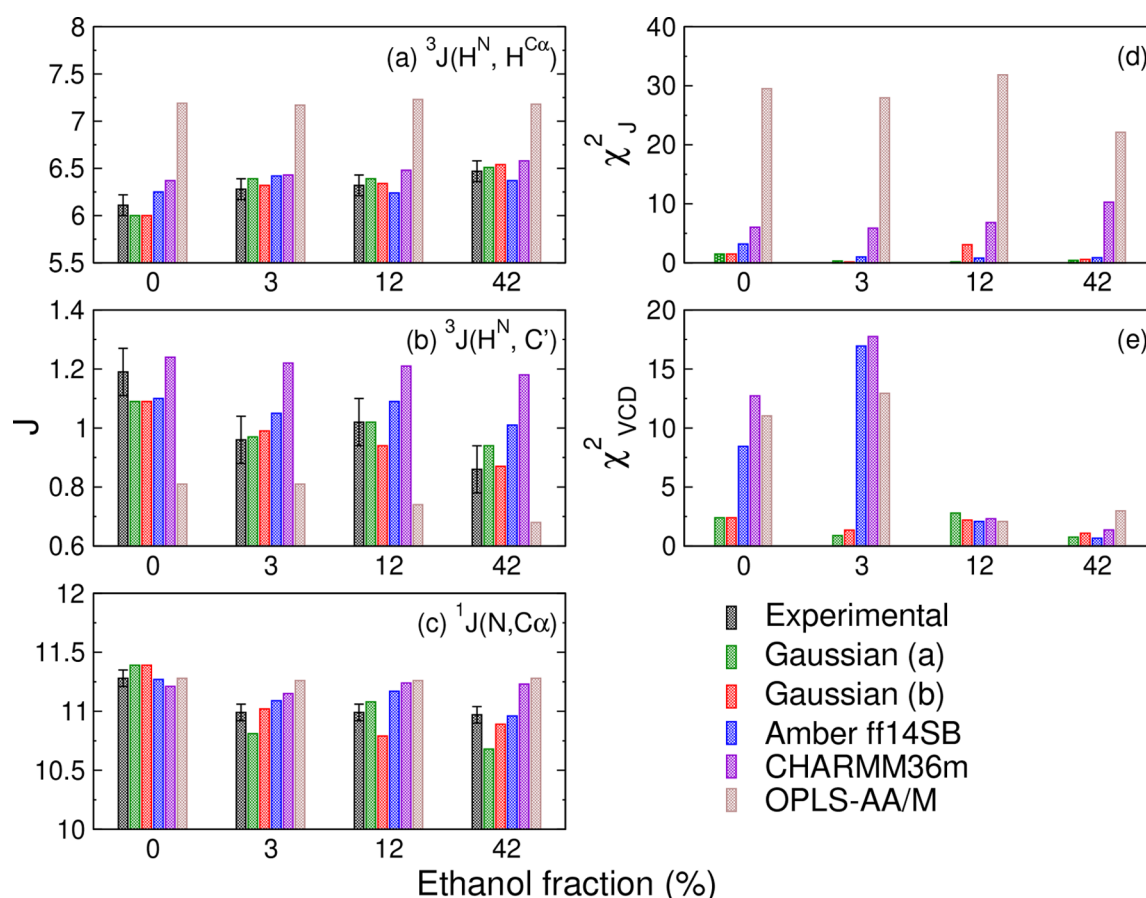


Figure 6. Comparison of Gaussian modeling and MD-derived spectroscopic data for alanine residue in cationic GAG to experimental results. Three J-coupling constants (a–c) and VCD amide I' profiles (Figures S1 and S2) are calculated from Ramachandran distributions obtained within Gaussian models (a) and (b) and derived from MD simulations with three force fields and compared to corresponding experimental data in pure water and aqueous solutions with 3, 12, and 42 mol % ethanol. (a–c) The error bars correspond to statistical uncertainties. The χ^2_J (d) and χ^2_{VCD} (e) values as defined in Methods provide a measure of overall deviations of Gaussian models and MD force fields from experimental data.

alanine residue in GAG are shifted along the negative ϕ direction (Figure 4) when compared to the corresponding pPII basins produced by the other two force fields (Figure 3 and Figure 5) and both Gaussian models (Figure 2) in all aqueous solutions under study. This result is consistent with the results of the previous study on GAG in pure water.³³ The inverse γ and classical γ turn mesostates, whose populations increase with the ethanol fraction in the both Gaussian models, are practically absent from the all MD-derived Ramachandran distributions.

Quantitative Assessment of MD Force Fields and Gaussian Models. We here evaluate Amber ff14SB, OPLS-AA/M, and CHARMM36m with respect to their capacity to reproduce the experimental J-coupling constant values and amide I' profiles for alanine residue in GAG when immersed in aqueous solutions with 3%, 12%, and 42% ethanol. We also assess Gaussian models (a) and (b), which offer benchmarks for comparison to MD results. The $^3J(H^N, H^{C\alpha})$, $^3J(H^N, C')$, and $^1J(N, C')$ values determined by the experiment, obtained from Gaussian models (a) and (b), and the respective MD-derived values are displayed in Tables S2, S3, and S4. Figure 6a–c shows that Amber ff14SB reproduces the three J-coupling constants quite well, whereas OPLS-AA/M and CHARMM36m compare to experimental J-coupling constants less favorably. The discrepancies stem mostly from an overestimation of $^3J(H^N, C')$ (Table S3) and $^1J(N, C\alpha)$ (Table

S4) by CHARMM36m and an even larger overestimation of $^3J(H^N, H^{C\alpha})$ (Table S2) by OPLS-AA/M. The overall performance of Gaussian models (a) and (b) and the three MD force fields is captured in the χ^2_J function, as defined in Methods, in Figure 6d. A comparison of χ^2_J values also demonstrates that Gaussian model (a) slightly outperforms Gaussian model (b). The MD-derived VCD amide I' profiles in Figure S2 (top row) demonstrate that the performance of all three force fields is the worst for 3% ethanol but improves for the ethanol fraction of 12% and even more so for 42% ethanol. This assessment is captured in the χ^2_{VCD} function values displayed in Figure 6e. This increasing agreement between experimental and MD-derived results may, however, be coincidental because the three force fields examined here tend to underestimate the rotational strength of amide I'. Because the latter decreases with increasing ethanol fraction while the calculated distributions remain almost intact, the discrepancy between experimental and calculated VCD signal is reduced. We then examined ethanol-induced changes to the MD-derived Ramachandran distributions of alanine residue in cationic GAG that might occur at higher ethanol fractions, for which spectroscopic data are not available. Table S2 shows experimental and MD-derived values of $^3J(H^N, H^{C\alpha})$ alongside their values obtained within Gaussian models (a) and (b). All MD simulations cause an increase of the $^3J(H^N, H^{C\alpha})$ constant upon an increase of ethanol fraction from 0% to

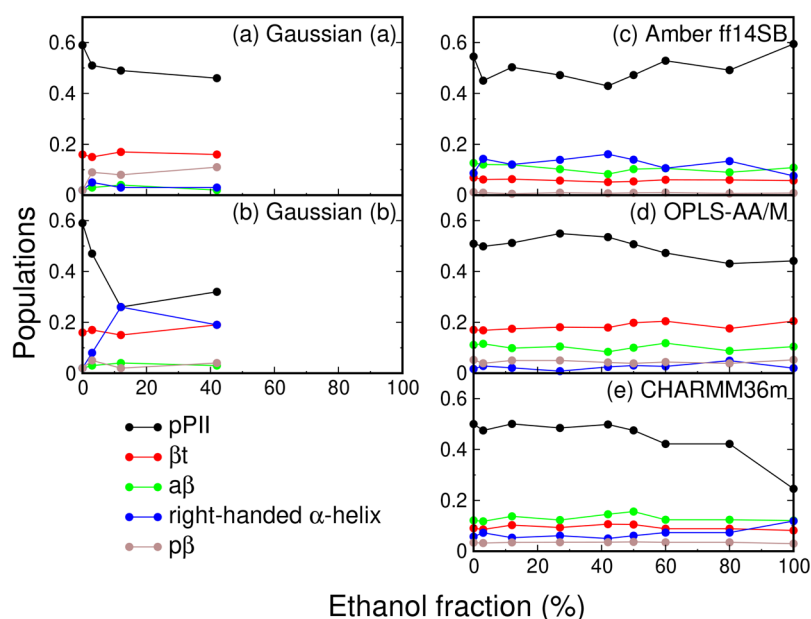


Figure 7. Mesostate populations for alanine residue in cationic GAG in aqueous solutions with different ethanol fractions. Populations of five mesostates are displayed in dependence on the ethanol fraction in aqueous solutions for Gaussian models (a) and (b) as well as for MD-derived conformational ensembles within (c) Amber ff14SB, (d) OPLS-AA/M, and (e) CHARMM36m.

50%, although this change is not as pronounced as the experimentally observed one.³⁷ Above 50% ethanol, the behavior of the MD-derived $^3J(H^N, H^{C\alpha})$ constant is strongly force field-dependent. Amber ff14SB produces a decrease, CHARMM36m predicts a significant increase, and OPLS-AA/M produces a moderate increase but overestimates the experimental $^3J(H^N, H^{C\alpha})$ value by a substantial amount for ethanol fractions between 3 and 42%.

MD Force Fields Predict Ethanol-Induced Changes in pPII and Right-Handed Helical Populations. To further examine the effect of ethanol on the Ramachandran distribution of alanine residue in cationic GAG, populations of mesostates defined in Methods are examined as a function of the ethanol fraction in the aqueous solution. Figure 7 shows mesostate populations calculated for Gaussian models (a) and (b) and MD-derived mesostate populations for all solutions under study.

Consistent with the observed ethanol-induced changes in the Ramachandran distribution in Figure 2, both Gaussian models (a) and (b) predict that the pPII population decreases with ethanol fraction (Figure 7a,b, black symbols); however, the pPII population decrease is larger within model (b). Whereas in Gaussian model (a), the decrease in the pPII population is concomitant with an increase in β -strand like conformations, specifically $p\beta$ (Figure 7a, black and brown symbols), in Gaussian model (b), the decrease in the pPII population coincides with an increase in the right-handed helical population (Figure 7b, black and blue symbols). MD simulations with Amber ff14SB result in the pPII population, which is reduced when pure water is replaced by an aqueous solution with 42% ethanol. The change in the pPII population is in line with Gaussian model (a), but unlike in either Gaussian models, this change is nonmonotonous (Figure 7c, black symbols). However, in Amber ff14SB the ethanol-induced decrease in the pPII population is concomitant with an increase in the right-handed helical population, which is qualitatively more in line with Gaussian model (b) than

Gaussian model (a). In contrast to Amber ff14SB, in MD simulations with OPLS-AA/M and CHARMM36m the pPII population does not strongly depend on the ethanol fraction of up to 42% (Figure 7d,e, black symbols), which is inconsistent with experimental data as well as with the results of Gaussian models.

Ethanol-induced changes in the pPII populations for ethanol fractions above 42%, for which experimental data is not available, strongly depend on the force field. In Amber ff14SB, the pPII population increases, albeit in a nonmonotonous way, when the ethanol fraction changes from 42% to 100%, whereby in pure ethanol the pPII population even exceeds the corresponding value in pure water (Figure 7c, black symbols). In contrast, in OPLS-AA/M the pPII population decreases when the ethanol fraction increases from 42% to 100% but this change is relatively modest (Figure 7d, black symbols). The pPII population in CHARMM36m simulations exhibits only a weak dependence on the ethanol fraction for fractions below 80%. When the ethanol fraction increases from 80% to 100%, CHARMM36m predicts a decrease of the pPII population from 0.50 to 0.25 (Figure 7e, black symbols), which is reflected by the decrease in $^3J(H^N, H^{C\alpha})$ constant (Table S3). This pPII population decrease appears to coincide with a small increase of the right-handed helical population (Figure 7e, black and blue symbols), which is again more in line with Gaussian model (b) than Gaussian model (a). It should be noted, however, that the overall sampling of the right helical region in CHARMM36m is not fully represented by the mesostate population since the asymmetric basin in the Ramachandran plot extends beyond the mesostate boundaries in the direction of its upper left corner (Figure 5).

To systematically examine ethanol-induced transitions among mesostates of alanine residue in GAG, the Pearson correlation coefficients between pairs of mesostate populations are calculated. The results are reported in Table S6, where the correlation coefficients that correspond to statistically significant anticorrelations are given in red font. These

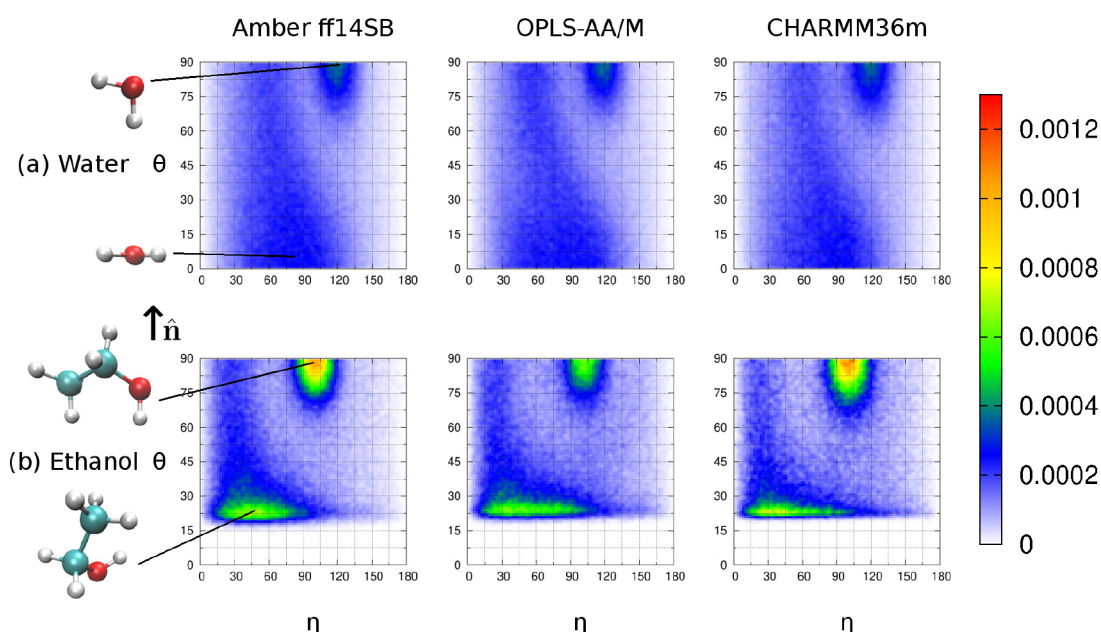


Figure 8. Water and ethanol orientation plots. These 2D plots show the average orientation of water and ethanol molecules within the solvation layer adjacent to the backbone of alanine residue in GAG in pure water and pure ethanol, respectively, as derived from MD simulations with three force fields. Two angles η and θ that define the orientation of each water (ethanol) molecule relative to the SAS normal $\hat{n}$ are defined in [Methods](#). Characteristic orientations of water and ethanol molecules (relative to $\hat{n}$) and the respective regions of the water and ethanol plots are displayed on the left. Red, white, and cyan beads represent oxygen, hydrogen, and carbon atoms, respectively.

anticorrelations indicate that in Gaussian model (b), Amber ff14SB and CHARMM36, conformational ensemble of alanine residue undergoes a transition between pPII and right-handed helical conformations, whereas Gaussian model (a) predicts ethanol-induced transition between the pPII and β t or β β mesostates. In OPLS-AA/M, the correlation coefficient between pPII and right-handed α -helix is equal to -0.6561 , which is borderline significant and also more aligned with Gaussian model (b) than Gaussian model (a).

Taking into consideration that all three MD force fields predict ethanol-induced changes that are, at least qualitatively, in line with Gaussian model (b), we asked if the redistribution between pPII and right-handed helical might indicate a real trend. A strong argument in favor of Gaussian model (a) can be deduced from the UV circular dichroism (CD) spectra of GAG in aqueous solutions with 48% and 55% ethanol, which were recorded for GAG concentrations below the respective critical gelation concentrations.³⁷ The CD spectra taken at room temperature indicate a predominant pPII state of the conformational ensemble of alanine residue in GAG. The dichroism values measured at the respective negative and positive maxima are only slightly less pronounced than they are in the spectrum of GAG in pure water. The rather modest ethanol dependence of the CD spectrum could reflect compensating influences of propensity changes (the pPII population decrease) concomitant with the change in the location of the pPII basin (whereby the center along the ψ -coordinate moves from 150° to 135°). A decrease of the pPII propensity from 0.8 (in water) to 0.35 (in the aqueous solution with 42% ethanol) would cause a large change in the CD spectrum, i.e., substantially reduced positive and negative maxima, which is at variance with the other spectroscopic data. Moreover, UV-CD spectra taken at different temperatures exhibit an isodichroic point which suggests a two-state transition. Thus, the UV-CD spectroscopy data reveal the

ethanol-induced transition of alanine residue from pPII to β mesostate, which is consistent with Gaussian model (a) rather than Gaussian model (b). Taking into consideration these additional spectroscopic data, it appears that all three MD force fields fail to capture the effect of ethanol on the conformational ensemble of alanine residue in GAG.

Characterization of the Solvation Layer and Solvent-Peptide Hydrogen Bonding. We here examine the properties of water and ethanol within the solvation layer around the GAG peptide. First, we derive water and ethanol orientation plots as described in [Methods](#). Water orientation plots were introduced by Meral and collaborators to visualize the average orientation of water in the hydration layer around the guest residue x in GxG peptides.³⁰ Recently, water orientation plots were used to examine orientational preferences of hydration water adjacent to the backbone of guest glycine, alanine, and proline in residues in GxG peptides.³⁶ Here, we extend this analysis to ethanol, specifically the part of ethanol that participates in HB formation (see [Methods](#) for details). [Figure 8a,b](#) shows water and ethanol orientation plots displaying the average orientation of water and ethanol molecules adjacent to the backbone of alanine residue of GAG in pure water and pure ethanol, respectively. In each type of the plot, the two most preferred orientations of water and ethanol are displayed as defined relative to the normal $\hat{n}$ to the SAS. Because water is smaller than ethanol, many more water than ethanol molecules are within the solvation layer in pure water than in pure ethanol, which explains why water orientation plots show a more homogeneous distribution of water orientations in comparison to ethanol orientation plots. Nonetheless, both water and ethanol orientation plots demonstrate the preference for water or ethanol orientations that are associated with hydrogen bonding to the functional groups of the alanine residue. The most populated region of both types of plots in the upper right

region corresponds to water or ethanol orientations that occur when the solvent forms HBs with the carbonyl group of alanine residue (solvent as a donor), whereas the increased population of the lower left region results from water or ethanol forming HBs with the amino group of alanine residue (solvent as an acceptor). In contrast to the water orientation plots, ethanol orientation plot exhibits a forbidden region for $0 < \theta < 15^\circ$, which occurs due to volume exclusion between the ethanol molecule and the C_α group of the alanine residue. The ethanol orientation plots demonstrate that two orientations of ethanol dominate the solvation layer adjacent to the backbone of alanine residues. Notably, the most populated ethanol orientations in the upper right region of these plots are shifted in the negative η -direction ($\sim 100^\circ$) relative to the corresponding water orientations ($\sim 120^\circ$) due to distinct values of the H–O–H angle of water and C–O–H angle of ethanol. Of three force fields, OPLS-AA/M produces the least populated preferred orientations in the upper right region of the plots for both ethanol and water.

We next asked if the MD force fields capture the peculiar solute–solvent interactions that the spectroscopic data revealed for the different functional groups of GAG, which lead to the hypothesis that the N-terminal and C-terminal regions of GAG are solvated preferentially by water and ethanol, respectively.⁴⁰ To this end, we examine hydrogen bonding of water and ethanol to GAG in the three MD force fields by calculating the average number of HBs between the solvent and each functional group of the three residues in GAG as a function of the ethanol fraction in the solution. Figure S4 demonstrates that the average number of HBs is nearly a linear function of the ethanol fraction for all functional groups in GAG except for CHARMM36m. In CHARMM36m simulations, the average number of HBs that water or ethanol forms with the NH_3^+ group of residue 1 clearly deviates from a linear function with a rather sharp increase in the average number of HBs between ethanol and the NH_3^+ group of residue 1 upon ethanol fraction increase from 80% to 100%. As expected, both water and ethanol form significantly more HBs with the NH_3^+ group of residue 1 than with the amino groups of residues 2 and 3. However, there are large differences in the average number of water–peptide and ethanol–peptide HBs among the three force fields. The lowest average number of HBs that the solvent (water and ethanol alike) forms with NH_3^+ and amino groups are observed within Amber ff14SB. OPLS-AA/M and CHARMM36m produce on average comparable numbers of HBs formed by NH_3^+ and amino groups. However, in OPLS-AA/M, water and ethanol show similar propensities for HB formation with these groups, while in CHARMM36m, water forms on average more HBs with these groups than ethanol. The force field-specific differences in HB formation are less pronounced for HBs between the solvent and the carbonyl groups. Notably, of the three force fields, Amber ff14SB simulations are characterized by on average the most HBs that the carbonyl groups of all three residues in GAG form with both water and ethanol.

We further examine the solvation layer around the GAG peptide by calculating the average molar fraction of ethanol within the solvation layer as a function of the bulk ethanol fraction (Figure 9a). Figure 9b shows a fraction of HBs that ethanol forms with the NH_3^+ and amide groups, and similarly, Figure 9c shows a fraction of HBs that ethanol forms with the carbonyl groups. Both HB fractions in Figure 9b,c are normalized to the total number of HBs that water and ethanol

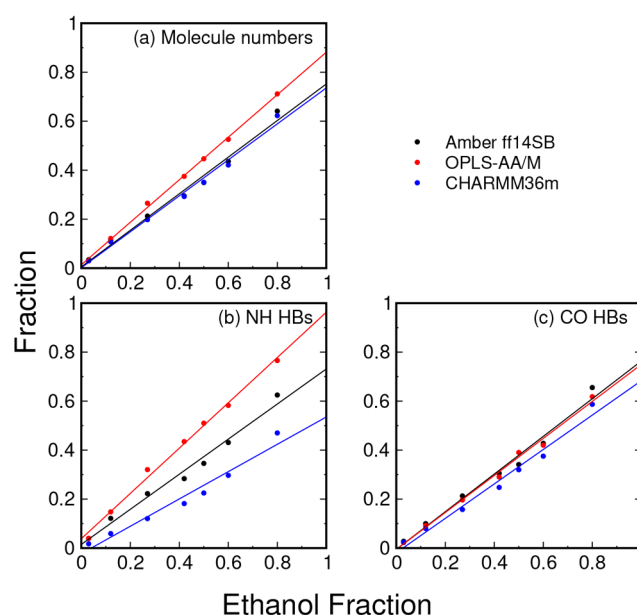


Figure 9. Characterization of the solvation layer around GAG peptide. (a) The fraction of ethanol molecules relative to all solvent molecules in the solvation layer, (b) the fraction of HBs between ethanol and amide groups of GAG relative to the total number of HBs formed by amide groups, and (c) the fraction of HBs between ethanol and carbonyl groups of GAG relative to the total number of HBs formed by carbonyl groups are plotted as a function of the bulk ethanol fraction in the simulations with three force fields. The straight lines correspond to linear fits with the resulting slopes displayed in Table 1.

form with the respective peptide group and are displayed in dependence on the bulk ethanol fraction. The results of the linear regression analysis on these three quantities versus the bulk ethanol fraction are displayed in Table 1. Notably, for all three force fields, all three slopes are below 1, indicating that water successfully displaces ethanol within the solvation layer and consequently forms on average more HBs with GAG than ethanol. The deviation of the slope of the linear fit in Figure 9a from 1 is a measure of how well water displaces ethanol in the solvation layer. This effect is the strongest for CHARMM36m. To further examine this effect in a way that is independent of the solvation layer definition, we calculated radial density functions of water and ethanol oxygen atoms from the three C_α atoms of GAG (Figure S3). Consistent with the results in Figure 9a, a comparison of water and ethanol radial density functions in Figure S3 shows that of the three force fields, CHARMM36m is associated with the strongest effect of water displacing ethanol from the region adjacent to the peptide. Of the three force fields, ethanol in OPLS-AA/M is the most capable of accessing the functional peptide groups and forming HBs with NH_3^+ and amide groups (Figure 9a,b). In MD simulations with Amber ff14SB and CHARMM36m, ethanol exhibits comparable albeit reduced (relative to OPLS-AA/M) abilities to access the solvation layer around GAG (Figure 9a), but ethanol in CHARMM36m forms on average fewer hydrogen bonds with the functional groups of GAG than ethanol in Amber ff14SB (Figure 9b,c).

DiGiuseppi et al. reported that while the two amide I' bands are clearly separable for GAG in pure water and in aqueous solutions with ethanol fractions below 30%, they merge into a single band profile at ethanol fraction of 42% (see Figure 5 in

Table 1. Slopes and Pearson Correlation Coefficients Resulting from the Linear Regression Analysis of Solvation Layer Characteristics Shown in Figure 9

	solvation EtOH		EtOH-NH HBs		EtOH-CO HBs	
	slope	R ²	slope	R ²	slope	R ²
Amber ff14SB	0.7469	0.9927	0.7163	0.9914	0.7688	0.9908
OPLS-AA/M	0.8679	0.9992	0.9245	0.9973	0.7555	0.9960
CHARMM36m	0.7342	0.9946	0.5576	0.9820	0.7024	0.9899

the paper by DiGiuseppi et al.⁴⁰). These spectral changes are referred to as a blue shift of the low wavenumber amide I' band, which is predominantly assignable to the CO-stretching mode of the C-terminal peptide group and indicates a preferential binding of ethanol to the CO and most likely also to the NH-group of the respective peptide group. A similar spectral shift was observed for the amide I band of *N*-methylacetamide,⁴⁰ implying that it reflects solely solute–solvent interactions (weakening of hydrogen bonding) and is unrelated to structural changes. The chemical shift of the amide proton of the N-terminal peptide group shows the expected temperature dependence, namely a linear downfield shift with increasing temperature even for 42% ethanol. In contrast, the chemical shift of the amide proton of the C-terminal peptide exhibits an abnormal temperature dependence above ethanol fractions of 10%. These experimental data strongly suggest that ethanol associates preferentially with the C-terminal peptide group, whereas water prefers to bind to the N-terminal half of the peptide. The above solvation analysis performed on MD trajectories derived from three force fields and, specifically, the analysis of water-GAG and ethanol-GAG hydrogen bonding does not capture the preferential solvation of the N-terminal and C-terminal part of GAG peptide by water and ethanol, respectively, indicated by spectroscopic data.

Emergence of Ethanol Clustering in MD Simulations.

Water/ethanol mixtures are known to be nonideal and characterized by a range of anomalies. Previous MD studies revealed a weak phase transition due to formation of ethanol clusters occurring at low ethanol fractions of 6–10%.⁶⁹ We here asked if ethanol-induced decrease in the pII population that is predicted by the experiment-based Gaussian models can be in any way associated with ethanol clustering. We here calculate the ethanol cluster size distribution for different water/ethanol mixtures to find out whether such clustering is present and to which extent this behavior depends on the force field. Our results in Figure S5 show that for all three force fields, the ethanol cluster size distribution exhibit a single maximum at a cluster size ~ 1 for aqueous solutions with ethanol fractions 3% and 12%. Ethanol clusters of much larger size only appear in solutions with 27% ethanol and grow rather significantly in number and size with an increasing ethanol fraction. The most probable cluster sizes are listed in Table S7. At 27% ethanol, two cluster sizes are listed because a large ethanol cluster coexists with “monomeric” ethanol molecules. As the ethanol fraction increases, the most probable cluster size also increases and approaches the total number of ethanol molecules in the simulation. Of the three force fields, OPLS-AA/M results in the largest ethanol clusters. CHARMM36m produces significantly smaller ethanol clusters than the other two force fields, which is reflected also in the most probable cluster sizes (Table S7). The onset of formation of larger ethanol clusters falls into region III of ethanol fractions, as introduced in the experimental study by Milorey and

collaborators,³⁷ which coincides with the most pronounced changes in the $^3J(H^N, H^{C\alpha})$ coupling constant that directly affects the pII population. Thus, formation of large ethanol clusters appears to coincide with the observed changes in the experimental value of $^3J(H^N, H^{C\alpha})$, which are responsible for pII population changes. However, while MD simulations capture ethanol clustering, the changes in the pII population are significantly smaller than those derived from experimental data.

Can MD Simulations of GAG Ensembles in Water/Ethanol Mixtures Capture Spectroscopic Data?

MD simulations described in this study thus far assume that GAG is monomeric in solutions under study, which may be the reason for the observed discrepancy between MD-derived and experimental data. In the experiment, cationic GAG forms a hydrogel with visible GAG fibers at a concentration of 200 mM only when the ethanol fraction exceeds $\sim 50\%$.^{37,38} However, 200 mM GAG at ethanol fractions up to 42% may still be prone to formation of transient oligomers without the capacity to self-assemble into fibers and form a gel, which would be difficult to detect and characterize experimentally. In contrast, early stages of oligomer formation by a short peptide such as GAG are easily accessible to MD simulations. We thus explored an ensemble of 16 GAG peptides in a $5 \times 5 \times 5$ Å box (corresponding to 200 mM concentration) in pure water and in the aqueous solution with 42% ethanol within the three force fields under study. The choice of the 42% ethanol fraction is informed by the observed emergence of large ethanol clusters in simulations of monomeric GAG (Figure S5). To assess the convergence of MD simulations of the ensemble of 16 GAG peptides and compare the Ramachandran distribution of alanine residue in a single GAG to that of the 200 mM GAG ensemble in a pure water and the aqueous solution with 42% ethanol, we monitor the average mesostate populations calculated within 100 ns-long time windows along MD trajectories. Figure S6 demonstrates that the mesostate populations do not change over time within the total simulation time of 300 ns, demonstrating that simulations have converged and the ensemble of 16 GAG peptides is in a steady state. The mesostate populations derived from the Ramachandran distributions using time frames between 200 and 300 ns (the last point in each panel of Figure S6) are displayed in Figure 10.

We first asked whether or not replacing a single GAG by a 200 mM GAG affects the mesostate populations in either solvent. In a pure water, the mesostate populations of monomeric GAG and 200 mM GAG are fairly comparable with an exception of a slight decrease in the pII population and even smaller decrease in the β t and $\alpha\beta$ populations (Figure 10, black and green bars). The pII decrease is the largest in Amber ff14SB simulations, which also produce a slight increase in the helical population. Larger changes in mesostate populations are noted when comparing monomeric and 200 mM GAG in the aqueous solution with 42% ethanol (Figure

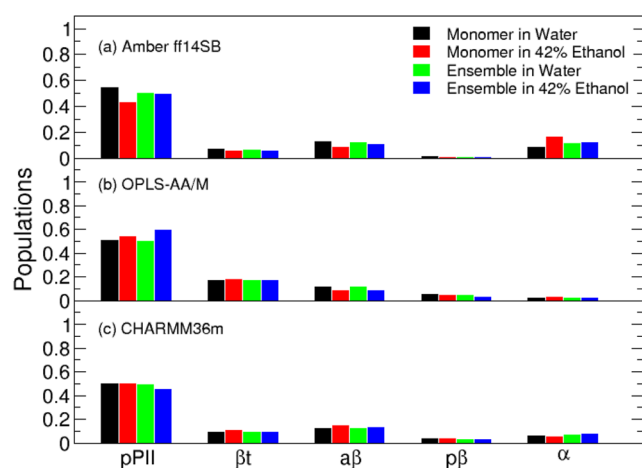


Figure 10. Mesostate populations of alanine residue in monomeric GAG and GAG ensembles in a pure water and in the aqueous solution with 42% ethanol derived by MD simulations with (a) Amber ff14SB, (b) OPLS-AA/M, and (c) CHARMM36m.

10, red and purple bars). Amber ff14SB simulations of a 200 mM GAG reveal increased pII, β_t , and $\alpha\beta$ alongside a decreased helical population. OPLS-AA/M simulations produce a more substantial increase in the pII population concomitant with slight decreased populations of all other mesostates when monomeric GAG is replaced by a 200 mM GAG ensemble. Simulations with CHARMM36m, in contrast, produce decreased pII and β populations at the expense of an increased right-handed helical population for alanine residue in 200 mM GAG relative to monomeric GAG. These findings indicate that the addition of ethanol to water affects the interactions among GAG peptides; however, the effect of these interactions on the conformational ensemble of alanine residue strongly depends on the force field.

We then asked if the comparison of mesostate populations of 200 mM GAG in pure water and aqueous solution with 42% ethanol could explain the discrepancies between MD data for monomeric GAG and spectroscopic data (Figure 10, green and purple bars). In Amber ff14SB, the pII population appears to be mostly unaffected by the addition of ethanol, although simulations with 42% ethanol result in a decrease in β populations and a slight increase in the right-handed helical population. In OPLS-AA/M, the pII population significantly increases while β and helical populations decrease in the presence of 42% ethanol. In CHARMM36m, the pII population decreases while β and helical populations slightly increase. While Amber ff14SB simulations of monomeric GAG capture, at least in a qualitative way, some aspects of experimental findings, i.e., a decrease in the pII population upon addition of ethanol, simulations of 200 mM GAG appear to be more or less insensitive to the presence of ethanol. This may be due the particular C-terminal capping in Amber ff14SB (CONH₂) which differs from the experimental capping (COOH). While this difference does not significantly affect the conformational ensemble of alanine residue in monomeric GAG, it may affect interpeptide interactions, which in turn affect the average conformational preferences of alanine residue. OPLS-AA/M again predicts ethanol-induced increase in the pII population, which is inconsistent with experimental data. Of the three force fields, only CHARMM36m produces results that are consistent with experimental findings. Moreover, in CHARMM36m, the ethanol-induced decrease in the

pII population stems from the interactions among GAG peptides. The extent of self-association in GAG ensembles, which could affect the conformational ensemble of alanine residue in GAG, is quantified in Figure 11, which shows the

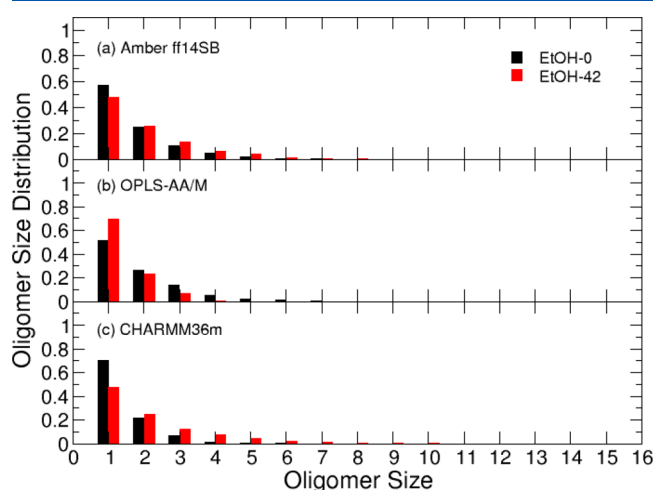


Figure 11. Oligomer size distributions of 200 mM GAG in water (black bars) and in aqueous solution with 42% ethanol (red bars) derived from (a) Amber ff14SB, (b) OPLS-AA/M, and (c) CHARMM36m simulations using time frames between 250 and 300 ns of respective MD trajectories.

size distribution of GAG oligomers formed in pure water and in water with 42% ethanol. GAG oligomer size distributions in both solvents and all three MD force fields are dominated by monomeric GAG, followed by populations of dimers and larger oligomers, which decrease with the size of the oligomer, consistent with experimental observation that GAG is mostly soluble at these conditions.

In MD simulations, the effect of ethanol on the GAG oligomer size distribution strongly depends on the force field. In Amber ff14SB simulations, the addition of 42% ethanol to water causes a decrease in the monomer population and a slight increase in the populations of dimers and larger oligomers, consistent with the expectation that ethanol in the solvent promotes GAG self-assembly (Figure 11a). However, the difference between the oligomer size distributions in water and water/ethanol mixture is rather small. The C-terminal capping (CONH₂) in Amber ff14SB might decrease the interpeptide hydrogen bonding propensity between the NH₃ and CONH₂ groups in comparison to the hydrogen bonding propensity between the NH₃ and COOH groups, so that 200 mM GAG simulations in Amber ff14SB are inconclusive. In OPLS-AA/M, ethanol has the opposite effect and causes increased GAG solubility reflected in an increased population of monomers at the expense of the populations of dimers and larger oligomers (Figure 11b). In CHARMM36m, GAG in the aqueous solution with 42% ethanol forms significantly higher populations of dimers and larger oligomers in comparison to GAG in pure water (Figure 11c). CHARMM36m simulations are associated with the largest GAG oligomer, a decapeptide. A comparison of mesostate populations in Figure 10 and GAG oligomer size distributions in Figure 11 suggests an inverse relationship between the pII populations and the propensity to form oligomers in 200 mM GAG simulations. These findings combined demonstrate that of the three force fields, only CHARMM36m captures the experimental observation

that the addition of 42% ethanol decreases the average pPII population of alanine residue in 200 mM GAG simulations. Moreover, CHARMM36m provides a plausible explanation that formation of soluble short-lived GAG oligomers may underlie the experimentally observed ethanol-induced decrease in pPII. A short movie of the final 25 ns of the MD trajectory corresponding to 200 mM GAG in the aqueous solution with 42% ethanol in the [Supporting Information](#) demonstrates that GAG oligomers are not very stable but rather short-lived. An example of a GAG tetramer in [Figure 12a](#) shows that the side

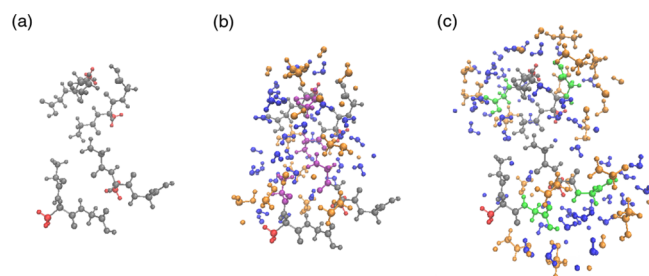


Figure 12. A snapshot of a GAG tetramer (a) without solvent molecules and with water (blue) and ethanol (orange) molecules that are within 5 Å of (b) the N-terminal glycine (purple) and (c) the C-terminal glycine (green). Alanine side chains are shown in red. The snapshots are derived from the MD trajectory of 200 mM GAG in the aqueous solution with 42% ethanol within CHARMM36m using VMD.⁶⁶

chains of alanine residue (red) face away from the center of the tetramer and do not appear to contribute to oligomer formation. Instead, [Figure 12b](#) demonstrates that the N-terminal glycines (purple) face toward each other at the center of the tetramer and are surrounded almost exclusively by water molecules (blue). In contrast, [Figure 12c](#) shows that ethanol molecules (orange) prefer the periphery of the GAG tetramer adjacent to the C-terminal glycines (green). Overall, this GAG tetramer conformation and the relative placement of water and ethanol molecules adjacent to the N-terminal and C-terminal end of GAG, respectively, are consistent with experimental observations.⁴⁰

CONCLUSIONS

The overarching motivation for this study originates in the surprising observation that cationic GAG in water/ethanol mixtures forms a rather strong gel.^{37,38} Gelation of this ternary mixture occurs only above a critical peptide concentration, which depends on the mole fraction of ethanol in the aqueous solution.^{37,38} The gel phase consists of unusually long peptide fibrils that form a sample spanning network. Spectroscopic data suggest that these fibrils do not exhibit the canonical β -sheet structure, which characterizes amyloid fibrils formed by amyloidogenic proteins, raising the question of what the driving force of the gelation process might be. Gel formation requires in addition to a high GAG concentration (~ 200 mM or above) also substantial ethanol content (above 50% ethanol mole fraction). Although GAG does not form fibrils at a lower ethanol fraction, ethanol nonetheless affects the conformational dynamics of alanine residue in GAG by reducing the pPII content at the expense of the β content. In this study, we assess three commonly used MD force fields with respect to their ability to capture experimentally observed ethanol-induced changes in conformational preferences of alanine

residue in GAG.^{37,40} These three force fields, Amber ff14SB, OPLS-AA/M, and CHARMM36m, have been recently evaluated with respect to their capacity to capture conformational dynamics of the central glycine residue in GGG³⁶ and the central alanine residue in GAG and AAA in pure water.³³

We here use three J-coupling constants ($^3J(H^N, H^{\alpha})$, $^3J(H^N, C')$, and $^1J(N, C')$) and VCD amide I' profiles for alanine residue of GAG in three different water–ethanol mixtures, which were reported to possess distinct thermodynamic and ethanol–water demixing properties, and reapply an experiment-based Gaussian Ramachandran distributions of alanine in cationic GAG.⁴⁰ Our analysis shows that in addition to the ethanol-induced pPII $\rightarrow \beta$ redistribution previously reported by DiGiuseppi and collaborators, spectroscopic data could also be accounted for by a more pronounced pPII $\rightarrow$ right-handed helix shift of conformational ensembles of alanine residue in GAG. We refer to the above solutions of Gaussian modeling as Gaussian models (a) and (b), respectively. Notably, the agreement with experimental data in Gaussian model (b) is slightly less satisfactory than in Gaussian model (a). Furthermore, earlier reported UV-CD spectra of GAG in water/ethanol mixtures measured right below the critical concentration of gelation offer evidence that strongly supports Gaussian model (a) and is inconsistent with Gaussian model (b). It is thus expected that MD simulations would account for ethanol-induced decrease in the pPII population concomitant with an increase in β -strand populations, in particular p β .

Assuming monomeric GAG, we conducted simulations of a single GAG in pure water, in aqueous solutions with 3%, 12%, and 42% ethanol to mimic experimental conditions, in aqueous solutions with 50%, 60%, and 80% ethanol and in pure ethanol within each of the three MD force fields. The assessment of the three MD force fields was performed by a direct comparison of MD-derived and experimental values of three J-coupling constants and amide I' profiles, whereby the results of Gaussian models (a) and (b) were also included in comparison. This comparison demonstrated that overall both Gaussian models outperform the MD force fields. Of the three MD force fields, only Amber ff14SB predicts a decrease in the pPII population of alanine residue in GAG when pure water is replaced by an aqueous solution with 42% ethanol; however, while the magnitude of the pPII reduction is consistent with Gaussian model (a), the pPII decrease is concomitant with an increase in the right-handed helical population rather than an increase in p β , which is more consistent with Gaussian model (b). In OPLS-AA/M simulations, the pPII population increases as a result of the addition of 42% ethanol, which is inconsistent with experimental data. Simulations with CHARMM36m do not show any significant changes for ethanol fractions up to $\sim 50\%$ ethanol. For solutions with higher ethanol fractions, there is no experimental data to facilitate a comparison. It is noteworthy, however, that the three MD force fields predict very different behavior at these high ethanol concentrations. Both OPLS-AA/M and CHARMM36m predict a decrease in pPII content with increasing ethanol fraction, but this decrease is significantly higher for CHARMM36m. In contrast, Amber ff14SB predicts an increase in the pPII population in pure ethanol relative to pure water. In both Amber ff14SB and CHARMM36m, an increase/decrease in the pPII population is concomitant with a decrease/increase in the right-handed helical population, which is at odds with Gaussian model (a), which favors pPII $\rightarrow \beta$ redistribution while keeping the right-handed helical

population more or less intact. In these two force fields, the pPII $\rightarrow$ helical transition is more likely to occur than the pPII $\rightarrow$ β transition, which may be associated with the free energy barriers of the respective transitions. These findings combined thus suggest that the free energy barrier between pPII and right-handed helical conformational basins may be too low while the free energy barrier between pPII and β may be too high to account for the spectroscopic data on alanine residue in GAG.

Finally, we asked whether or not the above limitations of MD force fields in capturing the spectroscopic data for alanine residue stem from our assumption that GAG is monomeric in pure water and in aqueous solutions with ethanol fractions up to 42%. Although at these low ethanol fractions 200 mM GAG does not form a hydrogel *in vitro*,^{37,38,40} it is plausible to expect that the presence of ethanol in the solution alters interpeptide interactions by increasing the propensity for GAG oligomer formation and that ethanol-induced quaternary contacts among GAG peptides would affect conformational preferences of alanine residue in GAG. We thus performed MD simulations of 200 mM GAG ensembles in pure water and in an aqueous solution with 42% ethanol. The discrepancy between MD-derived and experimental data in this case increased for both, Amber ff14SB and OPLS-AA/M; however, the results of CHARMM36m simulations are closer to the predictions of Gaussian model (a) with the remaining issue of the right-handed helical conformations, which are still too easily accessible in comparison to p β conformations. Further insights into the GAG propensity to form oligomers provide evidence that at least within CHARMM36m, the ethanol-induced decrease in the pPII population may be caused by increased propensity of GAG for oligomer formation. Importantly, a closer inspection of a transient GAG tetramer obtained in CHARMM36m simulations in water with 42% ethanol reveals that the four N-termini of GAG at the center of the tetramer are preferentially solvated by water, whereas the C-termini of the four GAG peptides at the periphery of the tetramer are preferentially solvated with ethanol, which is in line with the experiment-based solvation model proposed by DiGiuseppi and collaborators.⁴⁰

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.0c08245>.

Table S1, number of water and ethanol molecules in MD simulations; Table S2, experimental and calculated $^3J(H^N, H^C\alpha)$ constants for alanine residue in GAG at different ethanol fractions; Table S3, experimental and calculated $^3J(H^N, C')$ constants for alanine residue in GAG at different ethanol fractions; Table S4, experimental and calculated $^1J(N, C\alpha)$ constants for alanine residue in GAG at different ethanol fractions; Table S5, parameters for Gaussian models; Table S6, Pearson correlation coefficients between pairs of mesostate populations; Table S7, most probable ethanol cluster sizes; Figure S1, comparison of experimental VCD amide I' profiles, Raman, and IR spectra with those produced by Gaussian models, Figure S2, comparison of experimental VCD amide I' profiles, Raman, and IR spectra with those produced by MD simulations; Figure S3, radial distribution functions of water and ethanol

around the $C\alpha$ atoms of each residue; Figure S4, GAG-water and GAG-ethanol hydrogen bonding for each residue in GAG; Figure S5, ethanol cluster size distributions; and Figure S6, average mesostate populations of guest alanine in GAG produced from MD simulations of 200 mM GAG in water and 42% ethanol (PDF)

Formation of unstable GAG oligomers in a 42% ethanol fraction (MP4)

■ AUTHOR INFORMATION

Corresponding Author

Brigita Urbanc – Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, United States;

orcid.org/0000-0001-9159-7698; Email: bu25@drexel.edu; Fax: (215) 895-5934

Authors

Shuting Zhang – Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, United States

Brian Andrews – Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, United States

Reinhard Schweitzer-Stenner – Department of Chemistry, Drexel University, Philadelphia, Pennsylvania 19104, United States; orcid.org/0000-0001-5616-0722

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jpcb.0c08245>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants DMR-1707770 and MCB-1817650. Simulations reported in this work were run on hardware supported by Drexel's University Research Computing Facility.

■ REFERENCES

- (1) Brant, D. A.; Flory, P. J. J. The configuration of random polypeptide chains. II. Theory. *J. Am. Chem. Soc.* **1965**, *87*, 2791–2800.
- (2) Shi, Z. S.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163–240.
- (3) Avbelj, F.; Grdadolnik, S. G.; Grdadolnik, J.; Baldwin, R. L. Intrinsic backbone preferences are fully present in blocked amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1272–1277.
- (4) Grdadolnik, J.; Golic Grdadolnik, S.; Avbelj, F. Determination of conformational preferences of dipeptides using vibrational spectroscopy. *J. Phys. Chem. B* **2008**, *112*, 2712–2718.
- (5) Parchansky, V.; Kapitan, J.; Kaminsky, J.; Sebestik, J.; Bour, P. Ramachandran plot for alanine dipeptide as determined from Raman optical activity. *J. Phys. Chem. Lett.* **2013**, *4*, 2763–2768.
- (6) Toal, S. E.; Schweitzer-Stenner, R. Local order in the unfolded state: Conformational biases and nearest neighbor interactions. *Biomolecules* **2014**, *4*, 725–773.
- (7) Baxa, M. C.; Haddadian, E. J.; Jha, A. K.; Freed, K. F.; Sosnick, T. R. Context and force field dependence of the loss of protein backbone entropy upon folding using realistic denatured and native state ensembles. *J. Am. Chem. Soc.* **2012**, *134*, 15929–15936.
- (8) Schweitzer-Stenner, R. Conformational propensities and residual structures in unfolded peptides and proteins. *Mol. Biosyst.* **2012**, *8*, 122–133.
- (9) DeBartolo, J.; Jha, A.; Freed, K. F.; Sosnick, T. R. In *Ch.3: Local backbone preferences and nearest-neighbor effects in the unfolded and*

native states, in *Folding, misfolding and nonfolding of peptides and small proteins*; Schweitzer-Stenner, R., Ed.; John Wiley & Sons, Inc., 2012; pp 79–98.

(10) Schweitzer-Stenner, R.; Toal, S. E. Entropy reduction in unfolded peptides (and proteins) due to conformational preferences of amino acid residues. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22527–22536.

(11) Toal, S. E.; Kubatova, N.; Richter, C.; Linhard, V.; Schwalbe, H.; Schweitzer-Stenner, R. Randomizing the unfolded state of peptides (and proteins) by nearest neighbor interactions between unlike residues. *Chem. - Eur. J.* **2015**, *21*, 5173–5192.

(12) Schweitzer-Stenner, R.; Toal, S. E. Construction and comparison of the statistical coil states of unfolded and intrinsically disordered proteins from nearest-neighbor corrected conformational propensities of short peptides. *Mol. BioSyst.* **2016**, *12*, 3294–3306.

(13) Woutersen, S.; Hamm, P. Structure determination of trialanine in water using polarization sensitive two-dimensional vibrational spectroscopy. *J. Phys. Chem. B* **2000**, *104*, 11316–11320.

(14) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190–9195.

(15) Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. Structure and dynamics of the homologous series of alanine peptides: A joint molecular dynamics/NMR study. *J. Am. Chem. Soc.* **2007**, *129*, 1179–1189.

(16) Verbaro, D.; Ghosh, I.; Nau, W. M.; Schweitzer-Stenner, R. Discrepancies between Conformational Distributions of a Polyalanine Peptide in Solution Obtained from Molecular Dynamics Force Fields and Amide I' Band Profiles. *J. Phys. Chem. B* **2010**, *114*, 17201–17208.

(17) Hagarman, A.; Measey, T. J.; Mathieu, D.; Schwalbe, H.; Schweitzer-Stenner, R. Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I' band profiles and NMR scalar coupling constants. *J. Am. Chem. Soc.* **2010**, *132*, 540–551.

(18) Toal, S.; Meral, D.; Verbaro, D.; Urbanc, B.; Schweitzer-Stenner, R. The pH-independence of trialanine and the effects of termini blocking in short peptides: A combined vibrational, NMR, UVCD, and molecular dynamics study. *J. Phys. Chem. B* **2013**, *117*, 3689–3706.

(19) Swenson, C. A.; Formanek, R. Infrared study of poly-L-proline in aqueous solution. *J. Phys. Chem.* **1967**, *71*, 4073–4077.

(20) Rucker, A.; Pager, C.; Campbell, M.; Qualls, J.; Creamer, T. Host-guest scale of left-handed polyproline II helix formation. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 68–75.

(21) Ho, B. K.; Brasseur, R. The Ramachandran plots of glycine and pro-proline. *BMC Struct. Biol.* **2005**, *5*, 14.

(22) Tobias, D. J.; Brooks, C. L. Conformational equilibrium in the alanine dipeptide in the gas-phase and aqueous-solution: A comparison of theoretical results. *J. Phys. Chem.* **1992**, *96*, 3864–3870.

(23) Feig, M. Is alanine dipeptide a good model for representing the torsional preferences of protein backbones? *J. Chem. Theory Comput.* **2008**, *4*, 1555–1564.

(24) Cruz, V.; Ramos, J.; Martinez-Salazar, J. Water-mediated conformations of the alanine dipeptide as revealed by distributed umbrella sampling simulations, quantum mechanics based calculations, and experimental data. *J. Phys. Chem. B* **2011**, *115*, 4880–4886.

(25) Avbelj, F.; Baldwin, R. L. Role of backbone solvation and electrostatics in generating preferred peptide backbone conformations: Distributions of ϕ . *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 5742–5747.

(26) Garcia, A. E. Characterization of non- α helical conformations in Ala peptides. *Polymer* **2004**, *45*, 669–676.

(27) Gnanakaran, S.; Garcia, A. Helix-coil transition of alanine peptides in water: Force field dependence on the folded and unfolded structures. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 773–782.

(28) Kentsis, A.; Mezei, M.; Osman, R. Origin of the sequence-dependent polyproline II structure in unfolded peptides. *Proteins: Struct., Funct., Genet.* **2005**, *61*, 769–776.

(29) Tran, H.; Wang, X.; Pappu, R. Reconciling observations of sequence-specific conformational propensities with the generic polymeric behavior of denatured proteins. *Biochemistry* **2005**, *44*, 11369–11380.

(30) Meral, D.; Toal, S.; Schweitzer-Stenner, R.; Urbanc, B. Water-centered interpretation of intrinsic pPII propensities of amino acid residues: In vitro-driven molecular dynamics study. *J. Phys. Chem. B* **2015**, *119*, 13237–13251.

(31) Hinderaker, M. P.; Raines, R. T. An electronic effect on protein structure. *Protein Sci.* **2003**, *12*, 1188–1194.

(32) Law, P. B.; Daggett, V. The relationship between water bridges and the polyproline II conformation: a large-scale analysis of molecular dynamics simulations and crystal structures. *Protein Eng., Des. Sel.* **2010**, *23*, 27–33.

(33) Zhang, S.; Schweitzer-Stenner, R.; Urbanc, B. Do molecular dynamics force fields capture conformational dynamics of alanine in water? *J. Chem. Theory Comput.* **2020**, *16*, 510–527.

(34) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(35) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(36) Andrews, B.; Zhang, S.; Schweitzer-Stenner, R.; Urbanc, B. Glycine in water favors the polyproline II state. *Biomolecules* **2020**, *10*, 1121.

(37) Milorey, B.; Farrell, S.; Toal, S. E.; Schweitzer-Stenner, R. Demixing of water and ethanol causes conformational redistribution and gelation of the cationic GAG tripeptide. *Chem. Commun.* **2015**, *51*, 16498–16501.

(38) Farrell, S.; DiGuseppi, D.; Alvarez, N.; Schweitzer-Stenner, R. The interplay of aggregation, fibrillization and gelation of an unexpected low molecular weight gelator: glycylalanylglycine in ethanol/water. *Soft Matter* **2016**, *12*, 6096–6110.

(39) DiGuseppi, D. M.; Thursch, L.; Alvarez, N. J.; Schweitzer-Stenner, R. Exploring the gel phase of cationic glycylalanylglycine in ethanol/water. II. Spectroscopic, kinetic and thermodynamic studies. *J. Colloid Interface Sci.* **2020**, *573*, 123–134.

(40) DiGuseppi, D.; Milorey, B.; Lewis, G.; Kubatova, N.; Farrell, S.; Schwalbe, H.; Schweitzer-Stenner, R. Probing the conformation-dependent preferential binding of ethanol to cationic glycylalanylglycine in water/ethanol by vibrational and NMR spectroscopy. *J. Phys. Chem. B* **2017**, *121*, 5744–5758.

(41) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved peptide and protein torsional energetics with the OPLS-AA force field. *J. Chem. Theory Comput.* **2015**, *11*, 3499–3509.

(42) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71–73.

(43) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(44) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306–317.

(45) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(46) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(47) Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel

open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(48) Pall, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E. Tackling exascale software challenges in molecular dynamics simulations with GROMACS. Solving software challenges for exascale. In *Solving Software Challenges for Exascale*; Markidis, S., Laure, E., Eds.; Lecture Notes in Computer Science; EASC: Cham, Switzerland, 2014; Vol. 8759, pp 3–27.

(49) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(50) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(51) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.

(52) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(53) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, L.; Mackerell, A. D.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2009**, *31*, 671–690.

(54) Vanommeslaeghe, K.; MacKerell, A. D., Jr. Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.

(55) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr. Automation of the CHARMM general force field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.

(56) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D., Jr. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.

(57) Soteras Gutiérrez, I.; Lin, F.-Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L., III; MacKerell, A. D., Jr. Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand-protein interactions. *Bioorg. Med. Chem.* **2016**, *24*, 4812–4825.

(58) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force field benchmark of organic liquids: Density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theory Comput.* **2012**, *8*, 61–74.

(59) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(60) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(61) Schweitzer-Stenner, R. Distribution of conformations sampled by the central amino acid residue in tripeptides inferred from amide I' band profiles and NMR scalar coupling constants. *J. Phys. Chem. B* **2009**, *113*, 2922–2932.

(62) Hu, J.-S.; Bax, A. Determination of ϕ and χ_1 angles in proteins from ^{13}C - ^{13}C three-bond J couplings measured by three-dimensional heteronuclear NMR. How planar is the peptide bond? *J. Am. Chem. Soc.* **1997**, *119*, 6360–6368.

(63) Barfield, M.; Johnston, M. D. Solvent dependence of nuclear spin-spin coupling constants. *Chem. Rev.* **1973**, *73*, 53–73.

(64) Reppert, M.; Tokmakoff, A. Electrostatic frequency shifts in amide I vibrational spectra: Direct parameterization against experiment. *J. Chem. Phys.* **2013**, *138*, 134116.

(65) Wang, A. C.; Bax, A. Determination of the backbone dihedral angles ϕ in human ubiquitin from reparametrized empirical Karplus equations. *J. Am. Chem. Soc.* **1996**, *118*, 2483–2494.

(66) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(67) Wirmer, J.; Schwalbe, H. Angular dependence of $^1J(N_iC_{\alpha i})$ and $^2J(N_iC_{\alpha(i-1)})$ coupling constants measured in J-modulated HSQCs. *J. Biomol. NMR* **2002**, *23*, 47–55.

(68) DiGiuseppi, D.; Thursch, L.; Alvarez, N. J.; Schweitzer-Stenner, R. Exploring the thermal reversibility and tunability of a low molecular weight gelator using vibrational and electronic spectroscopy and rheology. *Soft Matter* **2019**, *15*, 3418–3431.

(69) Banerjee, S.; Ghosh, R.; Bagchi, B. Structural transformations, composition anomalies and a dramatic collapse of linear polymer chains in dilute ethanol-water mixtures. *J. Phys. Chem. B* **2012**, *116*, 3713–3722.